



HAL
open science

Robustness of iterated function systems of Lipschitz maps

Loïc Hervé, James Ledoux

► **To cite this version:**

Loïc Hervé, James Ledoux. Robustness of iterated function systems of Lipschitz maps. 2021. hal-03423198v1

HAL Id: hal-03423198

<https://hal.science/hal-03423198v1>

Preprint submitted on 9 Nov 2021 (v1), last revised 7 Sep 2022 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Robustness of iterated function systems of Lipschitz maps

Loïc HERVÉ and James LEDOUX *

Tuesday 9th November, 2021

Abstract

Let $\{X_n\}_{n \in \mathbb{N}}$ be a \mathbb{X} -valued iterated function system (IFS) of Lipschitz maps defined as: $X_0 \in \mathbb{X}$ and for $n \geq 1$, $X_n := F(X_{n-1}, \vartheta_n)$, where $\{\vartheta_n\}_{n \geq 1}$ are i.i.d.r.v. with common probability distribution ν and where $F(\cdot, \cdot)$ is Lipschitz continuous in the first variable. Under parametric perturbation of both F and ν , we are interested in the robustness of the V -geometrical ergodicity property of $\{X_n\}_{n \in \mathbb{N}}$, of its invariant probability measure and finally of the probability distribution of X_n . Specifically, we propose a pattern of assumptions for studying such robustness properties for an IFS. This pattern is implemented for the autoregressive processes with ARCH errors and for the roundoff IFS. Moreover, we provide a general set of assumptions, which cover the classical Feller-type hypotheses, for an IFS to be a V -geometrical ergodic process, together with an accurate bound for the rate of convergence.

AMS subject classification : 60J05, 47B07, 62F99

Keywords : Markov chain, Geometric ergodicity, Perturbation, Non-linear stochastic recursion, Autoregressive process.

1 Introduction

Let (\mathbb{X}, d) be a Polish space equipped with its Borel σ -algebra \mathcal{X} . The random variables (r.v.) are assumed to be defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and “i.i.d.” is the short-hand for “independent and identically distributed”. Throughout the paper we are concerned with iterated function systems of Lipschitz maps according to the following well-known definition.

Definition 1.1 (IFS of Lipschitz maps) *Let $(\mathbb{V}, \mathcal{V})$ be a measurable space, and let $\{\vartheta_n\}_{n \geq 1}$ be a sequence of \mathbb{V} -valued i.i.d. random variables, with common distribution denoted by ν . Let X_0 be a \mathbb{X} -valued r.v. which is assumed to be independent of the sequence $\{\vartheta_n\}_{n \geq 1}$. Finally*

*Université Univ Rennes, INSA Rennes, CNRS, IRMAR-UMR 6625, F-35000 Rennes, France. Loic.Herve@insa-rennes.fr, James.Ledoux@insa-rennes.fr

let $F : (\mathbb{X} \times \mathbb{V}, \mathcal{X} \otimes \mathcal{V}) \rightarrow (\mathbb{X}, \mathcal{X})$ be jointly measurable and Lipschitz continuous in the first variable. The associated iterated function system (IFS) is the sequence of random variables $\{X_n\}_{n \in \mathbb{N}}$ which, starting from X_0 , is recursively defined as follows:

$$\forall n \geq 1, \quad X_n := F(X_{n-1}, \vartheta_n). \quad (1)$$

Let $x_0 \in \mathbb{X}$ be fixed. For any $a \in [0, +\infty)$, we set $V_a(x) := (1 + d(x, x_0))^a$, and we denote by $(\mathcal{B}_a, |\cdot|_a)$ the weighted-supremum Banach space \mathcal{B}_{V_a} associated with $V_a(\cdot)$, that is

$$\mathcal{B}_a := \left\{ f : \mathbb{X} \rightarrow \mathbb{C} \text{ measurable such that } |f|_a := \sup_{x \in \mathbb{X}} \frac{|f(x)|}{V_a(x)} < \infty \right\}. \quad (2)$$

Note that $(\mathcal{B}_0, |\cdot|_0)$ corresponds to the usual Banach space of complex-valued bounded measurable functions on \mathbb{X} equipped with the supremum norm. The total variation distance between two probability distributions μ_0 and μ_1 on \mathbb{X} is defined by

$$\|\mu_0 - \mu_1\|_{TV} := \sup_{|f|_0 \leq 1} |\mu_0(f) - \mu_1(f)|$$

where $\mu_i(f) := \int_{\mathbb{X}} f(x) d\mu_i(x)$. Let $\{X_n\}_{n \in \mathbb{N}}$ be an IFS of Lipschitz maps. This is a Markov chain on \mathbb{X} with transition kernel P given by:

$$\forall x \in \mathbb{X}, \forall A \in \mathcal{X}, \quad P(x, A) = \mathbb{E}[1_A(F(x, \vartheta_1))] = \int_{\mathbb{V}} 1_A(F(x, v)) d\nu(v). \quad (3)$$

Recall that $\{X_n\}_{n \in \mathbb{N}}$ is V_a -geometrically ergodic if P has an invariant probability measure π such that $\pi(V_a) < \infty$ and if there exists $\rho_a \in (0, 1)$ and $C_a \in (0, +\infty)$ such that

$$\forall n \geq 1, \forall f \in \mathcal{B}_a, \quad |P^n f - \pi(f)1_{\mathbb{X}}|_a \leq C_a \rho_a^n |f|_a. \quad (4)$$

The V_a -geometric ergodicity of IFS has been extensively studied (see e.g. [DMPS18, GHL11, MT93, Als03, Wu04] and references therein). The common starting point in most of these works is that P satisfies the so-called drift condition under the well-known moment/contractive Conditions (\mathcal{C}_a) below (e.g. see [Duf97]), for which we introduce the following notations. If $\psi : (\mathbb{X}, d) \rightarrow (\mathbb{X}, d)$ is a Lipschitz continuous function, we define

$$L(\psi) := \sup \left\{ \frac{d(\psi x, \psi y)}{d(x, y)}, (x, y) \in \mathbb{X}^2, x \neq y \right\}. \quad (5)$$

For all $v \in \mathbb{V}$, set $L_F(v) := L(F(\cdot, v))$ to simplify. Note that the assumption in Definition 1.1 according that F is Lipschitz continuous in the first variable reads as $L(v) < \infty$ for any v . Then, for every $a \in [1, +\infty)$, Conditions (\mathcal{C}_a) write as follows:

Conditions (\mathcal{C}_a) . *The function $F(\cdot, \cdot)$ and the sequence $\{\vartheta_n\}_{n \geq 1}$ satisfy:*

$$\mathbb{E} [d(x_0, F(x_0, \vartheta_1))^a] < \infty \quad (6a)$$

$$\mathbb{E} [L_F(\vartheta_1)^a] < 1. \quad (6b)$$

The condition $a \geq 1$ in Conditions (\mathcal{C}_a) is just a technical assumption (for instance useful for applying Hölder inequality): in fact Conditions (\mathcal{C}_a) can be considered with $a > 0$ up to change the initial distance d by d^α for some $\alpha \in (0, 1)$. Let us specify Conditions (\mathcal{C}_a) for the so-called vector autoregressive models.

Example 1.1 (Vector autoregressive models) Assume that $\mathbb{X} := \mathbb{R}^q$ for some $q \geq 1$. Let $\|\cdot\|$ be any norm of \mathbb{R}^q , and define $d(x, y) := \|x - y\|$ the associated distance on \mathbb{R}^q . Consider $V_a(x) := (1 + \|x\|)^a$ with $a \in [1, +\infty)$ (here $x_0 := 0$), and let $\{X_n\}_{n \in \mathbb{N}}$ be the following IFS

$$X_0 \in \mathbb{R}^q, \quad \forall n \geq 1, \quad X_n := AX_{n-1} + \vartheta_n, \quad (7)$$

Here $F(x, v) := Ax + v$ where $A = (a_{ij})$ is a fixed real $q \times q$ -matrix. This is the well-known vector autoregressive model (VAR). We have $L_F(v) = \|A\|$ where $\|A\|$ denotes the induced norm of A corresponding to $\|\cdot\|$, and $d(0, F(0, v)) = \|v\|$. Consequently, Conditions (C_a) hold for $a \in [1, +\infty)$ provided that we have:

$$\mathbb{E}[\|\vartheta_1\|^a] < \infty \quad \text{and} \quad \|A\| < 1. \quad (8)$$

Moreover, if ϑ_1 has a probability density function on \mathbb{R}^q , then P is V_a -geometrically ergodic (see [MT93]).

The aim of this work is to use the results of [FHL13, HL14a, RS18] to investigate: first, the robustness of the V_a -geometrical ergodicity property (4); second, the sensitivity of the probability distribution (p.d.) of X_n and of the stationary distribution π , with respect to parametric variations of both the function F and the p.d. of the noise r.v. ϑ_n in (1). Thus, let us introduce the following definition.

Definition 1.2 (Parametric perturbation of IFS) The process $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ is a \mathbb{X} -valued IFS of Lipschitz maps given by

$$X_0^{(\theta)} \in \mathbb{X}, \quad \forall n \geq 1, \quad X_n^{(\theta)} := F_\xi(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}). \quad (9)$$

where $F_\xi : (\mathbb{X} \times \mathbb{V}, \mathcal{X} \otimes \mathcal{V}) \rightarrow (\mathbb{X}, \mathcal{X})$ and the sequence $\{\vartheta_n^{(\gamma)}\}_{n \geq 1}$ of \mathbb{V} -valued r.v. both satisfy the assumptions of Definition 1.1. The common parametric p.d. of $\{\vartheta_n^{(\gamma)}\}_{n \geq 1}$ is denoted by ν_γ , and the parameter $\theta := (\xi, \gamma)$ is Θ -valued, where Θ is assumed to be a subset of some metric space. The transition kernel of the Markov chain $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ is denoted by P_θ (see (3)), and μ_θ is the p.d. of $X_0^{(\theta)}$.

The Markov chain $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ must be thought of as a perturbed model of some ideal model $\{X_n^{(\theta_0)}\}_{n \in \mathbb{N}}$. Next, let us introduce the following assumptions:

(H₁) there exists $a \geq 1$ such that P_{θ_0} is V_a -geometrically ergodic with stationary distribution denoted by π_{θ_0} (so that Inequality (4) holds true for some $\rho_a \in (0, 1)$ and $C_a > 0$).

(H₂) $M_a := \sup_{\theta \in \Theta} \mathbb{E} \left[d \left(x_0, F_\xi(x_0, \vartheta_1^{(\gamma)}) \right)^a \right]^{1/a} < \infty$.

(H₃) $\kappa_a := \sup_{\theta \in \Theta} \mathbb{E} \left[L_{F_\xi}(\vartheta_1^{(\gamma)})^a \right]^{1/a} < 1$.

(H₄) $\Delta_\theta := \|P_\theta - P_{\theta_0}\|_{0,a} \xrightarrow{\theta \rightarrow \theta_0} 0$, where $\|P_\theta - P_{\theta_0}\|_{0,a} := \sup_{f \in \mathcal{B}_0, |f|_0 \leq 1} |P_\theta f - P_{\theta_0} f|_a$.

Assumption (H₁) is the natural starting point for our perturbation issues. Note that Assumptions (H₂)–(H₃) are nothing else but the uniform version (wrt θ) of Conditions (C_a). As a by-product it follows from (H₂)–(H₃) that each P_θ satisfies a drift condition with respect to the function V_a . More precisely, let $\kappa \in (\kappa_a, 1)$. Then the following uniform in $\theta \in \Theta$ drift condition holds true (see Appendix A):

$$\forall \theta \in \Theta, \quad P_\theta V_a \leq \delta_a V_a + K_a \quad \text{with} \quad \delta_a := \kappa^a \quad \text{and} \quad K_a := \frac{(1 + \kappa_a + M_a)^a (1 + M_a)^a}{(\kappa - \kappa_a)^a}. \quad (10)$$

This implies that, for every $\theta \in \Theta$, P_θ admits an invariant probability measure denoted by π_θ . However, first the continuity of $\theta \mapsto \pi_\theta$ with respect to the total variation distance, second the robustness of the V_a -geometrical ergodicity property under Assumption (H₁), are quite more difficult issues. In our context of parametric perturbation of IFS, the results of [FHL13, HL14a, RS18] can be used to obtain the following statements.

Theorem 1.1 *Under the assumptions (H₁)–(H₄), the following properties hold true.*

(P₁) *There exists an open neighbourhood \mathcal{V}_{θ_0} of θ_0 such that, for every $\theta \in \mathcal{V}_{\theta_0}$, P_θ is V_a -geometrically ergodic. More precisely there exists a positive constant R_a such that*

$$\forall \theta \in \mathcal{V}_{\theta_0}, \quad \forall n \geq 1, \quad \forall f \in \mathcal{B}_a, \quad |P_\theta^n f - \pi_\theta(f)1_{\mathbb{X}}|_a \leq R_a \rho_a^n |f|_a.$$

(P₂) $\lim_{\theta \rightarrow \theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{TV} = 0$. *More precisely:*

$$\forall \theta \in \Theta, \quad \|\pi_\theta - \pi_{\theta_0}\|_{TV} \leq \frac{e K_a D_a^{[\ln(\Delta_\theta^{-1})]^{-1}}}{(1 - \delta_a)(1 - \rho_a)} \Delta_\theta \ln(\Delta_\theta^{-1}) \quad (11)$$

provided that $\Delta_\theta \in (0, \exp(-1))$, where the constants ρ_a , C_a , δ_a and K_a are given in (4) and (10), and $D_a = 2C_a(K_a + 1)$.

(P₃) *We have for every $n \geq 1$ and for every $\theta \in \Theta$*

$$\|\mu_\theta P_\theta^n - \mu_{\theta_0} P_{\theta_0}^n\|_{TV} \leq C_a \rho_a^n \sup_{|f| \leq V} |\mu_\theta(f) - \mu_{\theta_0}(f)| + \frac{e G_a D_a^{[\ln(\Delta_\theta^{-1})]^{-1}}}{1 - \delta_a} \Delta_\theta \ln(\Delta_\theta^{-1})$$

provided that $\Delta_\theta \in (0, \exp(-1))$, with $G_a := \max\{K_a/(1 - \delta_a), \mu_{\theta_0}(V_a)\}$. In particular, if $X_0^{(\theta)}$ and $X_0^{(\theta_0)}$ have the same p.d., that is $\mu_\theta = \mu_{\theta_0}$, then: $\lim_{\theta \rightarrow \theta_0} \|\mu P_\theta^n - \mu P_{\theta_0}^n\|_{TV} = 0$.

In the general framework of V -geometrically ergodic Markov chains, Property (P₁) and the first statement in (P₂) are proved in [FHL13, Th. 1] by using the Keller-Liverani perturbation theorem [KL99]¹. Inequality (11) in (P₂) follows from [HL14a, Prop. 2.1] or [RS18, (3.19)]. The formulation [RS18, (3.19)] has been preferred to that in [HL14a, Prop. 2.1] in connection with Property (P₃). Property (P₃) is proved in [RS18, Th. 3.2] by using the Wasserstein

¹the real-valued parameter ε in [FHL13] may be replaced by the Θ -valued parameter θ

distance associated with a suitable distance on \mathbb{X} defined from the Lyapunov function V , as introduced in [HM11]. In the IFS context, Theorem 1.1 has been already illustrated in [FHL13, RS18] for real-valued AR(1) process in which the contracting coefficient is perturbed and the law of the noise is fixed. The goal of this work is to present further applications when both the function F and the p.d. ν of the noise in Definition 1.1 are perturbed, and to show that the weak continuity Assumption (H₄) remains well adapted when the p.d. of the noise varies. This last fact is highlighted by the following first simple application, where only the p.d. of the noise is perturbed.

Example 1.2 (IFS with perturbed noise) *Consider the generic IFS introduced in Definition 1.1 with noise probability distribution ν_0 . Its transition kernel P_{ν_0} is given by*

$$\forall f \in \mathcal{B}_0, \quad (P_{\nu_0}f)(x) = \int f(F(x, y)) d\nu_0(y).$$

Let us consider the specific perturbation scheme

$$X_0^{(\theta)} \in \mathbb{X}, \quad \forall n \geq 1, \quad X_n^{(\theta)} := F(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}).$$

where $\{\vartheta_n^{(\gamma)}\}_{n \geq 1}$ is a sequence of \mathbb{V} -valued i.i.d.r.v, with common parametric probability distribution denoted by ν_γ . That is, we consider an IFS with perturbed noise but fixed function F (e.g. the matrix A is fixed in the VAR model introduced in Example 1.1). For any $f \in \mathcal{B}_0$ such that $|f|_0 \leq 1$, we have

$$\forall x \in \mathbb{X}, \quad |(P_\nu f)(x) - (P_{\nu_0}f)(x)| \leq \|\nu - \nu_0\|_{TV}. \quad (12)$$

It follows that

$$\|P_\nu - P_{\nu_0}\|_{0,a} \leq \|P_\nu - P_{\nu_0}\|_{0,0} := \sup_{f \in \mathcal{B}_0, |f|_0 \leq 1} |P_\nu f - P_{\nu_0}f|_0 \leq \|\nu - \nu_0\|_{TV}.$$

Hence (H₄) is satisfied provided that $\lim \|\nu - \nu_0\|_{TV} = 0$.

In Section 2, a second application of Theorem 1.1, which again illustrates the interest of (H₄), is provided for real-valued AR(1) with ARCH(1) errors models given by $X_n := \alpha X_{n-1} + \sigma(X_{n-1})\vartheta_n$, for which the data α , $\sigma(\cdot)$ as well as the p.d. of the noise ϑ_1 are all together perturbed. A third application is presented in Section 3 in the framework of roundoff errors as introduced in [RRS98, BRR01]. In this case, the perturbed transition kernel is the form $\tilde{P}(x, A) := P(x, h^{-1}(A))$, where P is the transition kernel of a fixed IFS and where $h : \mathbb{X} \rightarrow \mathbb{X}$ is close to the identity map. The weak continuity assumption (H₄) is still proved to be well adapted as illustrated in Proposition 3.1 for VAR models defined in Example 1.1. Note that the function F in (9) is fixed in Example 1.2, so that we did not have to divide by $V(x)$ to prove (H₄). Indeed Inequality $\|P_\nu - P_{\nu_0}\|_{0,0} \leq \|\nu - \nu_0\|_{TV}$ is directly obtained and it automatically gives (H₄). When F in (9) is also perturbed as in Sections 2-3, the division by $V(x)$ in the definition $\|\cdot\|_{0,a}$ must be done to investigate (H₄). Finally we propose in Section 4 a new approach to prove the V_a -geometrical ergodicity of IFS of Lipschitz maps under Conditions (C_a), together with a bound on the spectral gap of P (i.e. the infimum bound of the positive real numbers ρ_a satisfying (4)).

Let us mention that this paper does not address the statistical issues when the model is misspecified. Indeed, we do not study the convergence properties of estimators of the parameters of the Markov model when the data are generated under the “wrong” model and the size n of the data grows is large (e.g. see [GW98, DM12] in the Markov context).

2 Robustness of AR(1) with ARCH(1) errors

Let $\{X_n\}_{n \in \mathbb{N}}$ be the following real-valued IFS which is usually called AR(1) with ARCH(1) errors process: X_0 is a given real-valued r.v. and

$$\forall n \geq 1, \quad X_n := \alpha X_{n-1} + \sigma(X_{n-1})\vartheta_n, \quad (13)$$

where $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ is defined by $\sigma(x) := (\beta + \lambda x^2)^{1/2}$ with constants $\alpha \in \mathbb{R}, \beta > 0, \lambda > 0$. It can be easily seen that the associated Markov kernel is of the form $P(x, dy) = p(x, y)dy$, provided that the probability distribution of ϑ_1 has a density function $\nu(\cdot)$ (see (15)). In this part, consider the following perturbed AR(1)-ARCH(1) real-valued $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ defined by

$$\forall n \geq 1, \quad X_n^{(\theta)} := F_\xi(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}) \quad (14)$$

where $F_\xi(x, v) = \alpha x + v(\beta + \lambda x^2)^{1/2}$ and $\{\vartheta_n^{(\gamma)}\}_{n \in \mathbb{N}}$ has common probability density function (p.d.f.) ν_γ . In such a context, we have $\theta = (\xi, \gamma)$ with $\xi := (\alpha, \beta, \lambda) \in \mathbb{R} \times (0, +\infty)^2$ and $\gamma \in \Gamma$ where Γ is some metric space (typically $\Gamma \subset \mathbb{R}$). Thus Θ is a subset of $\mathbb{R} \times (0, +\infty)^2 \times \Gamma$. Here $d(x; x_0) = |x - x_0|$ and $x_0 := 0$ so that $V_a(x) = (1 + |x|)^a$. The Markov kernel P_θ of $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ is given by $P_\theta(x, A) := \int_{\mathbb{R}} 1_A(y) p_\theta(x, y) dy$ ($A \in \mathcal{X}$) with

$$p_\theta(x, y) := (\beta + \lambda x^2)^{-1/2} \nu_\gamma \left(\frac{y - \alpha x}{(\beta + \lambda x^2)^{1/2}} \right). \quad (15)$$

Next, let us report the following observations with respect to basic quantities required in Assumptions (H₂) and (H₃).

1. It can be checked (see Lemma B.1) that

$$L_{F_\xi}(\vartheta_1) = \max(|\alpha - \sqrt{\lambda}\vartheta_1|; |\alpha + \sqrt{\lambda}\vartheta_1|) \quad (16)$$

Hence, the real number κ_a in (H₃) is

$$\kappa_a = \sup_{\theta \in \Theta} \left(\int_{\mathbb{R}} \max(|\alpha - \sqrt{\lambda}v|; |\alpha + \sqrt{\lambda}v|)^a \nu_\gamma(v) dv \right)^{1/a}. \quad (17)$$

2. The real number M_a in (H₂) is given by

$$M_a := \sup_{\theta \in \Theta} \sqrt{\beta} \mathbb{E} \left[|\vartheta_1^{(\gamma)}|^a \right]^{1/a} = \sup_{\theta \in \Theta} \sqrt{\beta} \left(\int_{\mathbb{R}} |v|^a \nu_\gamma(v) dv \right)^{1/a}. \quad (18)$$

Note that, if β lies in a compact set, then $M_a < \infty$ under the following uniform moment condition for the probability distribution of $\vartheta_1^{(\gamma)}$: $\sup_{\gamma \in \Gamma} \int_{\mathbb{R}} |v|^a \nu_\gamma(v) dv < \infty$.

Let us formulate the assumptions under which the conclusions of Theorem 1.1 hold true for $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$. Let $\theta_0 = (\alpha_0, \beta_0, \lambda_0, \gamma_0) \in \mathring{\Theta}$. We denote by $\mathbb{L}^1(\mathbb{R})$ the usual Lebesgue space and by $\|\cdot\|_{\mathbb{L}^1(\mathbb{R})}$ its norm.

(H'_{123}) There exists $a \geq 1$ such that

(a) For every $r > 0$, the function

$$y \mapsto g_{\theta_0, r}(y) := \inf_{x \in [-r, r]} p_{\theta_0}(x, y) = \inf_{x \in [-r, r]} (\beta_0 + \lambda_0 x^2)^{-1/2} \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}} \right)$$

is assumed to be positive on a non-negligible subset of $[-r, r]$ (with respect to Lebesgue's measure).

(b) $M_a < \infty$, where M_a is given in (18).

(c) $\kappa_a < 1$, where κ_a is given in (17).

(H'_4) $\lim_{\gamma \rightarrow \gamma_0} \|\nu_\gamma - \nu_{\gamma_0}\|_{\mathbb{L}^1(\mathbb{R})} = 0$.

Proposition 2.1 *Under Conditions (H'_{123})-(H'_4) for the AR(1)-ARCH(1) process given in (14), Properties (P_1)-(P_2)-(P_3) of Theorem 1.1 hold true.*

Proof. Let $\theta_0 = (\alpha_0, \beta_0, \lambda_0, \gamma_0) \in \mathring{\Theta}$ and let $a \geq 1$ provided by (H'_{123}). As already discussed Conditions (H'_{123})-(b)-(c) implies that Assumptions (H_2) and (H_3) of Theorem 1.1 hold in our context of AR(1)-ARCH(1) process. Moreover, use (43) to state that there exist $\delta_a < 1$, $K_a > 0$ and $r_a > 0$ such that

$$P_{\theta_0} V_a \leq \delta_a V_a + K_a 1_{[-r_a, r_a]}.$$

Next, Condition (H'_{123})-(a) ensures that

$$\forall x \in [-r_a, r_a], \forall A \in \mathcal{X}, \quad P_{\theta_0}(x, A) \geq \varphi_{r_a, \theta_0}(A)$$

with the positive measure $\varphi_{r, \theta_0}(dy) = g_{\theta_0, r}(y) dy$. In others words, $S = [-r, r]$ is a small set for P_{θ_0} . Moreover $\varphi_{r, \theta_0}(S) > 0$ from (H'_{123})-(a). Then Assumption (H_1) holds true, see [MT93][Bax05, Th 1.1]. The proof will be complete using the following lemma which asserts that Assumption (H_4) holds true under Condition (H'_4). \square

Lemma 2.1 *If $\lim_{\gamma \rightarrow \gamma_0} \|\nu_\gamma - \nu_{\gamma_0}\|_{\mathbb{L}^1(\mathbb{R})} = 0$ then*

$$\lim_{\theta \rightarrow \theta_0} \|P_\theta - P_{\theta_0}\|_{0, a} = 0.$$

Proof. Let $f \in \mathcal{B}_0$ such that $|f|_0 \leq 1$. We have

$$\begin{aligned} \forall x \in \mathbb{X}, \quad \frac{|(P_\theta f)(x) - (P_{\theta_0} f)(x)|}{V_a(x)} &= \frac{|\int_{\mathbb{R}} (p_\theta(x, y) - p_{\theta_0}(x, y)) f(y) dy|}{V_a(x)} \\ &\leq \frac{\int_{\mathbb{R}} |p_\theta(x, y) - p_{\theta_0}(x, y)| dy}{V_a(x)}. \end{aligned}$$

Let $\varepsilon > 0$. Since the last term is bounded from above by $2/V_a(x)$ and $\lim_{x \rightarrow +\infty} V_a(x) = +\infty$, there exists $B > 0$ such that

$$|x| > B \implies \forall \theta \in \Theta, \frac{|(P_\theta f)(x) - (P_{\theta_0})f(x)|}{V_a(x)} < \frac{\varepsilon}{2}. \quad (19)$$

It follows that the conclusion of the lemma holds true provided that, under the condition $\lim_{\gamma \rightarrow \gamma_0} \|\nu_\gamma - \nu_{\gamma_0}\|_{\mathbb{L}^1(\mathbb{R})} = 0$, we have

$$\forall A > 0, \lim_{\theta \rightarrow \theta_0} \sup_{|x| \leq A} \frac{\int_{\mathbb{R}} |p_\theta(x, y) - p_{\theta_0}(x, y)| dy}{V_a(x)} = 0. \quad (20)$$

Indeed, (19) and (20) with $A = B$ ensure that $\|P_\theta - P_{\theta_0}\|_{0,a} < \varepsilon$ when θ is sufficiently close to θ_0 . Let us prove (20). It follows from (15) that

$$\begin{aligned} & \int_{\mathbb{R}} |p_\theta(x, y) - p_{\theta_0}(x, y)| dy \\ &= \int_{\mathbb{R}} \left| (\beta + \lambda x^2)^{-1/2} \nu_\gamma \left(\frac{y - \alpha x}{(\beta + \lambda x^2)^{1/2}} \right) - (\beta_0 + \lambda_0 x^2)^{-1/2} \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}} \right) \right| dy \\ &\leq \int_{\mathbb{R}} (\beta + \lambda x^2)^{-1/2} \left| \nu_\gamma \left(\frac{y - \alpha x}{(\beta + \lambda x^2)^{1/2}} \right) - \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}} \right) \right| dy \end{aligned} \quad (21)$$

$$+ \int_{\mathbb{R}} \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}} \right) \left| (\beta + \lambda x^2)^{-1/2} - (\beta_0 + \lambda_0 x^2)^{-1/2} \right| dy. \quad (22)$$

First, using the change of variables $z = (y - \alpha x)/(\beta + \lambda x^2)^{1/2}$ in the integral (21) and the triangle inequality we obtain

$$\begin{aligned} (21) &= \int_{\mathbb{R}} \left| \nu_\gamma(z) - \nu_{\gamma_0} \left(\left(\frac{\beta + \lambda x^2}{\beta_0 + \lambda_0 x^2} \right)^{1/2} z + x \frac{\alpha - \alpha_0}{(\beta_0 + \lambda_0 x^2)^{1/2}} \right) \right| dz \\ &\leq \int_{\mathbb{R}} |\nu_\gamma(z) - \nu_{\gamma_0}(z)| dz + \int_{\mathbb{R}} |\nu_{\gamma_0}(z) - \nu_{\gamma_0}(b_\theta(x)z + a_\theta(x))| dz \end{aligned} \quad (23)$$

where

$$b_\theta(x) := \left(\frac{\beta + \lambda x^2}{\beta_0 + \lambda_0 x^2} \right)^{1/2} \quad \text{and} \quad a_\theta(x) := x \frac{\alpha - \alpha_0}{(\beta_0 + \lambda_0 x^2)^{1/2}}.$$

The first integral in (23) does not depend on x and is nothing else than $\|\nu_\gamma - \nu_{\gamma_0}\|_{\mathbb{L}^1(\mathbb{R})}$ which converge to 0 when $\gamma \rightarrow \gamma_0$ from the assumption. Now let $A > 0$ be fixed. It follows from Lemma B.2 that $\lim_{\gamma \rightarrow \gamma_0} \sup_{|x| \leq A} |b_\theta(x) - 1| = 0$ and $\lim_{\gamma \rightarrow \gamma_0} \sup_{|x| \leq A} a_\theta(x) = 0$. Then under Condition (H₄), Lemma B.3 allows us to conclude that the second integral in (23) is such that

$$\lim_{\gamma \rightarrow \gamma_0} \sup_{|x| \leq A} \int_{\mathbb{R}} |\nu_{\gamma_0}(z) - \nu_{\gamma_0}(b_\theta(x)z + a_\theta(x))| dz = 0.$$

Second, let us consider the integral (22). We must show that the supremum of this integral on $x \in [-A, A]$ converges to 0 when $\gamma \rightarrow \gamma_0$. We obtain for any $x \in \mathbb{R}$ such that $|x| \leq A$.

$$\begin{aligned}
(22) &= \left| (\beta + \lambda x^2)^{-1/2} - (\beta_0 + \lambda_0 x^2)^{-1/2} \right| \times \int_{\mathbb{R}} \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}} \right) dy \\
&= (\beta_0 + \lambda_0 x^2)^{-1/2} \left| \frac{1}{b_{\theta}(x)} - 1 \right| \times \int_{\mathbb{R}} \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}} \right) dy \\
&= \left| \frac{1 - b_{\theta}(x)}{b_{\theta}(x)} \right| \int_{\mathbb{R}} \nu_{\gamma_0}(z) dz \quad (\text{change of variables } z = (y - \alpha_0 x)/(\beta_0 + \lambda_0 x^2)^{1/2}) \\
&\leq \frac{|1 - b_{\theta}(x)|}{b_A}
\end{aligned}$$

with $b_A := (\beta/(\beta_0 + \lambda_0 A^2))^{1/2} \leq \min_{|x| \leq A} b_{\theta}(x)$ and $\int_{\mathbb{R}} \nu_{\gamma_0}(z) dz = 1$. We know that $\lim_{\gamma \rightarrow \gamma_0} \sup_{|x| \leq A} |b_{\theta}(x) - 1| = 0$ from Lemma B.2, so that the expected convergence holds. \square

Remark 2.1 *If the p.d.f. ν_{γ_0} of the noise of the unperturbed process is continuous on \mathbb{R} , then Condition (H₁₂₃)-(a) (stated to prove (H₁)) can be omitted. Actually, under the condition $\int_{\mathbb{R}} L_{F_{\xi_0}}(v)^a \nu_{\gamma_0}(v) dv < 1$ which is contained in (H₁₂₃)-(c), Assumption (H₁) holds with any real number ρ_a (and the associated constant C_a) such that*

$$\left(\int_{\mathbb{R}} L_{F_{\xi_0}}(v)^a \nu_{\gamma_0}(v) dv \right)^{1/a} < \rho_a < 1.$$

Indeed the kernel $p_{\theta_0}(x, y)$ given by (15) is continuous, so that Remark 4.3 and Proposition 4.2 ensure that, under Conditions (H₁₂₃)-(b) and (H₁₂₃)-(c), P_{θ_0} fulfils (4) for any ρ_a satisfying the above condition. In other words, only Conditions (H₁₂₃)-(b) and (H₁₂₃)-(c) with $\Theta = \{\theta_0\}$ are required to obtain (H₁) (and only (H₁)) if the p.d.f. ν_{γ_0} is continuous on \mathbb{R} .

Remark 2.2 *It is well-known from Scheffé's lemma [Sch47] that the almost everywhere point-wise convergence of the p.d.f. ν_{γ} to the p.d.f. ν_{γ_0} when $\gamma \rightarrow \gamma_0$ asserts the $\mathbb{L}^1(\mathbb{R})$ -convergence required in (H₄) (see Lemma 2.1).*

3 Robustness of roundoff IFS models

As introduced in [RRS98, BRR01], the effect of roundoff errors in computer simulations using a Markov chain with transition kernel P yields to consider a modified Markov chain with perturbed transition kernel of the form $\tilde{P}(x, A) := P(x, h^{-1}(A))$, where $h : \mathbb{X} \rightarrow \mathbb{X}$ is such that $h(x)$ is close to x . In our context let us consider an \mathbb{X} -valued IFS as defined in Definition 1.1 where the common p.d. of ϑ_n is denoted by ν . Let $(h_{\theta})_{\theta \in \Theta}$ be a family of functions on \mathbb{X} such that $h_{\theta} \rightarrow id$ when $\theta \rightarrow \theta_0$ in a sense to be specified later, where id denotes the identity map on \mathbb{X} and Θ is a subset of a metric space. Then the associated roundoff IFS $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ is defined by

$$X_0^{(\theta)} \in \mathbb{X}, \quad \forall n \geq 1, X_n^{(\theta)} = F_{\theta}(X_{n-1}^{(\theta)}, \vartheta_n)$$

where $F_\theta(x, v) := h_\theta(F(x, v))$ and $F_{\theta_0}(x, v) = id(F(x, v)) = F(x, v)$. The perturbed Markov kernels associated with $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ (or $(h_\theta)_{\theta \in \Theta}$) are given by

$$\forall f \in \mathcal{B}_0, \forall x \in \mathbb{X}, \quad (P_\theta f)(x) = P(f \circ h_\theta)(x) = \int_{\mathbb{V}} f(h_\theta \circ \Psi(x, v)) d\nu(v). \quad (24)$$

When the unperturbed Markov kernel P_{θ_0} is assumed to be V -geometrically ergodic, the first natural question is to know whether P_θ remains V -geometrically ergodic for θ close to θ_0 . The simplest way used in [RRS98] to study this question is to assume that $h_\theta \rightarrow id$ uniformly on \mathbb{R}^q when $\theta \rightarrow \theta_0$ (i.e. $\forall x \in \mathbb{R}^q, \|h_\theta(x) - x\| \leq \varepsilon(\theta)$ with $\lim_{\theta \rightarrow \theta_0} \varepsilon(\theta) = 0$). However, as mentioned in [BRR01], this assumption is too restrictive in practice since the roundoff errors for some $x \in \mathbb{R}^q$ is obviously proportional to x . The authors in [BRR01] introduced the following weaker assumption $\|h_\theta(x) - x\| \leq \varepsilon(\theta)\|x\|$ with $\lim_{\theta \rightarrow \theta_0} \varepsilon(\theta) = 0$, and proved that the V -geometric ergodicity property is stable for the roundoff Markov kernels under some mild assumptions on the function V . Below, as a by-product of Theorem 1.1, we find again this result in the specific instance of the roundoff process associated with a VAR(1) model $\{X_n\}_{n \in \mathbb{N}}$, but more importantly the sensitivity of the probability distribution of $X_n^{(\theta)}$ and of the stationary distribution of $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ when $\theta \rightarrow \theta_0$ is addressed too. These two issues are not investigated in [BRR01].

Let $\{X_n\}_{n \in \mathbb{N}}$ be a \mathbb{R}^q -valued VAR model as defined in Example 1.1. To simplify we assume that, for some $p \geq 1$, Θ is an open subset of \mathbb{R}^p containing $\theta_0 := 0$ (the null vector of \mathbb{R}^p), and we consider a family $(h_\theta)_{\theta \in \Theta}$ of functions on $\mathbb{X} := \mathbb{R}^q$ such that $h_0 = id$, where id denotes the identity map on \mathbb{R}^q . Thus the roundoff process $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ associated with $F_\theta(x, y) := h_\theta(Ax + v)$ is the Markov chain with transition kernel P_θ (see (24))

$$\forall f \in \mathcal{B}_0, \forall x \in \mathbb{X}, \quad (P_\theta f)(x) = \int_{\mathbb{R}^q} f(h_\theta(Ax + v)) d\nu(v). \quad (25)$$

If $g : \mathbb{R}^q \rightarrow \mathbb{R}^q$ is differentiable and if $z \in \mathbb{R}^q$, we denote by $\nabla g(z)$ the Jacobian matrix of g at z , and we set $\|\nabla g\|_\infty := \sup_{z \in \mathbb{R}^q} \|\nabla g(z)\|$, where $\|\cdot\|$ here denotes the induced matrix-norm of Example 1.1. To simplify the norms chosen on \mathbb{R}^q and \mathbb{R}^p are both denoted by $\|\cdot\|$. We introduce the following assumptions in order to apply Theorem 1.1 to $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$:

(H'1) $\|A\| < 1$ and there exists $a \geq 1$ such that $\mathbb{E}[\|\vartheta_1\|^a] < \infty$.

(H'2) $\sup_{\theta \in \Theta} \int_{\mathbb{R}^q} \|h_\theta(v)\|^a \nu(v) dv < \infty$

(H'3) For every $\theta \in \Theta$, h_θ is differentiable on \mathbb{R}^q with $\sup_{\theta \in \Theta} \|\nabla h_\theta\|_\infty < \|A\|^{-1}$,

(H'4) (a) The p.d. of ϑ_1 admits a bounded continuous p.d.f. ν satisfying the following condition: there exists $M > 0$ such that for every $z_1, z_2 \in \mathbb{R}^q$

$$M \leq \|z_1\| \leq \|z_2\| \implies \nu(z_2) \leq \nu(z_1).$$

(b) For every $\theta \in \Theta$, the map h_θ is a \mathcal{C}^1 -diffeomorphism on \mathbb{R}^q with inverse function denoted by g_θ , and the following conditions hold:

i. $\exists c \in (0, 1), \forall \theta \in \Theta, \forall z \in \mathbb{R}^q, \|g_\theta(z) - z\| \leq c\|z\|$.

- ii. $\forall z \in \mathbb{R}^q, \lim_{\theta \rightarrow 0} g_\theta(z) = z.$
- iii. $\sup_{\theta \in \Theta} \|\nabla g_\theta\|_\infty < \infty,$ and $\lim_{\theta \rightarrow 0} \nabla g_\theta = id$ uniformly on each centred balls of $\mathbb{R}^q,$ that is:

$$\forall A > 0, \forall \eta > 0, \exists \alpha > 0, \forall \theta \in \Theta, \|\theta\| < \alpha, \sup_{\|z\| \leq A} \|\nabla g_\theta(z) - id\| < \eta.$$

Proposition 3.1 *Under Conditions (H'1)–(H'4) for a VAR process as defined in Example 1.1, Properties (P1)–(P3) of Theorem 1.1 hold true with any real number $\rho_a \in (\|A\|, 1)$ (and with the associated constant C_a).*

Remark 3.1 *Conditions in (H'4)-(b) focus on the inverse function g_θ of h_θ because g_θ naturally occurs in the proof after a change of variable. Note that, as in [BRR01], the uniform convergence $\lim_{\theta \rightarrow 0} g_\theta = id$ (or $\lim_{\theta \rightarrow 0} h_\theta = id$) is not required on the whole space \mathbb{R}^q in the above assumptions. For instance the roundoff functions $h_\theta(x) = x + \theta x$ (simple perturbation of id on \mathbb{R}) satisfy the above assumptions, but neither the convergence $\lim_{\theta \rightarrow 0} g_\theta = id,$ nor the convergence $\lim_{\theta \rightarrow 0} h_\theta = id,$ are uniform on $\mathbb{R}.$*

Proof. Recall that $\theta_0 = 0$ here. Using the computations in Example 1.1 it follows from Proposition 4.2 and Remark 4.3 that Assumption (H1) holds for every real number $\rho_a \in (\|A\|, 1)$ since P_0 is associated with a continuous kernel $K(x, y)$ (use (26) with $\theta = 0$). Next, for any $\theta \in \Theta$ and $z \in \mathbb{R}^q,$ set $\Gamma_\theta(z) = |\det \nabla g_\theta(z)|.$ Then P_θ is from (24)

$$\forall f \in \mathcal{B}_0, \forall x \in \mathbb{R}^q, (P_\theta f)(x) = \int_{\mathbb{R}^q} f(z) \nu(g_\theta(z) - Ax) \Gamma_\theta(z) dz \quad (26)$$

from the change of variable $z = h_\theta(Ax + v).$ Recall that P_θ is the transition kernel of the \mathbb{R}^q -valued IFS $\{X_n^{(\theta)}\}_{n \in \mathbb{N}}$ associated with $F_\theta(x, y) := h_\theta(Ax + y).$ Then (H'2) is nothing else but (H2) (here $x_0 = 0$), while (H3) is implied by (H'3) from Taylor's inequality applied to $h_\theta.$

Next we prove (H4). For every $r > 0,$ let $B(0, r) = \{z \in \mathbb{R}^q : \|z\| \leq r\}.$ Let $f \in \mathcal{B}_0$ such that $|f|_0 \leq 1,$ and let $x \in \mathbb{R}^q.$ Fix $\varepsilon > 0.$ First let $K \equiv K(\varepsilon) > 0$ be such that $(1 + K)^{-a} < \varepsilon/2.$ Then

$$\forall x \in \mathbb{R}^q \setminus B(0, K), \frac{|(P_\theta f)(x) - (P_0 f)(x)|}{V(x)} \leq \frac{2}{V(x)} < \varepsilon. \quad (27)$$

Now we assume that $x \in B(0, K).$ Note that

$$|(P_\theta f)(x) - (P_0 f)(x)| \leq \int_{\mathbb{R}^q} |\nu(g_\theta(z) - Ax) \Gamma_\theta(z) - \nu(z - Ax)| dz \quad (28)$$

since $g_0 = id.$ Set $d := 2/(1 - c)$ where c is given in (H'4)-(b)-(i). Note that $\|Ax\| \leq K$ and that (H'4)-(b)-(i) provides: $\forall z \in \mathbb{R}^q, \|g_\theta(z)\| \geq (1 - c)\|z\|.$ Then we have for every $z \in \mathbb{R}^q$ such that $\|z\| \geq dK$

$$\|g_\theta(z) - Ax\| \geq \|g_\theta(z)\| - \|Ax\| \geq (1 - c)\|z\| - K \geq (1 - c)\|z\| - \frac{1}{d}\|z\| \geq \frac{1 - c}{2}\|z\|.$$

It follows from (H'₄)-(a) that we have for every $\theta \in \Theta$

$$\|z\| \geq B \equiv B(\varepsilon) := \max(dM, dK) \implies \nu(g_\theta(z) - Ax) \leq \nu(d^{-1}z).$$

Since the function $z \mapsto \nu(d^{-1}z)$ is Lebesgue-integrable on \mathbb{R}^q , we can choose $C \equiv C(\varepsilon) > 0$ such that $\int_{\|z\| \geq C} \nu(d^{-1}z) dz \leq \varepsilon/2(\gamma + 1)$ where

$$\gamma := \sup_{\theta \in \Theta} \sup_{z \in \mathbb{R}^q} \Gamma_\theta(z).$$

Note that $\gamma < \infty$ from the first condition of (H'₄)-(b)-(iii) and from the continuity of $\det(\cdot)$. Set $D = \max(B, C)$. We deduce from the triangular inequality that for every $\theta \in \Theta$

$$\int_{\|z\| \geq D} |\nu(g_\theta(z) - Ax) \Gamma_\theta(z) - \nu(z - Ax)| dz \leq (\gamma + 1) \int_{\|z\| \geq C} \nu(d^{-1}z) dz \leq \frac{\varepsilon}{2}. \quad (29)$$

Now we investigate the integrand in (28) for $z \in B(0, D)$ (recall that $x \in B(0, K)$). First, setting $m := \sup_{u \in \mathbb{R}^q} \nu(u)$, we have for every $z \in B(0, D)$ and for every $x \in B(0, K)$

$$|\nu(g_\theta(z) - Ax) \Gamma_\theta(z) - \nu(z - Ax)| \leq \gamma |\nu(g_\theta(z) - Ax) - \nu(z - Ax)| + m |\Gamma_\theta(z) - 1|. \quad (30)$$

We have: $\forall z \in B(0, D)$, $\|g_\theta(z)\| \leq (1+c)D$ (use (H'₄)-(b)-(i)). From the standard statement for uniform convergence of differentiable functions, we deduce from Conditions (H'₄)-(b)-(ii) and (H'₄)-(b)-(iii) that $\lim_{\theta \rightarrow 0} g_\theta = id$ uniformly on $B(0, D)$. Let ℓ_D denote the volume of $B(0, D)$ with respect to Lebesgue's measure on \mathbb{R}^q . From the previous uniform convergence and from the uniform continuity of ν on $B(0, (1+c)D+K)$, there exists an open neighbourhood V_0 of $\theta = 0$ in \mathbb{R}^p such that

$$\forall \theta \in V_0, \forall z \in B(0, D), \forall x \in B(0, K), \quad |\nu(g_\theta(z) - Ax) - \nu(z - Ax)| < \frac{\varepsilon}{4\gamma\ell_D}.$$

Moreover there exists an open neighbourhood $V'_0 \subset V_0$ of $\theta = 0$ in \mathbb{R}^p such that

$$\forall \theta \in V'_0, \forall z \in B(0, D), \quad |\Gamma_\theta(z) - 1| < \frac{\varepsilon}{4m\ell_D}$$

from (H'₄)-(b)-(iii) and from the uniform continuity of the function $\det(\cdot)$ on every compact subset of the set $\mathcal{M}_q(\mathbb{R})$ of real $q \times q$ -matrices. Then it follows from (30) that

$$\forall \theta \in V'_0, \forall z \in B(0, D), \forall x \in B(0, K), \quad |\nu(g_\theta(z) - Ax) \Gamma_\theta(z) - \nu(z - Ax)| \leq \frac{\varepsilon}{2\ell_D}.$$

Integrating this inequality on $B(0, D)$ gives

$$\forall \theta \in V'_0, \forall x \in B(0, K), \quad \int_{\|z\| \leq D} |\nu(g_\theta(z) - Ax) \Gamma_\theta(z) - \nu(z - Ax)| dz \leq \frac{\varepsilon}{2}. \quad (31)$$

We deduce from (28), (29) and (31) that

$$\forall \theta \in V'_0, \forall x \in B(0, K), \quad \frac{|(P_\theta f)(x) - (P_0 f)(x)|}{V(x)} \leq |(P_\theta f)(x) - (P_0 f)(x)| \leq \varepsilon$$

This inequality combined with (27) gives (H₄). □

4 V_a -geometric ergodicity of IFS

For $a \geq 1$, define for any $x \in \mathbb{X}$, $p(x) := 1 + d(x, x_0)$, so that $V_a(x) := p(x)^a$, and let us introduce the following space \mathcal{L}_a :

$$\mathcal{L}_a := \left\{ f : \mathbb{X} \rightarrow \mathbb{C} : m_a(f) := \sup \left\{ \frac{|f(x) - f(y)|}{d(x, y) (p(x) + p(y))^{a-1}}, (x, y) \in \mathbb{X}^2, x \neq y \right\} < \infty \right\}. \quad (32)$$

Such Lipschitz-weighted spaces have been introduced in [LP83] to obtain the quasi-compactness of Lipschitz kernels (see also [MR89, Hen93, Duf97, Ben98, HH01]). Note that, for $f \in \mathcal{L}_a$, we have for all $x \in \mathbb{X}$: $|f(x)| \leq |f(x_0)| + 2^{a-1} m_a(f) V_a(x)$ so that $|f|_a < \infty$ for any $f \in \mathcal{L}_a$. Hence $\mathcal{L}_a \subset \mathcal{B}_a$. Moreover \mathcal{L}_a is a Banach space when equipped with the norm

$$\forall f \in \mathcal{L}_a, \quad \|f\|_a := m_a(f) + |f|_a. \quad (33)$$

Let $\{X_n\}_{n \in \mathbb{N}}$ be an IFS of Lipschitz maps as in Definition 1.1. For all $x \in \mathbb{X}$ and $v \in \mathbb{V}$, we set $F_v x := F(x, v)$. Recall that we have set $L_F(v) := L(F_v)$ in Introduction. Since F is fixed in this section, we simply write $L(v)$ for $L_F(v)$. Similarly, for every $(v_1, \dots, v_n) \in \mathbb{V}^n$ ($n \in \mathbb{N}^*$), define:

$$F_{v_n:v_1} := F_{v_n} \circ \dots \circ F_{v_1} \quad \text{and} \quad L(v_n : v_1) := L(F_{v_n:v_1}). \quad (34)$$

By hypothesis we have $L(v) < \infty$, thus $L(v_n : v_1) < \infty$. Note that, for each $a \geq 1$, the limit

$$\hat{\kappa}_a := \lim_{n \rightarrow +\infty} \mathbb{E} [L(F_{\vartheta_n:\vartheta_1})^a]^{\frac{1}{na}}$$

exists in $[0, +\infty]$, since the sequence $(\mathbb{E}[L(\vartheta_n : \vartheta_1)^a])_{n \in \mathbb{N}^*}$ is sub-multiplicative. In this section we first present standard contraction/moment Conditions $(\widehat{\mathcal{C}}_a)$ (counterpart of (\mathcal{C}_a) of Introduction) for P given in (3) to have a geometric rate of convergence on \mathcal{L}_a (see Proposition 4.1). Then the passage to the V_a -geometric ergodicity is addressed in Proposition 4.2.

Conditions $(\widehat{\mathcal{C}}_a)$. For some $a \in [1, +\infty)$:

$$\mathbb{E} [d(x_0, F(x_0, \vartheta_1))^a] < \infty \quad (35a)$$

$$\hat{\kappa}_a < 1. \quad (35b)$$

Note that Condition (35b) is equivalent to the following one

$$\exists N \in \mathbb{N}^*, \mathbb{E} [L(\vartheta_N : \vartheta_1)^a] < 1 \quad (36)$$

and Conditions (\mathcal{C}_a) of Introduction corresponds to (35a) and to (36) with $N = 1$.

The properties of the next proposition can be derived from the results of [Duf97, Chapter 6], also see [Ben98] for the existence and uniqueness of the invariant distribution. For convenience, in Appendix C, the properties (37a) and (37b) are proved (with explicit constants) under the assumption $\mathbb{E}[L(\vartheta_1)^a] < 1$.

Proposition 4.1 ([[Du97](#), Chapter 6]) *Under Conditions $(\widehat{\mathcal{C}}_a)$, P has a unique invariant distribution on $(\mathbb{X}, \mathcal{X})$, denoted by π , and we have $\pi(d(x_0, \cdot)^a) < \infty$. Moreover the Markov kernel P continuously acts on \mathcal{L}_a , and for any $\kappa \in (\widehat{\kappa}_a, 1)$, there exists positive constants $c \equiv c_\kappa$ and $c' \equiv c'_\kappa$ such that:*

$$\forall f \in \mathcal{L}_a, \forall n \geq 1, \quad |P^n f - \pi(f)1_{\mathbb{X}}|_a \leq c \kappa^n m_a(f) \quad (37a)$$

$$\forall f \in \mathcal{L}_a, \forall n \geq 1, \quad \|P^n f - \pi(f)1_{\mathbb{X}}\|_a \leq c' \kappa^n \|f\|_a. \quad (37b)$$

In particular, if $\kappa_{1,a} := \mathbb{E}[L(\vartheta_1)^a]^{1/a} < 1$, then

$$\forall f \in \mathcal{L}_a, \forall n \geq 1, \quad |P^n f - \pi(f)1_{\mathbb{X}}|_a \leq c_1 \kappa_{1,a}^n m_a(f), \quad (38)$$

where the constant c_1 is defined by $c_1 := \xi^{(a-1)/a} \|\pi\|_1 (1 + \|\pi\|_a)^{a-1}$, with

$$\xi := \sup_{n \geq 1} \sup_{x \in \mathbb{X}} \frac{(P^n V_a)(x)}{V_a(x)} < \infty \quad \text{and} \quad \|\pi\|_b := \left(\int_{\mathbb{X}} V_b(y) d\pi(y) \right)^{\frac{1}{b}} \quad \text{for } b := 1, a.$$

Under Conditions $(\widehat{\mathcal{C}}_a)$, Property (37a) with $f := V_a$ and $n := 1$ gives $PV_a \leq \xi_1 V_a$ for some $\xi_1 \in (0, +\infty)$, so that P continuously acts on \mathcal{B}_a . But it is worth noticing that Property (37a) (or (38)) does not provide the V_a -geometric ergodicity (4) since (37a) (or (38)) is only established for $f \in \mathcal{L}_a$. Under Conditions $(\widehat{\mathcal{C}}_a)$ Alsmeyer proved in [[Als03](#), Prop. 5.2] that, if $\{X_n\}_{n \in \mathbb{N}}$ is Harris recurrent and the support of π has a non-empty interior, then $\{X_n\}_{n \in \mathbb{N}}$ is V_a -geometrically ergodic. Under Conditions $(\widehat{\mathcal{C}}_a)$ $\{X_n\}_{n \in \mathbb{N}}$ is shown to be V_a -geometrically ergodic in [[Wu04](#), Prop. 7.2] provided that P and P^N for some $N \geq 1$ are Feller and strongly Feller respectively. An alternative approach is proposed in Proposition 4.1 below. The bound (39) is the same as in [[Wu04](#), Prop. 7.2], but the Feller-type assumptions of [[Wu04](#)] are replaced by the following one : $P^\ell : \mathcal{B}_0 \rightarrow \mathcal{B}_a$ for some $\ell \geq 1$ is compact (see Remark 4.2 for comparisons).

Proposition 4.2 *Let us assume that Conditions $(\widehat{\mathcal{C}}_a)$ hold true and that $P^\ell : \mathcal{B}_0 \rightarrow \mathcal{B}_a$ for some $\ell \geq 1$ is compact. Then P is V_a -geometrically ergodic, and the spectral gap $\rho_{V_a}(P)$ of P on \mathcal{B}_a (i.e. the infimum bound of the positive real numbers ρ_a such that Property (4) holds true) satisfies the following bound:*

$$\rho_{V_a}(P) \leq \widehat{\kappa}_a. \quad (39)$$

Proof. To avoid confusion, we simply denote by P the action of $P(x, dy)$ on \mathcal{B}_a , and we denote by $P|_{\mathcal{L}_a}$ the restriction of P on \mathcal{L}_a . Let δ and κ be such that $\widehat{\kappa}_a < \kappa < \delta < 1$. Then there exists $N \in \mathbb{N}^*$ such that $c \kappa^N m_a(V_a) \leq \delta^N$, where $c \equiv c_\kappa$ is defined in (37a). Then Property (37a) applied to $f := V_a$ gives: $P^N V_a \leq \delta^N V_a + \pi(V_a)$. We deduce from [[HL14a](#), Prop. 5.4 and Rk. 5.5] that P is a power bounded quasi-compact operator on \mathcal{B}_a and that its essential spectral radius $r_{ess}(P)$ satisfies $r_{ess}(P) \leq \widehat{\kappa}_a$ since δ is arbitrarily close to $\widehat{\kappa}_a$ (e.g. see [[Hen93](#)] for the definition of the quasi-compactness and of the essential spectral radius of a linear bounded operator). From these properties it follows that the adjoint operator P^* of P is quasi-compact on the dual space \mathcal{B}'_a of \mathcal{B}_a and that $r_{ess}(P^*) \leq \widehat{\kappa}_a$.

Next, let us establish that P is V_a -geometrically ergodic from Conditions (b) in [[GHL11](#), Prop. 2.1]. Let $r_0 \in (\widehat{\kappa}_a, 1)$. Prove that $\lambda := 1$ is the only eigenvalue of P on \mathcal{B}_a such that

$r_0 \leq |\lambda| \leq 1$. Let $\lambda \in \mathbb{C}$ be such an eigenvalue. Then λ is also an eigenvalue of P^* since P and P^* have the same spectrum and $r_{ess}(P^*) \leq \widehat{\kappa}_a < |\lambda|$. Thus there exists $f \in \mathcal{B}'_a$ such that $f' \circ P = \lambda f'$. But f' is also in \mathcal{L}'_a since we have: $\forall f \in \mathcal{L}_a$, $|\langle f', f \rangle| \leq \|f'\|_{\mathcal{B}'_a} |f|_a \leq \|f'\|_{\mathcal{B}'_a} \|f\|_a$. This proves that λ is an eigenvalue of the adjoint of $P|_{\mathcal{L}_a}$. Hence λ is a spectral value of $P|_{\mathcal{L}_a}$. More precisely λ is an eigenvalue of $P|_{\mathcal{L}_a}$ since, from (37b), $P|_{\mathcal{L}_a}$ is quasi-compact on \mathcal{L}_a and $r_{ess}(P|_{\mathcal{L}_a}) \leq \widehat{\kappa}_a < r_0 \leq |\lambda|$. Finally we have $\lambda = 1$. Indeed, if $\lambda \neq 1$, then any $f \in \mathcal{L}_a$ satisfying $Pf = \lambda f$ is such that $\pi(f) = 0$, thus $f = 0$ from (37b) (pick $\kappa \in (\widehat{\kappa}_a, r_0)$).

Now prove that 1 is a simple eigenvalue of P on \mathcal{B}_a . Using the previous property and the fact that P is power bounded and quasi-compact on \mathcal{B}_a , we know that $P^n \rightarrow \Pi$ with respect to the operator norm on \mathcal{B}_a , where Π is the finite rank eigen-projection on $\text{Ker}(P - I) = \text{Ker}(P - I)^2$. The last equality holds since P is power bounded on \mathcal{B}_a . Set $m := \dim \text{Ker}(P - I)$. From [Wu04, Prop. 4.6] (see also [Her08, Th. 1]), there exist m linearly independent nonnegative functions $f_1, \dots, f_m \in \text{Ker}(P - I)$ and probability measures $\mu_1, \dots, \mu_m \in \text{Ker}(P^* - I)$ satisfying $\mu_k(V_a) < \infty$ such that: $\forall f \in \mathcal{B}_a$, $\Pi f = \sum_{k=1}^m \mu_k(f) f_k$. That 1 is a simple eigenvalue of P on \mathcal{B}_a then follows from the first assertion of Proposition 4.1.

From [GHL11, Prop. 2.1] and the previous results, we have proved that, for any $r_0 \in (\widehat{\kappa}_a, 1)$, we have $\rho_{V_a}(P) \leq r_0$. Thus $\rho_{V_a}(P) \leq \widehat{\kappa}_a$. \square

Remark 4.1 *Inequality (39) means that, for any real number $\rho \in (\widehat{\kappa}_a, 1)$, there exists a constant $C \equiv C_\rho$ such that*

$$\forall n \geq 1, \forall f \in \mathcal{B}_a, |P^n f - \pi(f) 1_{\mathbb{X}}|_a \leq C \rho^n |f|_a.$$

Unfortunately neither the proof of Proposition 4.1, nor that of [Wu04, Prop. 7.2], give any information on the constant C . Computing such an explicit constant C is an intricate issue which is not addressed in this work (e.g. see [MT94, LT96, Bax05, HL14a, HL14b] and the reference therein). Mention that explicit bounds on ρ and C are also provided in [GP14] for a parametrized family of transition kernels.

Remark 4.2 *Let us just mention the following classical statement. Assume that every closed ball of \mathbb{X} is compact. Let $\{X_n\}_{n \in \mathbb{N}}$ be a Markov chain such that its transition kernel P satisfies the following hypothesis: there exist a positive measure η on $(\mathbb{X}, \mathcal{X})$ and a measurable function $K : \mathbb{X}^2 \rightarrow [0, +\infty)$ such that:*

$$\forall x \in \mathbb{X}, P(x, dy) = K(x, y) d\eta(y). \quad (40)$$

If P^ℓ is strongly Feller for some $\ell \geq 1$, then $P^{2\ell}$ is compact from \mathcal{B}_0 to \mathcal{B}_a (e.g. see [GHL11, Lemma 3]). Hence, when P admits a kernel as in (40), then assuming that P^N is strongly Feller for some N in [Wu04, Prop. 7.2] is more restrictive than the compactness hypothesis of Proposition 4.2. A detailed comparison with the approach [Wu04, Prop. 7.2] is provided in [GHL11] for general Markov kernels.

Remark 4.3 *If $\{X_n\}_{n \in \mathbb{N}}$ is an IFS of Lipschitz maps as in Definition 1.1 such that its transition kernel P satisfies Assumption (40) with K continuous in the first variable, then P is*

strongly Feller, thus P^2 is compact from \mathcal{B}_0 to \mathcal{B}_a , so that the conclusion (39) of Proposition 4.2 holds true under Conditions $(\widehat{\mathcal{C}}_a)$. Indeed we have for all $(x, x') \in \mathbb{X}^2$ and for any $f \in \mathcal{B}_0$

$$|(Pf)(x') - (Pf)(x)| \leq \int_{\mathbb{X}} |K(x', y) - K(x, y)| d\eta(y).$$

Since $K(\cdot, \cdot) \geq 0$, $\int K(\cdot, y) d\eta(y) = 1$, and $\lim_{x' \rightarrow x} K(x', y) = K(x, y)$, we deduce from Scheffé's theorem that $\lim_{x' \rightarrow x} \int_{\mathbb{X}} |K(x', y) - K(x, y)| d\eta(y) = 0$. This proves the desired statement. Note that the previous argument even shows that $\{Pf, |f|_0 \leq 1\}$ is equicontinuous, so that the compactness of $P : \mathcal{B}_0 \rightarrow \mathcal{B}_1$ can be directly proved from Ascoli's theorem.

Remark 4.4 In the proof of Proposition 4.2 the drift inequality $P^N V_a \leq \delta^N V_a + \pi(V_a)$ has been written with any $\delta \in (\widehat{\kappa}_a, 1)$ by using Property (37a) of Proposition 4.1 in order to deduce the bound $r_{ess}(P) \leq \widehat{\kappa}_a$ on the essential spectral radius of P (acting on \mathcal{B}_a). This bound was sufficient since the remainder of the proof of Proposition 4.2 is based on Property (37b) from which we deduce the bound $r_{ess}(P|_{\mathcal{L}_a}) \leq \widehat{\kappa}_a$. Actually the drift inequality $P^N V_a \leq \delta^N V_a + K$ with any $\delta \in (\widehat{\kappa}_a^a, 1)$ et with some $K > 0$ can be derived from Conditions $(\widehat{\mathcal{C}}_a)$ by adapting the proof in Appendix B (with here $P_{\theta_0} = P$ and $\Theta = \{\theta_0\}$). Then the more accurate bound $r_{ess}(P) \leq \widehat{\kappa}_a^a$ can be derived from [HL14a, Prop. 5.4 and Rem. 5.5] under the compactness assumption of Proposition 4.2. See [GHL11] for details and see also [Wu04, Prop. 7.2] which provides the same bound under Feller-type assumptions.

5 Conclusion

Theorem 1.1 has been first applied in Section 2 for the real-valued AR(1) with ARCH(1) errors models (see Proposition 2.1), and second in Section 3 for roundoff errors of a VAR model (see Proposition 3.1). Although these applications have been presented for specific IFS, it is worth noticing that they give a general pattern to investigate the issues (P₁)-(P₂)-(P₃) of Introduction for others instances of \mathbb{R}^q -valued IFS, provided that the noise ν_γ in Definition 1.2 admits a p.d.f. with respect to Lebesgue's measure on \mathbb{R}^q and that the change of variable $v \mapsto z = F_\xi(x, v)$ is feasible for every $x \in \mathbb{R}^q$, where $F_\xi(\cdot, \cdot)$ is the perturbed function involved in Definition 1.2.

For instance, denote by $GL_q(\mathbb{R})$ the set of invertible real $q \times q$ -matrices, and consider the following general perturbed \mathbb{R}^q -valued IFS $\{X_n\}_{n \in \mathbb{N}}$ given by

$$\forall n \geq 1, \quad X_n = \psi_\xi(X_{n-1}) + B_\xi(X_{n-1}) \vartheta_n^{(\gamma)}$$

with some parametrized maps $\psi_\xi : \mathbb{R}^q \rightarrow \mathbb{R}^q$ and $B_\xi : \mathbb{R}^q \rightarrow GL_q(\mathbb{R})$, and with an i.i.d. sequence $\{\vartheta_n^{(\gamma)}\}_{n \geq 1}$ of \mathbb{R}^q -valued r.v. with common parametric p.d.f. denoted by ν_γ (hence $\theta = (\xi, \gamma)$). Then, noticing that for every $x \in \mathbb{R}^q$ the change of variable $v \mapsto z := \psi_\xi(x) + B_\xi(x)v$ is valid and leads to $P_\theta(x, A) := \int_{\mathbb{R}} 1_A(z) p_\theta(x, z) dz$ ($A \in \mathcal{X}$) with

$$p_\theta(x, z) := |\det B_\xi(x)|^{-1} \nu_\gamma(B_\xi(x)^{-1}(z - \psi_\xi(x))) \quad (41)$$

the following remarks are relevant to investigate the assumptions (H₁)-(H₄) of Theorem 1.1.

(R1) If the p.d.f. ν_{γ_0} of the unperturbed IFS (corresponding to $\theta_0 = (\xi_0, \gamma_0)$) is continuous on \mathbb{R}^q , then it follows from Remark 4.3 and Proposition 4.2 that P_{θ_0} is V_a -geometrically ergodic provided that the unperturbed IFS satisfies Conditions (\mathcal{C}_a) . More precisely, in this case, Assumption (H_1) holds with any real number ρ_a (and the associated constant C_a) such that

$$\mathbb{E} \left[L_{F_{\xi_0}} (\vartheta_1^{(\gamma_0)})^a \right]^{1/a} < \rho_a < 1 \quad \text{where} \quad F_{\xi_0}(x, v) = \psi_{\xi_0}(x) + B_{\xi_0}(x) v.$$

(R2) The assumptions in (R1), that is the moment/contractive conditions (6a) and (6b) related to $\theta_0 = (\xi_0, \gamma_0)$, involve some expectations which simply write with the above function F_{ξ_0} and with the p.d.f. ν_{γ_0} . Then Conditions (H_2) and (H_3) consists in assuming that these expectations are respectively bounded and strictly less than 1 in a uniform way on the parameters $\theta := (\xi, \gamma)$ near $\theta_0 = (\xi_0, \gamma_0)$ (up to reduce the set Θ).

(R3) Thanks to Formula (41) proving (H_4) remains to check that for every $A > 0$

$$\lim_{\theta \rightarrow \theta_0} \sup_{\|x\| \leq A} \int_{\mathbb{R}^q} \frac{|p_\theta(x, z) - p_{\theta_0}(x, z)|}{(1 + \|x\|)^a} dz = 0$$

(indeed note that the previous integral is less than $2/(1 + A)^a$ for $\|x\| > A$).

Note that such a material enables to investigate the general class of non-linear AR-ARCH models introduced in [MS10, and references therein].

Similarly in the context of roundoff errors, let $\{X_n\}_{n \in \mathbb{N}}$ be a fixed IFS of the form

$$\forall n \geq 1, \quad X_n = \psi(X_{n-1}) + B(X_{n-1}) \vartheta_n$$

where $\psi : \mathbb{R}^q \rightarrow \mathbb{R}^q$ and $B : \mathbb{R}^q \rightarrow \text{GL}_q(\mathbb{R})$ are fixed here, as well as the common p.d.f. ν of $\{\vartheta_n\}_{n \geq 1}$. Consider a roundoff error family $(h_\theta)_{\theta \in \Theta}$ with h_θ close to $h_0 = id$ when $\theta \rightarrow 0$ (see Section 3). Then the roundoff transition kernel $P_\theta(x, A) = P(x, h_\theta^{-1}(A))$ writes as $P_\theta(x, A) := \int_{\mathbb{R}} 1_A(z) p_\theta(x, z) dz$ with

$$p_\theta(x, z) := \Gamma_\theta(z) \nu(B(x)^{-1}(g_\theta(z) - \psi(x))) \quad (42)$$

from the change of variable $v \mapsto z := h_\theta(\psi(x) + B(x)v)$, where g_θ denotes the inverse function of h_θ and $\Gamma_\theta(z) := |\det B_\xi(x)|^{-1} |\det \nabla g_\theta(z)|$. Again the following remarks are relevant to investigate the assumptions (H_1) – (H_4) . If the p.d.f. ν of $\{X_n\}_{n \in \mathbb{N}}$ is continuous on \mathbb{R}^q , then it follows from Remark 4.3 and Proposition 4.2 that $\{X_n\}_{n \in \mathbb{N}}$ is V_a -geometrically ergodic under Conditions (\mathcal{C}_a) . More precisely, in this case, Assumption (H_1) holds with any real number ρ_a (and the associated constant C_a) such that

$$\mathbb{E} [L_F(\vartheta_1)^a]^{1/a} < \rho_a < 1 \quad \text{where} \quad F(x, v) = \psi(x) + B(x) v.$$

The previous remarks (R2) and (R3) obviously extends to the roundoff context.

A Proof of (10)

We suppose that Assumptions (H₂)-(H₃) are fulfilled, and we prove the drift inequality (10) of Introduction. More precisely, let $\kappa \in (\kappa_a, 1)$, and prove that the following strengthened condition holds:

$$\forall \theta \in \Theta, \quad P_\theta V_a \leq \delta_a V_a + K_a 1_{[-r_a, r_a]} \quad (43)$$

where the constants $\delta_a < 1$ and $K_a > 0$ are given in (10), and where $r_a := \frac{1+M_a+\kappa_a-\kappa}{\kappa-\kappa_a}$. We have for any $\theta \in \Theta$ and any $x \in \mathbb{X}$

$$\begin{aligned} \left(\frac{(P_\theta V_a)(x)}{V_a(x)} \right)^{1/a} &= \left(\mathbb{E} \left[\left(\frac{1 + d(F_\xi(x, \vartheta_1^{(\gamma)}); x_0)}{1 + d(x; x_0)} \right)^a \right] \right)^{1/a} \\ &\leq \left(\mathbb{E} \left[\left(\frac{1 + d(F_\xi(x, \vartheta_1^{(\gamma)}); F_\xi(x_0, \vartheta_1^{(\gamma)})) + d(F_\xi(x_0, \vartheta_1^{(\gamma)}); x_0)}{1 + d(x; x_0)} \right)^a \right] \right)^{1/a} \\ &\leq \left(\mathbb{E} \left[\left(\frac{1}{1 + d(x; x_0)} + L_{F_\xi}(\vartheta_1^{(\gamma)}) + \frac{d(F_\xi(x_0, \vartheta_1^{(\gamma)}); x_0)}{1 + d(x; x_0)} \right)^a \right] \right)^{1/a} \\ &\leq \frac{1}{1 + d(x; x_0)} + \mathbb{E}[L_{F_\xi}(\vartheta_1^{(\gamma)})^a]^{1/a} + \frac{\mathbb{E}[d(F_\xi(x_0, \vartheta_1^{(\gamma)}); x_0)^a]^{1/a}}{1 + d(x; x_0)} \quad (\text{Holder inequality}). \end{aligned}$$

It follows from Assumption (H₂)-(H₃) that

$$\forall \theta \in \Theta, \quad \forall x \in \mathbb{X}, \quad \left(\frac{(P_\theta V_a)(x)}{V_a(x)} \right)^{1/a} \leq \frac{1}{1 + d(x; x_0)} + \kappa_a + \frac{M_a}{1 + d(x; x_0)}. \quad (44)$$

Let $r_a > 0$ as above defined, so that we have for every $x \in \mathbb{X}$ such that $d(x; x_0) > r_a$

$$\frac{1 + M_a}{1 + d(x; x_0)} \leq \frac{1 + M_a}{1 + r_a} = \kappa - \kappa_a$$

Then we obtain that for every $\theta \in \Theta$ and for every $x \in \mathbb{X}$ such that $d(x; x_0) > r_a$

$$(P_\theta V_a)(x) \leq \kappa^a V_a(x). \quad (45)$$

Moreover, for every $\theta \in \Theta$ and for every $x \in \mathbb{X}$ such that $d(x; x_0) \leq r_a$, we deduce from (44) that

$$(P_\theta V_a)(x) \leq F_a V_a(x) \leq F_a(1 + r_a)^a \quad (46)$$

where $F_a := (1 + \kappa_a + M_a)^a$. Finally combining (45) and (46) provides (43), thus (10), with $\delta_a := \kappa_a^a < 1$ and $K_a := F_a(1 + r_a)^a > 0$.

B Complements on the proof of Proposition 2.1

First we prove Property (16).

Lemma B.1 Let $(\alpha, \beta, \lambda) \in \mathbb{R} \times (0, +\infty)^2$ and: $\forall(x, v) \in \mathbb{R}^2, F(x, v) = \alpha x + v\sqrt{\beta + \lambda x^2}$. Then we have for every $v \in \mathbb{R}$

$$L(v) := \sup_{(x,y) \in \mathbb{R}^2, x \neq y} \frac{|F(x, v) - F(y, v)|}{|x - y|} = \max(|\alpha - \sqrt{\lambda}v|; |\alpha + \sqrt{\lambda}v|). \quad (47)$$

Proof. Let $v \in \mathbb{R}$ be fixed, and define: $\forall x \in \mathbb{R}, F_v(x) := F(x, v)$. Then

$$\forall x \in \mathbb{R}, \quad F'_v(x) = \alpha + \frac{\lambda xv}{(\beta + \lambda x^2)^{1/2}} \quad \text{and} \quad F''_v(x) = \frac{\lambda \beta v}{(\beta + \lambda x^2)^{3/2}}.$$

Property (47) is obvious if $v = 0$. Assume that $v > 0$. Then F'_v is strictly increasing, so that

$$\inf_{x \in \mathbb{R}} F'_v(x) = \lim_{x \rightarrow -\infty} F'_v(x) = \alpha - \sqrt{\lambda}v \leq \alpha + \sqrt{\lambda}v = \lim_{x \rightarrow +\infty} F'_v(x) = \sup_{x \in \mathbb{R}} F'_v(x).$$

Then $L(v) \leq \max(|\alpha - \sqrt{\lambda}v|; |\alpha + \sqrt{\lambda}v|)$ follows from Taylor's inequality. If $v < 0$, then F'_v is strictly decreasing, so that

$$\inf_{x \in \mathbb{R}} F'_v(x) = \lim_{x \rightarrow +\infty} F'_v(x) = \alpha + \sqrt{\lambda}v \leq \alpha - \sqrt{\lambda}v = \lim_{x \rightarrow -\infty} F'_v(x) = \sup_{x \in \mathbb{R}} F'_v(x),$$

and the same conclusion holds. That $L(v) \geq \max(|\alpha - \sqrt{\lambda}v|; |\alpha + \sqrt{\lambda}v|)$ follows from the inequality $L(v) \geq |F'_v(x)|$ for any $x \in \mathbb{R}$, which is easily deduced from the definition of $L(v)$ in (47). Hence we obtain that $L(v) \geq \lim_{x \pm \infty} |F'_v(x)|$. The proof of (47) is complete. \square

Lemma B.2 Let $(\alpha_0, \beta_0, \lambda_0) \in \mathbb{R} \times (0, +\infty)^2$. For any $(\alpha, \beta, \lambda) \in \mathbb{R} \times (0, +\infty)^2$ and $x \in \mathbb{R}$, define

$$b_{\beta, \lambda}(x) := \left(\frac{\beta + \lambda x^2}{\beta_0 + \lambda_0 x^2} \right)^{1/2} \quad \text{and} \quad a_\alpha(x) := x \frac{\alpha - \alpha_0}{\sqrt{\beta_0 + \lambda_0 x^2}}.$$

Then for any $A > 0$

$$\lim_{(\beta, \lambda) \rightarrow (\beta_0, \lambda_0)} \sup_{|x| \leq A} |b_{\beta, \lambda}(x) - 1| = 0 \quad \text{and} \quad \lim_{\alpha \rightarrow \alpha_0} \sup_{|x| \leq A} a_\alpha(x) = 0.$$

Proof. Let $A > 0$. We have for any $x \in \mathbb{R}$ such that $|x| \leq A$

$$|b_{\beta, \lambda}(x)^2 - 1| = \left| \frac{\beta + \lambda x^2}{\beta_0 + \lambda_0 x^2} - 1 \right| = \left| \frac{\beta - \beta_0 + (\lambda - \lambda_0)x^2}{\beta_0 + \lambda_0 x^2} \right| \leq \frac{1}{\beta_0} [|\beta - \beta_0| + |\lambda - \lambda_0|A^2].$$

Therefore we have $\lim_{(\beta, \lambda) \rightarrow (\beta_0, \lambda_0)} \sup_{|x| \leq A} |b_{\beta, \lambda}(x)^2 - 1| = 0$. Since $1 + b_{\beta, \lambda}(x) \geq 1$, we have $|b_{\beta, \lambda}(x) - 1| \leq |b_{\beta, \lambda}(x)^2 - 1|$, so that the first convergence of Lemma is proved. The second one holds since $\sup_{|x| \leq A} |a_\alpha(x)| \leq A |\alpha - \alpha_0| / \sqrt{\beta_0}$. \square

The following lemma is an easy extension of the classical continuity property of the map $f \mapsto f(\cdot + a)$ from \mathbb{R} to $\mathbb{L}^1(\mathbb{R})$.

Lemma B.3 For any $f \in \mathbb{L}^1(\mathbb{R})$, we have

$$\lim_{(a,b) \rightarrow (0,1)} \int_{\mathbb{R}} |f(a+bz) - f(z)| dz = 0.$$

Proof. First, if $g \in \mathcal{C}_K(\mathbb{R})$, then the desired convergence follows from Lebesgue's theorem. Second, if $f \in \mathbb{L}^1(\mathbb{R})$, then we have for every $g \in \mathcal{C}_K(\mathbb{R})$ and for every $(a, b) \in \mathbb{R}^2$ such that $b \geq 1/2$

$$\begin{aligned} & \int_{\mathbb{R}} |f(a+bz) - f(z)| dz \\ & \leq \int_{\mathbb{R}} |f(a+bz) - g(a+bz)| dz + \int_{\mathbb{R}} |g(a+bz) - g(z)| dz + \int_{\mathbb{R}} |g(z) - f(z)| dz \\ & = \frac{1}{b} \int_{\mathbb{R}} |f(y) - g(y)| dy + \int_{\mathbb{R}} |g(a+bz) - g(z)| dz + \int_{\mathbb{R}} |g(z) - f(z)| dz \\ & \leq \int_{\mathbb{R}} |g(a+bz) - g(z)| dz + 3\|f - g\|_{\mathbb{L}^1(\mathbb{R})}. \end{aligned}$$

Then we conclude by using the density of $\mathcal{C}_K(\mathbb{R})$ in $\mathbb{L}^1(\mathbb{R})$. \square

C Proof of (37a)–(37b) under the assumption $\mathbb{E}[L(\vartheta_1)^a] < 1$

Below we assume that $\kappa_{1,a} = \mathbb{E}[L(\vartheta_1)^a]^{1/a} < 1$, and we prove Properties (37a) and (37b) of Proposition 4.1 with explicit constants. Under the general assumption $\widehat{\kappa}_a < 1$ of Conditions $(\widehat{\mathcal{C}}_a)$, the proof of (37a)–(37b) is similar (replace P by P^N with N such that $\mathbb{E}[L(F_{\vartheta_N: \vartheta_1})^a] < 1$).

That the constant ξ in Proposition 4.1 is finite can be easily deduced from the drift inequality (43) which holds here with $P_{\theta_0} = P$, $\Theta = \{\theta_0\}$, and with $\kappa_{1,a}$ in place of κ_a . Now let us introduce some notations. If μ is a probability measure on \mathbb{X} and $X_0 \sim \mu$, we make a slight abuse of notation in writing $\{X_n^\mu\}_{n \in \mathbb{N}}$ for the associated IFS given in Definition 1.1. We simply write $\{X_n^x\}_{n \in \mathbb{N}}$ when $\mu := \delta_x$ is the Dirac mass at some $x \in \mathbb{X}$. We denote by \mathcal{M}_a the set of all the probability measures μ on \mathbb{X} such that $\|\mu\|_a := (\int_{\mathbb{X}} V_a(y) d\mu(y))^{1/a} < \infty$. Finally, for $n \in \mathbb{N}$ and for any probability measures μ_1 and μ_2 on \mathbb{X} , define:

$$\Delta_n(\mu_1, \mu_2) := d(X_n^{\mu_1}, X_n^{\mu_2}) (p(X_n^{\mu_1}) + p(X_n^{\mu_2}))^{a-1}.$$

Lemma C.1 We have: $\forall n \geq 1, \forall (\mu_1, \mu_2) \in \mathcal{M}_a \times \mathcal{M}_a$

$$\mathbb{E}[\Delta_n(\mu_1, \mu_2)] \leq \xi^{\frac{a-1}{a}} \kappa_{1,a}^n \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})] (\|\mu_1\|_a + \|\mu_2\|_a)^{a-1}. \quad (48)$$

Furthermore we have for all $f \in \mathcal{L}_a$:

$$\mathbb{E}[|f(X_n^{\mu_1}) - f(X_n^{\mu_2})|] \leq \xi^{\frac{a-1}{a}} m_a(f) \kappa_{1,a}^n \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})] (\|\mu_1\|_a + \|\mu_2\|_a)^{a-1}. \quad (49)$$

Proof. Note that $X_n^\mu = F_{\vartheta_n: \vartheta_1} X_0^\mu$ from Definition 1.1. If $a := 1$, then (48) follows from the independence of the ϑ_n 's and from the definition of $L(v)$ and $\kappa_{1,a}$. Now assume that $a \in (1, +\infty)$. Without loss of generality, one can suppose that the sequence $\{\vartheta_n\}_{n \geq 1}$ is independent from $(X_0^{\mu_1}, X_0^{\mu_2})$. Also note that, if $\mu \in \mathcal{M}_a$, then we have

$$\mathbb{E}[p(X_n^\mu)^a] = \int_{\mathbb{X}} (P^n V_a)(x) d\mu(x) \leq \xi \|\mu\|_a^a.$$

From Holder's inequality (use $1 = 1/a + (a-1)/a$), we obtain

$$\begin{aligned} \mathbb{E}[\Delta_n(\mu_1, \mu_2)] &= \mathbb{E}\left[d(F_{\vartheta_n: \vartheta_1} X_0^{\mu_1}, F_{\vartheta_n: \vartheta_1} X_0^{\mu_2}) (p(X_n^{\mu_1}) + p(X_n^{\mu_2}))^{a-1}\right] \\ &\leq \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})] \mathbb{E}\left[L(\vartheta_n : \vartheta_1) (p(X_n^{\mu_1}) + p(X_n^{\mu_2}))^{a-1}\right] \\ &\leq \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})] \mathbb{E}[L(\vartheta_n : \vartheta_1)^a]^{\frac{1}{a}} \mathbb{E}[(p(X_n^{\mu_1}) + p(X_n^{\mu_2}))^a]^{\frac{a-1}{a}} \\ &\leq \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})] \mathbb{E}[L(\vartheta_1)^a]^{\frac{n}{a}} \xi^{\frac{a-1}{a}} (\|\mu_1\|_a + \|\mu_2\|_a)^{a-1}. \end{aligned}$$

This proves (48). Property (49) follows from (48) and the definition of $m_a(f)$. \square

Now we prove (38). Property (49), applied to $\mu_1 := \delta_x$ and $\mu_2 := \pi$ gives

$$\begin{aligned} |P^n f(x) - \pi(f)| &\leq \mathbb{E}[|f(X_n^x) - f(X_n^\pi)|] \\ &\leq \xi^{\frac{a-1}{a}} m_a(f) \kappa_{1,a}^n \mathbb{E}[d(x, X_0^\pi)] (\|\delta_x\|_a + \|\pi\|_a)^{a-1}. \end{aligned}$$

Next observe that $\|\delta_x\|_a = p(x)$ and

$$\mathbb{E}[d(x, X_0^\pi)] \leq \mathbb{E}[d(x, x_0) + d(x_0, X_0^\pi)] \leq p(x) + \pi(d(x_0, \cdot)) \leq p(x) \|\pi\|_1.$$

Hence $\mathbb{E}[d(x, X_0^\pi)] (\|\delta_x\|_a + \|\pi\|_a)^{a-1} \leq p(x)^a \|\pi\|_1 (1 + \|\pi\|_a)^{a-1}$. Thus, under the assumption $\mathbb{E}[L(\vartheta_1)^a] < 1$, Property (37a) (namely (38)) is proved.

Finally, to prove (37b) under the assumption $\mathbb{E}[L(\vartheta_1)^a] < 1$, it remains to study $m_a(P^n f)$ for $f \in \mathcal{L}_a$. Inequality (49) applied to $\mu_1 := \delta_x$ and $\mu_2 := \delta_y$ gives:

$$\forall f \in \mathcal{L}_a, \quad |P^n f(x) - P^n f(y)| \leq \xi^{\frac{a-1}{a}} m_a(f) \kappa_{1,a}^n d(x, y) (p(x) + p(y))^{a-1}.$$

Thus $m_a(P^n f) \leq \xi^{\frac{a-1}{a}} m_a(f) \kappa_{1,a}^n$. Since $m_a(1_X) = 0$, this gives

$$m_a(P^n f - \pi(f)1_X) \leq \xi^{\frac{a-1}{a}} m_a(f) \kappa_{1,a}^n.$$

Combining the last inequality with (38) gives (37b).

References

- [Als03] G. Alsmeyer. On the Harris recurrence of iterated random Lipschitz functions and related convergence rate results. *J. Theoret. Probab.*, 16(1):217–247, 2003.

- [Bax05] P. H. Baxendale. Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.*, 15(1B):700–738, 2005.
- [Ben98] M. Benda. A central limit theorem for contractive stochastic dynamical systems. *J. Appl. Probab.*, 35:200–205, 1998.
- [BRR01] L. Breyer, G. O. Roberts, and J. S. Rosenthal. A note on geometric ergodicity and floating-point roundoff error. *Statist. Probab. Lett.*, 53(2):123–127, 2001.
- [DM12] R. Douc and E. Moulines. Asymptotic properties of the maximum likelihood estimation in misspecified hidden Markov models. *Ann. Statist.*, 40(5):2697 – 2732, 2012.
- [DMPS18] R. Douc, E. Moulines, P. Priouret, and P. Soulier. *Markov chains*. Springer Series in Operations Research and Financial Engineering. Springer, 2018.
- [Duf97] M. Duflo. *Random Iterative Models*. Applications of Mathematics. Springer-Verlag, Berlin Heidelberg, 1997.
- [FHL13] D. Ferré, L. Hervé, and J. Ledoux. Regular perturbation of V -geometrically ergodic Markov chains. *J. Appl. Probab.*, 50(1):184–194, 2013.
- [GHL11] D. Guibourg, L. Hervé, and J. Ledoux. Quasi-compactness of Markov kernels on weighted-supremum spaces and geometrical ergodicity. arXiv : 1110.3240, 2011.
- [GP14] L. Galtchouk and S. Pergamenschikov. Geometric ergodicity for classes of homogeneous Markov chains. *Stochastic Process. Appl.*, 124(10):3362–3391, 2014.
- [GW98] P. E. Greenwood and W. Wefelmeyer. Maximum likelihood estimator and Kullback-Leibler information in misspecified Markov chain models. *Theory Probab. Appl.*, 1998.
- [Hen93] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.*, 118:627–634, 1993.
- [Her08] L. Hervé. Quasi-compactness and mean ergodicity for Markov kernels acting on weighted supremum normed spaces. *Ann. Inst. H. Poincaré Probab. Statist.*, 44(6):1090–1095, 2008.
- [HH01] H. Hennion and L. Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, volume 1766 of *Lecture Notes in Math*. Springer, 2001.
- [HL14a] L. Hervé and J. Ledoux. Approximating Markov chains and V -geometric ergodicity via weak perturbation theory. *Stochastic Process. Appl.*, 124(1):613–638, 2014.
- [HL14b] L. Hervé and J. Ledoux. Spectral analysis of Markov kernels and application to the convergence rate of discrete random walks. *Adv. in Appl. Probab.*, 46(4):1036–1058, 2014.

- [HM11] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In Robert Dalang, Marco Dozzi, and Francesco Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117, Basel, 2011. Springer Basel.
- [KL99] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze (4)*, XXVIII:141–152, 1999.
- [LP83] É. Le Page. Théorèmes de renouvellement pour les produits de matrices aléatoires. Équations aux différences aléatoires. In *Séminaires de probabilités Rennes 1983*, Publ. Sémin. Math., page 116. Univ. Rennes I, Rennes, 1983.
- [LT96] R. B. Lund and R. L. Tweedie. Geometric convergence rates for stochastically ordered Markov chains. *Math. Oper. Res.*, 21(1):182–194, 1996.
- [MR89] X. Milhaud and A. Raugi. étude de l'estimateur du maximum de vraisemblance dans le cas d'un processus autorégressif : convergence, normalité asymptotique, vitesse de convergence. *Ann. Inst. H. Poincaré Probab. Statist.*, 25:383–428, 1989.
- [MS10] M. Meitz and P. Saikkonen. A note on the geometric ergodicity of a nonlinear AR-ARCH model. *Statist. Probab. Lett.*, 80(7):631–638, 2010.
- [MT93] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag London Ltd., London, 1993.
- [MT94] S. P. Meyn and R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Probab.*, 4:981–1011, 1994.
- [RRS98] G. O. Roberts, J. S. Rosenthal, and P. O. Schwartz. Convergence properties of perturbed Markov chains. *J. Appl. Probab.*, 35(1):1–11, 1998.
- [RS18] D. Rudolf and N. Schweizer. Perturbation theory for Markov chains via Wasserstein distance. *Bernoulli*, 24(4A):2610–2639, 2018.
- [Sch47] H. Scheffé. A useful convergence theorem for probability distributions. *Ann. Math. Statist.*, 18(3):434–438, 1947.
- [Wu04] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probab. Theory Related Fields*, 128(2):255–321, 2004.