Robustness of iterated function systems of Lipschitz maps

Loïc HERVÉ and James LEDOUX *

Wednesday 7$^{\text{th}}$ September, 2022

Abstract

Let $\{X_n\}_{n\in\mathbb{N}}$ be an $\mathbb{X}$-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 independent and identically distributed random variables with common probability distribution ν, $F(\cdot,\cdot)$ is Lipschitz continuous in the first variable and X_0 is independent of $\{\vartheta_n\}_{n\geq 1}$. 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 autoregressive conditional heteroscedastic errors, and for IFS under roundoff error or under thresholding/truncation. Moreover, we provide a general set of assumptions covering the classical Feller-type hypotheses, for an IFS to be a V-geometrical ergodic process. An accurate bound for the rate of convergence is also provided.

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 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*

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

X_0 be a $\mathbb{X}$-valued r.v. which is assumed to be independent of the sequence $\{\vartheta_n\}_{n\geq 1}$. Finally let $F : (\mathbb{X}\times\mathbb{V}, \mathcal{X}\otimes\mathcal{V}) \to (\mathbb{X}, \mathcal{X})$ 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:

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

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 associated with $V_a(\cdot)$, that is

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

Note that $(\mathcal{B}_0, |\cdot|_0)$ corresponds to the Banach space of complex-valued bounded measurable functions on $\mathbb{X}$ equipped with the supremum norm. The total variation distance between two probability distributions (p.d.) 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 $\mu_i(f) := \int_{\mathbb{X}} f(x)\, d\mu_i(x)$, $i = 0, 1$. 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:

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

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

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

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

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

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

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

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

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

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

Example 1.1 (Vector AutoRegressive model (VAR)) *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. 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*

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

Here $F(x,v) := Ax+v$ where $A = (a_{ij})$ is a fixed real $q\times q$-matrix. This is called a vector or multivariate autoregressive model. We have $L_F(v) = \|A\|$ where $\|A\|$ denotes the induced norm of A corresponding to $\|\cdot\|$, and $d\big(0,F(0,v)\big) = \|v\|$. Consequently, Condition $(\mathcal{C}_a)$ holds for some $a \in [1,+\infty)$ provided that

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

Moreover, if ϑ_1 has a probability density function (p.d.f.) on $\mathbb{R}^q$, then P is V_a-geometrically ergodic. More precisely Inequality (4) holds for any $\rho_a \in (\|A\|,1)$, see Remark 4.2.

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

Definition 1.2 (Parametric perturbation of IFS) *Let us introduce the parameter $\theta := (\xi,\gamma)$ taking values in a subset Θ of some metric space. Let $F_\xi : (\mathbb{X}\times\mathbb{V}, \mathcal{X}\otimes\mathcal{V}) \to (\mathbb{X},\mathcal{X})$ and let $\{\vartheta_n^{(\gamma)}\}_{n\geq 1}$ be a sequence of $\mathbb{V}$-valued r.v. both satisfying the assumptions of Definition 1.1. The common parametric p.d. of $\{\vartheta_n^{(\gamma)}\}_{n\geq 1}$ is denoted by ν_γ. For any $\theta \in \Theta$, the process $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ is the $\mathbb{X}-$valued IFS of Lipschitz maps given by*

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

The transition kernel of the Markov chain $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ is denoted by P_θ, 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 $\{X_n^{(\theta_0)}\}_{n\in\mathbb{N}}$ with $\theta_0 \in \mathring{\Theta}$, where $\mathring{\Theta}$ denotes the interior of Θ. Next, pick $\theta_0 \in \mathring{\Theta}$ and let us introduce the following assumptions:

- (H$_1$) There exists $a \geq 1$ such that P_{θ_0} is V_a-geometrically ergodic with stationary distribution denoted by π_{θ_0}, that is P_{θ_0} satisfies (4) for some $\rho_a \in (0,1)$ and $C_a > 0$.
- (H$_2$) $M_a := \sup_{\theta\in\Theta} \mathbb{E}\left[d\left(x_0, F_\xi\big(x_0, \vartheta_1^{(\gamma)}\big)\right)^a\right]^{1/a} < \infty$.
- (H$_3$) $\kappa_a := \sup_{\theta\in\Theta} \mathbb{E}\left[L_{F_\xi}\big(\vartheta_1^{(\gamma)}\big)^a\right]^{1/a} < 1$.
- (H$_4$) $\Delta_\theta := \|P_\theta - P_{\theta_0}\|_{0,a} \xrightarrow[\theta\to\theta_0]{} 0$, where $\|P_\theta - P_{\theta_0}\|_{0,a} := \sup\limits_{f\in\mathcal{B}_0,\ |f|_0\leq 1} |P_\theta f - P_{\theta_0} f|_a$.

Assumption (H_1) is the natural starting point for our perturbation issues. Note that the assumptions (H_2)–(H_3) are nothing else but the uniform version with respect to θ of Condition $(\mathcal{C}_a)$. As a by-product it follows from (H_2)–(H_3) 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):

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

This implies that, for every $\theta \in \Theta$, P_θ admits an invariant probability measure denoted by π_θ. The following natural questions are much more difficult to address: Is the map $\theta \mapsto \pi_\theta$ continuous with respect to the total variation distance? Under Assumption (H_1), do the perturbed transition kernels P_θ satisfy the V_a-geometrical ergodicity when θ is close to θ_0? In our context of parametric perturbation of IFS, these questions are addressed in the following theorem using the results of [FHL13, HL14a, RS18].

Theorem 1.1 *Under the assumptions* (H_1)–(H_4)*, the following properties hold.*

(P_1) *For every* $\rho \in (\rho_a, 1)$ *there exist an open neighbourhood* $\mathcal{V}_{\theta_0}$ *of* θ_0 *and a positive constant* R *such that*

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

(P_2) $\lim_{\theta \to \theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{TV} = 0$. *More precisely:*

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

provided that $\Delta_\theta \in (0, \exp(-1))$*, where the constants* ρ_a*,* C_a*,* δ_a *and* K_a *are given in* (H_1) *and* (10)*, and* $D_a = 2C_a(K_a+1)$*.*

(P_3) *We have for every* $n \geq 1$ *and for every* $\theta \in \Theta$

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

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

In the general framework of V-geometrically ergodic Markov chains, Property (P_1) and the first statement in (P_2) are proved in [FHL13, Th. 1] by using the Keller-Liverani perturbation theorem [KL99][1]. Inequality (11) in (P_2) 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 (P_3). Property (P_3) is proved in [RS18, Th. 3.2] by using the Wasserstein

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

distance associated with a suitable metric on $\mathbb{X}$ defined from the Lyapunov function V, as introduced in [HM11]. The goal of this work is to present various 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 (H_4) is well suited to such a case. This last claim 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 p.d. 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

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

where $\{\vartheta_n^{(\gamma)}\}_{n\geq 1}$ is a sequence of $\mathbb{V}$-valued i.i.d. r.v, with common parametric p.d. 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

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

It follows that

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

Hence (H_4) *is satisfied provided that* $\lim \|\nu_\gamma - \nu_0\|_{TV} = 0$.

In Section 2, a second application of Theorem 1.1, which again illustrates the interest of (H_4), is provided for the real-valued Markov chain $X_n := \alpha X_{n-1} + \sigma(X_{n-1})\vartheta_n$, for which all the data α, $\sigma(\cdot)$ and the p.d. of the noise ϑ_1 are perturbed. This Markov chain is called an AutoRegressive process of order 1 with AutoRegressive Conditional Heteroscedastic errors of order 1 (AR(1)-ARCH(1)). Such autoregressive models with conditional heteroscedastic errors were introduced to allow the conditional variance of a time series models to depend on past information. It turns out that such processes fit very well to many types of econometrics and financial data where stochastic volatility must be taken into account (e.g. see [Tsa10]). Note that the perturbation results of Section 2 can be extended to multivariate AR(p)-ARCH(q) processes with any order (p,q) (see [MS10]) thanks to the material provided in Section 5. In Section 3, a third application is presented in the framework of roundoff errors. In applied mathematics, any analytic material must be run on computer to get practical answers. This concerns simulation, approximation, numerical schemes and so on. Thus, when a Markov model is implemented on computer, the original transition kernel P is replaced with a perturbed one, say $\tilde{P}$, and their difference may have a great impact on the results. Such changes in computer simulations induced by floating point roundoff error were discussed in [RRS98, BRR01]. In this case, the perturbed transition kernel takes the form $\tilde{P}(x,A) := P(x, h^{-1}(A))$, where P is the transition kernel of a fixed IFS and where $h : \mathbb{X} \to \mathbb{X}$ is close to the identity map. The weak continuity assumption (H_4) 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_ξ in (9) is fixed in Example 1.2, so that we did not have to divide by $V(x)$ to prove (H_4). Indeed the inequality $\|P_{\nu_\gamma} - P_{\nu_0}\|_{0,0} \leq \|\nu_\gamma - \nu_0\|_{TV}$ in Example 1.2 is directly obtained and it automatically gives (H_4). When F_ξ in (9) is not fixed as in Sections 2-3, the division by $V(x)$ in the definition $\|\cdot\|_{0,a}$ must be done to investigate (H_4). In Section 4 we propose a new approach to prove the V_a-geometrical ergodicity of IFS of Lipschitz maps under Condition ($\mathcal{C}_a$), together with a bound on the spectral gap of P (i.e. the infimum bound of the positive real numbers ρ_a satisfying (4)). In Section 5 further applications of Theorem 1.1 are presented. The goal of Subsection 5.1 is to show that the arguments developed in Sections 2 and 3 for specific IFS of Lipschitz maps naturally extend to more general IFS. In Subsection 5.2 we apply Theorem 1.1 to the case where the function F and the p.d.f. ν of $\{\vartheta_n\}_{n\geq 1}$ in (1) are perturbed by thresholding and by truncation respectively.

The perturbation theory for Markov chains is a natural issue which has been widely developed in the last decades. As mentioned in [SS00, p. 1126] (see also [FHL13]), the strong continuity assumption introduced in [Kar86], which involves the iterates of both perturbed and unperturbed transition kernels, does not hold in general for V-geometrically ergodic Markov chains, excepted for particular perturbed transition kernels (e.g. when $P_\theta = P_0 + \theta D$, see [AANQ04]) and for uniformly ergodic Markov chains (i.e. when (4) holds with $a = 0$), see [Mit05, MA10, AFEB16, JMMD]. Similar questions arise for dynamical systems, and [Kel82, p. 316] seems to be the first work which introduced a weaker continuity assumption using two norms as in (H_4) (instead of a single one in the standard theory). Then, the Keller-Liverani perturbation theorem [KL99, Bal00, Liv04] has proved to be very powerful for studying the behaviour of the Sinai-Ruelle-Bowen measures of certain perturbed dynamical systems (e.g. see [Bal00, Th. 2.10] and [GL06, Th. 2.8]). In the context of V-geometrically ergodic Markov chains, Keller's approach is used in [SS00] and the Keller-Liverani theorem is applied in [FHL13, HL14a]. The recent works [RS18, MARS20, and references therein] combine Keller's approach and the elegant idea of [HM11] using the Wasserstein distance associated with a suitable metric on $\mathbb{X}$ defined from the Lyapunov function V. Perturbation issues have been also investigated in the framework of roundoff errors [RRS98, BRR01] (see Section 3) and in the special case of reversible transition kernels as in Markov Chain Monte Carlo methods, e.g. see [MALR16, NR] and the references therein. The purpose of this paper is to show that the material developed in [FHL13, HL14a, RS18] is very well suited to the perturbation of general IFS. In the IFS context, Assumption (H_4) has so far only been investigated in [FHL13, RS18] for the perturbation of univariate AR(1) processes $X_n := \alpha X_{n-1} + \vartheta_n$ with respect to the contracting coefficient α. Our work shows that Assumption (H_4) allows us to deal with perturbation schemes of the general IFS (1) with respect to both function F and p.d. of ϑ_1.

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 growths is large (e.g. see [GW98, DM12] in the Markov context).

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

According to Definition 1.2, we consider the following perturbed AR(1)-ARCH(1) real-valued $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ defined by: $X_0^{(\theta)}$ is a given real-valued r.v. and

$$\forall n \geq 1, \quad X_n^{(\theta)} := F_\xi(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}) \tag{13}$$

where $F_\xi(x,v) = \alpha x + v\,(\beta + \lambda x^2)^{1/2}$ with constants $\alpha \in \mathbb{R}, \beta > 0, \lambda > 0$, $\{\vartheta_n^{(\gamma)}\}_{n\in\mathbb{N}}$ has common p.d.f. ν_γ and is independent of $X_0^{(\theta)}$. Therefore, 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

$$p_\theta(x,y) := \big(\beta + \lambda x^2\big)^{-1/2}\, \nu_\gamma\bigg(\frac{y - \alpha x}{\big(\beta+\lambda x^2\big)^{1/2}}\bigg). \tag{14}$$

Next, let us report the following observations with respect to basic quantities required in the assumptions (H_2) and (H_3) .

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

$$L_{F_\xi}(\vartheta_1) = \max\big(|\alpha - \sqrt{\lambda}\vartheta_1|; |\alpha + \sqrt{\lambda}\vartheta_1|\big). \tag{15}$$

   Hence, the real number κ_a in (H_3) is

$$\kappa_a = \sup_{\theta\in\Theta}\bigg(\int_{\mathbb{R}} \max\big(|\alpha - \sqrt{\lambda}v|; |\alpha+\sqrt{\lambda}v|\big)^a \nu_\gamma(v)dv\bigg)^{1/a}. \tag{16}$$

2. The real number M_a in (H_2) is given by

$$M_a := \sup_{\theta\in\Theta} \sqrt{\beta}\,\mathbb{E}\Big[\big|\vartheta_1^{(\gamma)}\big|^a\Big]^{1/a} = \sup_{\theta\in\Theta}\sqrt{\beta}\,\bigg(\int_{\mathbb{R}} |v|^a \nu_\gamma(v)dv\bigg)^{1/a}. \tag{17}$$

   Note that, if β lies in a compact set, then $M_a < \infty$ under the following uniform moment condition for the p.d.f. 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.

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

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

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

   is positive on a subset of $[-r,r]$ which has a positive Lebesgue's measure.

(b) $M_a < \infty$, where M_a is given in (17).

(c) $\kappa_a < 1$, where κ_a is given in (16).

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

Proposition 2.1 *Under the conditions* (H'$_{123}$)–(H'$_4$) *for the AR(1)-ARCH(1) processes given in (13), the assertions* (P$_1$)-(P$_2$)-(P$_3$) *of Theorem 1.1 hold.*

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 the conditions (H'$_{123}$)-(b)-(c) imply that the assumptions (H$_2$) and (H$_3$) of Theorem 1.1 hold. Moreover, use (47) 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 $\varphi_{r,\theta_0}(dy) = g_{\theta,r}(y)\, dy$. In others words, $S = [-r_a, r_a]$ 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 following lemma asserts that Assumption (H$_4$) holds true under Condition (H'$_4$), so that the proof is complete. □

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

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

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

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

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

$$|x| > B \implies \forall \theta \in \Theta,\ \frac{|(P_\theta f)(x) - (P_{\theta_0}) f(x)|}{V_a(x)} < \frac{\varepsilon}{2}. \tag{18}$$

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

$$\forall A > 0, \quad \lim_{\theta \to \theta_0} \sup_{|x| \leq A} \frac{\int_{\mathbb{R}} |p_\theta(x,y) - p_{\theta_0}(x,y)| dy}{V_a(x)} = 0. \tag{19}$$

Indeed, (18) and (19) with $A = B$ ensure that $\|P_\theta - P_{\theta_0}\|_{0,a} < \varepsilon$ when θ is sufficiently close to θ_0. Let us prove (19). It follows from (14) that

$$
\begin{aligned}
&\int_{\mathbb{R}} |p_\theta(x,y) - p_{\theta_0}(x,y))| dy \\
&= \int_{\mathbb{R}} \left| \left(\beta + \lambda x^2\right)^{-1/2} \nu_\gamma \left(\frac{y - \alpha x}{\left(\beta + \lambda x^2\right)^{1/2}} \right) - \left(\beta_0 + \lambda_0 x^2\right)^{-1/2} \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{\left(\beta_0 + \lambda_0 x^2\right)^{1/2}} \right) \right| dy \\
&\leq \int_{\mathbb{R}} \left(\beta + \lambda x^2\right)^{-1/2} \left| \nu_\gamma \left(\frac{y - \alpha x}{\left(\beta + \lambda x^2\right)^{1/2}} \right) - \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{\left(\beta_0 + \lambda_0 x^2\right)^{1/2}} \right) \right| dy \qquad (20) \\
&\quad + \int_{\mathbb{R}} \nu_{\gamma_0} \left(\frac{y - \alpha_0 x}{\left(\beta_0 + \lambda_0 x^2\right)^{1/2}} \right) \left| \left(\beta + \lambda x^2\right)^{-1/2} - \left(\beta_0 + \lambda_0 x^2\right)^{-1/2} \right| dy. \qquad (21)
\end{aligned}
$$

First, using the change of variables $z = (y - \alpha x)/(\beta + \lambda x^2)^{1/2}$ in the integral (20) and the triangle inequality we obtain

$$
\begin{aligned}
(20) &= \int_{\mathbb{R}} \left| \nu_\gamma(z) - \nu_{\gamma_0} \left(\left(\frac{\beta + \lambda x^2}{\beta_0 + \lambda_0 x^2} \right)^{1/2} z + x \, \frac{\alpha - \alpha_0}{(\beta_0 + \lambda_0 x^2)^{1/2}} \right) \right| dz \\
&\leq \int_{\mathbb{R}} |\nu_\gamma(z) - \nu_{\gamma_0}(z)| \, dz + \int_{\mathbb{R}} \left| \nu_{\gamma_0}(z) - \nu_{\gamma_0}\left(b_{\beta,\lambda}(x) z + a_\alpha(x)\right) \right| dz \qquad (22)
\end{aligned}
$$

where

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

The first integral in (22) does not depend on x and is equal to $\|\nu_\gamma - \nu_{\gamma_0}\|_{\mathbb{L}^1(\mathbb{R})}$ which converges to 0 when $\gamma \to \gamma_0$ from the assumption. Now let $A > 0$ be fixed. It follows from Lemma B.2 that $\lim_{(\beta,\lambda) \to (\beta_0,\lambda_0)} \sup_{|x| \leq A} |b_{\beta,\lambda}(x) - 1| = 0$ and $\lim_{\alpha \to \alpha_0} \sup_{|x| \leq A} a_\alpha(x) = 0$. Then under Condition (H'$_4$), Lemma B.3 allows us to conclude that the second integral in (22) is such that

$$
\lim_{(\alpha,\beta,\lambda) \to (\alpha_0,\beta_0,\lambda_0)} \sup_{|x| \leq A} \int_{\mathbb{R}} \left| \nu_{\gamma_0}(z) - \nu_{\gamma_0}\left(b_{\beta,\lambda}(x) z + a_\alpha(x)\right) \right| dz = 0.
$$

Second, let us consider the integral (21). We must show that the supremum of this integral on $x \in [-A, A]$ converges to 0 when $(\beta, \lambda) \to (\beta_0, \lambda_0)$. We obtain for any $x \in \mathbb{R}$ such that $|x| \leq A$:

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

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

Remark 2.1 *If the p.d.f. ν_{γ_0} of the noise of the unperturbed AR(1)-ARCH(1) process is continuous on $\mathbb{R}$, then Condition* (H'$_{123}$)-(a) *(stated to prove* (H$_1$)*) can be omitted in Proposition 2.1. Actually, under the condition $\int_{\mathbb{R}} L_{F_{\xi_0}}(v)^a\, \nu_{\gamma_0}(v)\, dv < 1$ which is contained in* (H'$_{123}$)-(c)*, Assumption* (H$_1$) *holds with any real number ρ_a (and the associated constant C_a) such that*

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

Indeed the kernel $p_{\theta_0}(x,y)$ given by (14) is continuous, so that Remark 4.3 and Proposition 4.2 ensure that, under the conditions (H'$_{123}$)-(b) *and* (H'$_{123}$)-(c)*, P_{θ_0} satisfies* (4) *for any ρ_a satisfying the above condition. In other words, if the p.d.f. ν_{γ_0} is continuous on $\mathbb{R}$, then only the conditions* (H'$_{123}$)-(b) *and* (H'$_{123}$)-(c) *with $\Theta = \{\theta_0\}$ are useful to obtain* (H$_1$).

Remark 2.2 *It is well-known from Scheffé's lemma [Sch47] that the almost everywhere pointwise convergence of the p.d.f. ν_γ to the p.d.f. ν_{γ_0} when $\gamma \to \gamma_0$ provides the $\mathbb{L}^1(\mathbb{R})$-convergence required in* (H'$_4$).

3 Robustness of IFS under roundoff error

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

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

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/roundoff transition kernels associated with $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ (or $(h_\theta)_{\theta\in\Theta}$) are given by

$$\forall f \in \mathcal{B}_0,\ \forall x \in \mathbb{X}, \quad (P_\theta f)(x) = P(f\circ h_\theta)(x) = \int_{\mathbb{V}} f\big((h_\theta \circ F)(x,v)\big)\, d\nu(v). \tag{23}$$

When the 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 $h_\theta \to id$ uniformly on $\mathbb{R}^q$ when $\theta \to \theta_0$ (i.e. $\forall x \in \mathbb{R}^q$, $\|h(x) - x\| \le \varepsilon(\theta)$ with $\lim_{\theta\to\theta_0} \varepsilon(\theta) = 0$). 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. In [BRR01], the authors introduced the following weaker assumption $\|h(x) - x\| \le \varepsilon(\theta)\|x\|$ with $\lim_{\theta\to\theta_0} \varepsilon(\theta) = 0$, 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 model $\{X_n\}_{n\in\mathbb{N}}$, but more importantly the sensitivity of the p.d. of $X_n^{(\theta)}$ and of the stationary distribution of $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ when $\theta \to \theta_0$ 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 $h_0 = id$. 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 (23))

$$\forall f \in \mathcal{B}_0,\ \forall x \in \mathbb{X}, \quad (P_\theta f)(x) = \int_{\mathbb{R}^q} f\big(h_\theta(Ax+v)\big)d\nu(v). \tag{24}$$

If $g : \mathbb{R}^q \to \mathbb{R}^q$ 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. For the sake of simplicity 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 $\mathbb{E}\big[\|\vartheta_1\|^a\big] < \infty$.

(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 monotonicity-type 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. $\forall z \in \mathbb{R}^q,\ \lim_{\theta\to 0} g_\theta(z) = z$.

iii. $\sup_{\theta\in\Theta} \|\nabla g_\theta\|_\infty < \infty$, and $\lim_{\theta\to 0} \nabla g_\theta = id$ uniformly on each balls of $\mathbb{R}^q$ centred at 0, that is:

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

Proposition 3.1 *Under the conditions* (H"$_1$)–(H"$_4$) *for a VAR process as defined in Example 1.1, the assertions* (P$_1$)–(P$_3$) *of Theorem 1.1 hold for every real number* $\rho_a \in (\|A\|, 1)$ *(and associated constant* C_a*).*

Remark 3.1 *Condition* (H"$_4$)-(b) *focuses 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* $\lim_{\theta\to 0} g_\theta = id$ *(or* $\lim_{\theta\to 0} h_\theta = id$*) 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* $\lim_{\theta\to 0} g_\theta = id$*, nor the convergence* $\lim_{\theta\to 0} h_\theta = id$*, are uniform on* $\mathbb{R}$*.*

Proof. Recall that $\theta_0 = 0$ here. We know that Assumption (H$_1$) holds, see Remark 4.2. Next, for any $\theta \in \Theta$ and $z \in \mathbb{R}^q$, set $\Gamma_\theta(z) = |\det \nabla g_\theta(z)|$. Then, using (24), P_θ has the form

$$\forall f \in \mathcal{B}_0,\ \forall x \in \mathbb{R}^q, \quad (P_\theta f)(x) = \int_{\mathbb{R}^q} f(z)\, \nu\big(g_\theta(z) - Ax\big)\, \Gamma_\theta(z)\, dz \tag{25}$$

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 $F_\theta(x, v) := h_\theta(Ax + v)$. Then (H"$_2$) is nothing else but (H$_2$) (here $x_0 = 0$), while (H$_3$) is implied by (H"$_3$) from Taylor's inequality applied to h_θ.

Next we prove (H$_4$). For every $r > 0$, let $B(0, r) = \{z \in \mathbb{R}^q : \|z\| \le r\}$. Let $f \in \mathcal{B}_0$ be such that $|f|_0 \le 1$, 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

$$\forall x \in \mathbb{R}^q \setminus B(0, K), \quad \frac{\big|(P_\theta f)(x) - (P_0 f)(x)\big|}{V(x)} \le \frac{2}{V(x)} < \varepsilon. \tag{26}$$

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

$$\big|(P_\theta f)(x) - (P_0 f)(x)\big| \le \int_{\mathbb{R}^q} \big|\nu\big(g_\theta(z) - Ax\big)\, \Gamma_\theta(z) \, - \nu(z - Ax)\big|\, dz \tag{27}$$

since $g_0 = id$. Set $d := 2/(1 - c)$ where c is given in (H"$_4$)-(b)-i. Note that $\|Ax\| \le K$ and that (H"$_4$)-(b)-i provides: $\forall z \in \mathbb{R}^q,\ \|g_\theta(z)\| \ge (1 - c)\|z\|$. Then we have for every $z \in \mathbb{R}^q$ such that $\|z\| \ge dK$

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

It follows from (H"$_4$)-(a) that we have for every $\theta \in \Theta$

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

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 $\int_{\|z\|\ge C} \nu(d^{-1}z)\, dz \le \varepsilon/2(\gamma + 1)$ 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 $\theta \in \Theta$

$$\int_{\|z\|\ge D} \big|\nu\big(g_\theta(z) - Ax\big)\, \Gamma_\theta(z) \, - \nu(z - Ax)\big|\, dz \le (\gamma + 1) \int_{\|z\|\ge C} \nu(d^{-1}z)\, dz \ \le \frac{\varepsilon}{2}. \tag{28}$$

Now we investigate the integrand in (27) 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 $x \in B(0,K)$

$$\left|\nu\big(g_\theta(z) - Ax\big)\,\Gamma_\theta(z) \;-\nu(z-Ax)\right| \leq \gamma\left|\nu\big(g_\theta(z)-Ax\big) - \nu(z-Ax)\right| + m\left|\Gamma_\theta(z)-1\right|. \quad (29)$$

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 the conditions (H"$_4$)-(b)-ii and (H"$_4$)-(b)-iii that $\lim_{\theta\to 0} g_\theta = id$ 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 $\mathcal{V}_0$ of $\theta=0$ in $\mathbb{R}^p$ such that

$$\forall \theta \in \mathcal{V}_0,\ \forall z \in B(0,D),\ \forall x \in B(0,K), \quad \left|\nu\big(g_\theta(z)-Ax\big) - \nu(z-Ax)\right| < \frac{\varepsilon}{4\gamma\ell_D}.$$

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

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

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 (29) that

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

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

$$\forall \theta \in \mathcal{V}_0',\ \forall x \in B(0,K), \quad \int_{\|z\|\leq D} \left|\nu\big(g_\theta(z)-Ax\big)\,\Gamma_\theta(z) \;-\nu(z-Ax)\right| dz \leq \frac{\varepsilon}{2}. \quad (30)$$

We deduce from (27), (28) and (30) that

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

This inequality combined with (26) gives (H$_4$). □

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 $\mathcal{L}_a$:

$$\mathcal{L}_a := \left\{ f:\mathbb{X}\to\mathbb{C} \ :\ m_a(f) := \sup\left\{ \frac{|f(x)-f(y)|}{d(x,y)\,(p(x)+p(y))^{a-1}},\ (x,y)\in\mathbb{X}^2,\ x\neq y \right\} < \infty \right\}. \quad (31)$$

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

$$\forall f \in \mathcal{L}_a, \quad \|f\|_a := m_a(f) + |f|_a. \tag{32}$$

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 Section 1. 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:

$$F_{v_n:v_1} := F_{v_n} \circ \cdots \circ F_{v_1} \quad \text{and} \quad L(v_n : v_1) := L(F_{v_n:v_1}). \tag{33}$$

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

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

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 Condition $(\widehat{\mathcal{C}}_a)$ (counterpart of $(\mathcal{C}_a)$ in Section 1) 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.

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

$$\mathbb{E}\left[d\big(x_0, F(x_0, \vartheta_1)^a\big)\right] < \infty \tag{34a}$$

$$\widehat{\kappa}_a < 1. \tag{34b}$$

Note that Condition (34b) is equivalent to the following one

$$\exists N \in \mathbb{N}^*,\ \mathbb{E}\left[L(\vartheta_N : \vartheta_1)^a\right] < 1 \tag{35}$$

and Condition $(\mathcal{C}_a)$ in Section 1 corresponds to (34a) and to (35) 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 (36a) and (36b) are proved with explicit constants under the assumptions (34a) and (35) with $N = 1$ (i.e. $\mathbb{E}[L(\vartheta_1)^a] < 1$).

Proposition 4.1 ([Duf97, Chapter 6]) *Under Condition* $(\widehat{\mathcal{C}}_a)$*,* 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:*

$$\forall f \in \mathcal{L}_a,\ \forall n \geq 1, \quad |P^n f - \pi(f)1_{\mathbb{X}}|_a \leq c\,\kappa^n\, m_a(f) \tag{36a}$$

$$\forall f \in \mathcal{L}_a,\ \forall n \geq 1, \quad \|P^n f - \pi(f)1_{\mathbb{X}}\|_a \leq c'\,\kappa^n\, \|f\|_a. \tag{36b}$$

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

$$\forall f \in \mathcal{L}_a,\ \forall n \geq 1, \quad |P^n f - \pi(f)1_{\mathbb{X}}|_a \leq c_1\, {\kappa_{1,a}}^n\, m_a(f), \tag{37}$$

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

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

Under Condition $(\widehat{\mathcal{C}}_a)$, Property (36a) 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 (36a) (or (37)) does not provide the V_a-geometric ergodicity (4) since (36a) (or (37)) is only established for $f \in \mathcal{L}_a$. Under Condition $(\widehat{\mathcal{C}}_a)$ 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 Condition $(\widehat{\mathcal{C}}_a)$, the Markov chain $\{X_n\}_{n\in\mathbb{N}}$ 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 4.2 below. The bound (38) is the same as in [Wu04, Prop. 7.2], but the Feller-type assumptions of [Wu04] are replaced with the following one: $P^\ell : \mathcal{B}_0 \to \mathcal{B}_a$ for some $\ell \geq 1$ is compact (see Remark 4.2 for comparisons).

Proposition 4.2 *Let us assume that Condition $(\widehat{\mathcal{C}}_a)$ holds true and that $P^\ell : \mathcal{B}_0 \to \mathcal{B}_a$ 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:*

$$\rho_{V_a}(P) \leq \widehat{\kappa}_a. \tag{38}$$

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 to 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 (36a). Then Property (36a) 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 bounded linear 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 [HL14b, 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 (36b), $P_{|\mathcal{L}_a}$ 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 (36b) (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 $P^n \to \Pi$ with respect to the operator norm on $\mathcal{B}_a$, where Π is the finite rank eigen-projection on $\mathrm{Ker}(P-I) = \mathrm{Ker}(P-I)^2$. The last equality holds since P is power bounded on $\mathcal{B}_a$. Set $m := \dim \mathrm{Ker}(P-I)$. From [Wu04, Prop. 4.6] (see also [Her08, Th. 1]), there exist m linearly independent nonnegative functions $f_1, \ldots, f_m \in \mathrm{Ker}(P-I)$ and probability measures $\mu_1, \ldots, \mu_m \in \mathrm{Ker}(P^*-I)$ 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 [HL14b, 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 $\rho_{V_a}(P) \leq \widehat{\kappa}_a$. □

Remark 4.1 *Inequality (38) means that, for any real number $\rho \in (\widehat{\kappa}_a, 1)$, there exists a constant $C \equiv C_\rho$ such that*

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

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). It is worth mentioning that explicit bounds on ρ and C are also provided in [GP14] for a parametrized family of transition kernels.

Remark 4.2 *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 $K : \mathbb{X}^2 \to [0, +\infty)$ such that:*

$$\forall x \in \mathbb{X}, \quad P(x, dy) = K(x, y)\, d\eta(y). \tag{39}$$

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, if P admits a kernel as in (39), 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. Finally, note that the transition kernel P of an VAR process (see Example 1.1) is always strongly Feller. Indeed, let $f \in \mathcal{B}_0$ such that $\|f\|_0 \leq 1$. Then we have

$$\forall (x, x') \in \mathbb{R}^q \times \mathbb{R}^q, \qquad \big|(Pf)(x') - (Pf)(x)\big| \leq \int_{\mathbb{R}^q} \big|\nu(y - A(x' - x)) - \nu(y)\big|\, dy.$$

Since $t \mapsto \nu(\cdot - t)$ is continuous from $\mathbb{R}^q$ to the Lebesgue space $\mathbb{L}^1(\mathbb{R}^q)$, it follows that P is strongly Feller. Thus the V_a-geometric ergodicity of P claimed in Example 1.1 follows from Proposition 4.2. See also [Wu04, Sect. 8].

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 (39) 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 conclusions of Proposition 4.2 hold true under Condition $(\widehat{\mathcal{C}}_a)$. Indeed we have for all $(x, x') \in \mathbb{X}^2$ and for any $f \in \mathcal{B}_0$*

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

Since $K(\cdot,\cdot) \geq 0$, $\int K(\cdot, y) d\eta(y) = 1$, and $\lim_{x' \to x} K(x', y) = K(x, y)$, we deduce from Scheffé's theorem that $\lim_{x'\to x} \int_{\mathbb{X}} |K(x', y) - K(x, y)|\, d\eta(y) = 0$. 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 $P : \mathcal{B}_0 \to \mathcal{B}_1$ 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 (36a) 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 (36b) from which we deduce the bound $r_{ess}(P_{|\mathcal{L}_a}) \leq \widehat{\kappa}_a$. Actually, for any $\delta \in (\widehat{\kappa}_a^{\,a}, 1)$, the drift inequality $P^N V_a \leq \delta^N V_a + K$ with some $N \geq 1$ and $K > 0$ can be derived from Condition $(\widehat{\mathcal{C}}_a)$ 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 also [Wu04, Prop. 7.2] which provides the same bound under Feller-type assumptions.*

5 Further applications

Theorem 1.1 has been applied in Section 2 for the real-valued AR(1) with ARCH(1) errors models (see Proposition 2.1), and in Section 3 for roundoff errors of an VAR model (see Proposition 3.1). Although these applications have been presented for specific IFS, it is worth noticing that they give a general road map to investigate the issues (P_1)-(P_2)-(P_3) of Section 1 for other instances of $\mathbb{R}^q$-valued IFS, provided that the p.d. of the noise ν_γ in Definition 1.2 admits an p.d.f. with respect to Lebesgue's measure on $\mathbb{V} = \mathbb{R}^q$ 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. In Subsection 5.1 we propose two examples to support this claim. Finally in Subsection 5.2 we discuss the robustness of IFS of Lipschitz maps under perturbation by some thresholding and truncation.

5.1 A general non-linear time series model

Denoting by $\mathrm{GL}_q(\mathbb{R})$ the set of invertible real $q \times q-$matrices, consider an IFS $\{X_n\}_{n\in\mathbb{N}}$ of the form

$$\forall n \geq 1, \quad X_n = \psi(X_{n-1}) + B(X_{n-1})\,\vartheta_n \tag{40}$$

where $\psi : \mathbb{R}^q \to \mathbb{R}^q$, $B : \mathbb{R}^q \to \mathrm{GL}_q(\mathbb{R})$ and where the random variables $\{\vartheta_n\}_{n\geq 1}$ have common p.d.f. ν. If $B(x) = I_q$ for any $x \in \mathbb{R}^q$ where I_q is the identity $q \times q$-matrix, this Markov chain is called a functional-coefficient AR model. The Markov model (40) encompasses a very large class of non-linear time series models (e.g. see [MT93, Chap. 2], [Tsa10, Chap. 4]), [CP02, Cli07a, Cli07b, Cli07a, MS10, and references therein].

As a generalization of Section 2, consider the following general parametric perturbation of the $\mathbb{R}^q-$valued IFS $\{X_n\}_{n\in\mathbb{N}}$ defined in (40):

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

with some parametrized maps $\psi_\xi : \mathbb{R}^q \to \mathbb{R}^q$ and $B_\xi : \mathbb{R}^q \to \mathrm{GL}_q(\mathbb{R})$, 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

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

the following remarks are relevant to investigate the assumptions (H_1)–(H_4) of Theorem 1.1.

(R1) If the p.d.f. ν_{γ_0} of the unperturbed IFS (corresponding to some $\theta_0 = (\xi_0, \gamma_0)$), as well as the functions ψ_{ξ_0} and B_{ξ_0}, are 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 Condition ($\mathcal{C}_a$). More precisely, in this case, Assumption (H_1) holds with any real number ρ_a (and the associated constant C_a) such that

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

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

(R3) Thanks to Formula (41), Condition (H_4) holds provided that for every $A > 0$

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

since the previous integral is less than $2/(1+A)^a$ for $\|x\| > A$. Moreover the above integral on $\mathbb{R}^q$ can be decomposed on some ball of $\mathbb{R}^q$ and on its complementary in order to use uniform continuity and decay properties of the kernel $p_\theta(\cdot,\cdot)$ (see the proof of Proposition 5.1 in Appendix D).

Next, as a generalization of Section 3, consider the IFS defined by (40) under roundoff error. If $(h_\theta)_{\theta\in\Theta}$ is the roundoff family with h_θ close to $h_0 = id$ when $\theta \to 0$, 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

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

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 $\Gamma_\theta(z) := \big|\det B_\xi(x)\big|^{-1} |\det \nabla g_\theta(z)|$. Using here the kernels in (42), the remarks (R1)–(R3) then hold.

5.2 Robustness of IFS under thresholding/truncation

Here we consider $\mathbb{X} := \mathbb{R}^d$ ($d \ge 1$) equipped with the euclidean norm $\|\cdot\|$, and $\mathbb{V} := \mathbb{R}^q$ ($q \ge 1$) equipped with some norm still denoted by $\|\cdot\|$ for the sake of simplicity. Let $\{X_n\}_{n\in\mathbb{N}}$ be an IFS of Lipschitz maps

$$X_0 \in \mathbb{R}^d, \quad \forall n \ge 1, \quad X_n := F(X_{n-1}, \vartheta_n) \tag{43}$$

with $F : \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d$ and $\{\vartheta_n\}_{n\geq 1}$ satisfying the assumptions of Definition 1.1. Suppose that the p.d. of ϑ_1 is absolutely continuous with respect to Lebesgue's measure on $\mathbb{R}^q$, with p.d.f. denoted by ν. Assume that $\{X_n\}_{n\in\mathbb{N}}$ is V-geometrically ergodic. Then, a natural question is: What happens if we consider a perturbation of the IFS (43) by some thresholding and/or truncation? Such an issue may be raised as soon as a numerical implementation of the model is considered. Thus, let us investigate the robustness of the IFS (43) when thresholding the function F on the infinite set $\mathbb{X}$ and truncating the p.d.f. ν on $\mathbb{R}^q$. More precisely, for any $\xi \in (0, +\infty)$ let $\Phi_\xi : \mathbb{R}^d \to \mathbb{R}^d$ be the following thresholding function at level ξ:

$$\forall x \in \mathbb{R}^d, \quad \Phi_\xi(x) = \min\big(\frac{\xi}{\|x\|}\,,\,1\big)\,x = \begin{cases} x & \text{if } \|x\| \leq \xi \\ \xi\frac{x}{\|x\|} & \text{if } \|x\| > \xi. \end{cases} \tag{44}$$

Moreover, for any $\gamma \in (0, +\infty)$, define the truncated p.d.f. ν_γ at level γ by:

$$\forall v \in \mathbb{R}^q, \quad \nu_\gamma(v) = c_\gamma\,\nu(v)\,1_{B(0,\gamma)}(v) \qquad \text{with} \quad c_\gamma := \bigg(\int_{B(0,\gamma)} \nu(v)\,dv\bigg)^{-1}$$

where $B(0, \gamma)$ denotes the ball centred at 0 with radius γ in $\mathbb{R}^q$. Then, according to Definition 1.2, we consider the following perturbed IFS $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$

$$X_0^{(\theta)} \in \mathbb{X}, \quad \forall n \geq 1, \quad X_n^{(\theta)} := F_\xi(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}) \qquad \text{with} \quad F_\xi(x, v) := \Phi_\xi\big(F(x, v)\big) \tag{45}$$

where the sequence $\{\vartheta_n^{(\gamma)}\}_{n\geq 1}$ of $\mathbb{R}^q$-valued i.i.d. r.v. is assumed to admit the common p.d.f. ν_γ. Note that the stability of quantitative bounds for Markov chains via truncation rather than thresholding is studied in [MARS20]. However it is worth mentioning that we cannot set $\Phi_\xi(x) = 0$ for $x \in \mathbb{R}^d$ such that $\|x\| > \xi$ as in [MARS20, Subsection 3.2, Th. 9]) since the resulting perturbed process is no more an IFS of Lipschitz maps. Morevover, note that the study of $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ does not fit to the framework of Section 3. Indeed, the family $\{F_\xi, \xi > 0\}$ does not satisfy the assumptions of Section 3 since Φ_ξ is neither bijective nor differentiable. By contrast, each function Φ_ξ is 1−Lipschitz (i.e. $L(\Phi_\xi) = 1$) and this property is well suited to our perturbation approach. Therefore, the next Proposition 5.1 is stated in the general framework of Definition 1.1 up to the slight condition of absolute continuity of the p.d. of ϑ_1 with respect to Lebesgue's measure on $\mathbb{R}^q$. The proof of Proposition 5.1 is postponed to Appendix D.

Proposition 5.1 *Assume that the unperturbed IFS $\{X_n\}_{n\in\mathbb{N}}$ given in (43) satisfies Definition 1.1 with ϑ_1 having an p.d.f. on $\mathbb{R}^q$. Moreover suppose that Assumption* (H$_1$) *holds for some $a \geq 1$ and that*

$$\widetilde{M}_a := \mathbb{E}\left[\|F(0, \vartheta_1)\|^a\right]^{1/a} < \infty, \quad \widetilde{\kappa}_a := \mathbb{E}\left[L_F(\vartheta_1)^a\right]^{1/a} < 1 \quad \textit{and} \quad \mathbb{E}[\|\vartheta_1\|^a] < \infty. \tag{46}$$

Let $\kappa_a \in (\widetilde{\kappa}_a, 1)$, and let $\Theta := (0, +\infty) \times (\gamma_0, +\infty)$ with $\gamma_0 > 0$ defined by the condition:

$$\forall \gamma > \gamma_0, \quad c_\gamma \leq (\kappa_a/\widetilde{\kappa}_a)^a\,.$$

Then the perturbed IFS $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ defined by (45) with $\theta \in \Theta$ satisfies the assertions (P$_1$)–(P$_3$) *of Theorem 1.1 with*

$$\Delta_\theta \to 0 \quad \textit{when } \xi \to +\infty \textit{ and } \gamma \to +\infty.$$

More precisely, for every $\varepsilon \in (0,2)$ *define* $A_\varepsilon = 2^a\varepsilon^{-a} - 1$. *Then we have* $\Delta_\theta \leq \varepsilon$ *provided that* $\theta := (\xi, \gamma) \in \Theta$ *is such that*

$$|c_\gamma - 1| + \left(1 + \left(\frac{\kappa_a}{\widetilde{\kappa}_a}\right)^a\right)\left(\frac{\mathbb{E}\big[\|\vartheta_1\|^a\big]}{\gamma^a} + \frac{(2A_\varepsilon\widetilde{\kappa}_a)^a + (2\widetilde{M}_a)^a}{\xi^a}\right) \leq \varepsilon.$$

A Proof of (10)

Suppose that the assumptions (H$_2$)-(H$_3$) are fulfilled. Then we prove the drift inequality (10) in Section 1. In fact, for any $\kappa \in (\kappa_a, 1)$, we prove that the following strengthened inequality holds:

$$\forall \theta \in \Theta, \quad P_\theta V_a \leq \delta_a V_a + K_a 1_{[-r_a, r_a]} \tag{47}$$

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

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

It follows from the assumptions (H$_2$)-(H$_3$) that

$$\forall \theta \in \Theta,\ \forall x \in \mathbb{X}, \quad \left(\frac{(P_\theta V_a)(x)}{V_a(x)}\right)^{1/a} \leq \frac{1}{1 + d(x; x_0)} + \kappa_a + \frac{M_a}{1 + d(x; x_0)}. \tag{48}$$

For any $\kappa \in (\kappa_a, 1)$, set $r_a := (1 + M_a + \kappa_a - \kappa)/(\kappa - \kappa_a) > 0$. Then we have for every $x \in \mathbb{X}$ such that $d(x; x_0) > r_a$

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

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

$$(P_\theta V_a)(x) \leq \kappa^a\, V_a(x). \tag{49}$$

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 (48) that

$$(P_\theta V_a)(x) \leq F_a\, V_a(x) \leq F_a(1 + r_a)^a \tag{50}$$

where $F_a := (1 + \kappa_a + M_a)^a$. Finally combining (49) and (50) provides (47), thus (10), with $\delta_a := {\kappa_a}^a < 1$ and $K_a := F_a(1 + r_a)^a > 0$.

B Complements on Proposition 2.1

First we prove Property (15).

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* $v \in \mathbb{R}$

$$L(v) := \sup_{(x,y)\in\mathbb{R}^2, x\neq y} \frac{\left|F(x, v) - F(y, v)\right|}{|x - y|} = \max\left(|\alpha - \sqrt{\lambda}v|; |\alpha + \sqrt{\lambda}v|\right). \tag{51}$$

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

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

Property (51) is obvious if $v = 0$. Assume that $v > 0$. Then F_v' is strictly increasing, so that

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

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

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

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 (51). Hence we obtain that $L(v) \geq \lim_{x\pm\infty} |F_v'(x)|$. The proof of (51) is complete. □

Next, we prove the two following lemmas used in the proof of Proposition 2.1.

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 for any* $x \in \mathbb{R}$, *define*

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

Then for any $A > 0$

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

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

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

Therefore we have $\lim_{(\beta,\lambda)\to(\beta_0,\lambda_0)} \sup_{|x|\leq A} |b_{\beta,\lambda}(x)^2 - 1| = 0$. 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 is proved. The second one holds since $\sup_{|x|\leq A} \left|a_\alpha(x)\right| \leq A\,|\alpha - \alpha_0|/\sqrt{\beta_0}$. □

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*

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

Proof. Let $\mathcal{C}_K(\mathbb{R})$ be the set of continuous functions on $\mathbb{R}$ with compact support. 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 $b \geq 1/2$

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

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

C Proof of (36a)–(36b) under the assumptions (34a) and (35) with $N = 1$

Thoughout this section, the conditions (34a) and (35) with $N = 1$ are assumed to hold. Note that (35) with $N = 1$ is $\kappa_{1,a} = \mathbb{E}[L(\vartheta_1)^a]^{1/a} < 1$. We prove the properties (36a) and (36b) of Proposition 4.1 with explicit constants. Under the general assumption $\widehat{\kappa}_a < 1$ of Condition $(\widehat{\mathcal{C}}_a)$, the proof of (36a)-(36b) is similar (replace P with P^N with N such that $\mathbb{E}[L(\vartheta_N : \vartheta_1)^a] < 1$).

That the constant ξ in Proposition 4.1 is finite can be easily deduced from the drift inequality (47) 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:

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

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

$$\mathbb{E}\big[\Delta_n(\mu_1,\mu_2)\big] \leq \xi^{\frac{a-1}{a}}\, \kappa_{1,a}^n\, \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})]\, \big(\|\mu_1\|_a + \|\mu_2\|_a\big)^{a-1}. \tag{52}$$

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

$$\mathbb{E}\big[|f(X_n^{\mu_1}) - f(X_n^{\mu_2})|\big] \leq \xi^{\frac{a-1}{a}}\, m_a(f)\, \kappa_{1,a}^n\, \mathbb{E}[d(X_0^{\mu_1}, X_0^{\mu_2})]\, \big(\|\mu_1\|_a + \|\mu_2\|_a\big)^{a-1}. \tag{53}$$

Proof. Note that $X_n^\mu = F_{\vartheta_n:\vartheta_1} X_0^\mu$ from Definition 1.1 and the notations introduced in (33). If $a := 1$, then (52) 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

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

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

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

This proves (52). Property (53) follows from (52) and the definition of $m_a(f)$. □

Now recall that we consider the case $N = 1$ in (35). Let us prove the inequality (36a) in this case (that is (37)). Property (53), applied to $\mu_1 := \delta_x$ and $\mu_2 := \pi$ gives

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

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

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

Hence $\mathbb{E}[d(x, X_0^\pi)]\,(\|\delta_x\|_a + \|\pi\|_a)^{a-1} \leq p(x)^a \|\pi\|_1\,(1 + \|\pi\|_a)^{a-1}$. This proves the expected inequality.

Finally, to prove (36b), it remains to study $m_a(P^n f)$ for $f \in \mathcal{L}_a$. Inequality (53) applied to $\mu_1 := \delta_x$ and $\mu_2 := \delta_y$ for any $(x,y) \in \mathbb{X}^2$ gives:

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

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

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

Combining the last inequality with (37) gives (36b).

D Proof of Proposition 5.1

For every $\theta := (\xi,\gamma) \in \Theta := (0,+\infty) \times (\gamma_0,+\infty)$ we have

$$\mathbb{E}\left[\|F_\xi\big(0,\vartheta_1^{(\gamma)}\big)\|^a\right] = \int_{\mathbb{R}^d} \|\Phi_\xi\big(F(0,v)\big)\|^a \,\nu_\gamma(v)\,dv \leq \left(\frac{\kappa_a}{\widetilde{\kappa}_a}\right)^a \int_{\mathbb{R}^q} \|F(0,v)\|^a \,\nu(v)\,dv < \infty$$

so that Assumption (H$_2$) holds with $x_0 = 0$ and $M_a := \widetilde{M}_a \kappa_a/\widetilde{\kappa}_a$. Moreover note that $L_{\Phi_\xi} \leq 1$, so that

$$\forall v \in \mathbb{V}, \quad L_{F_\xi}(v) \leq L_F(v).$$

Hence we have for every $\theta = (\xi,\gamma) \in \Theta$

$$\mathbb{E}\left[L_{F_\xi}\big(\vartheta_1^{(\gamma)}\big)^a\right] = \int_{\mathbb{R}^d} L_{F_\xi}(v)^a \,\nu_\gamma(v)\,dv \leq c_\gamma \int_{\mathbb{R}^d} L_F(v)^a \,\nu(v)\,dv \leq c_\gamma \widetilde{\kappa}_a^a \leq {\kappa_a}^a.$$

Thus Assumption (H$_3$) holds. It remains to check Assumption (H$_4$) and to specify the error term Δ_θ. Let P (respectively P_θ) denote the transition kernel of the unperturbed IFS $\{X_n\}_{n\in\mathbb{N}}$ (respectively of the perturbed IFS $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$). Let $\varepsilon > 0$, and let $f \in \mathcal{B}_0$ be such that $|f|_0 \leq 1$. First note that we have for every $x \in \mathbb{R}^d$ satisfying $\|x\| > A_\varepsilon$

$$\frac{\big|(P_\theta f)(x) - (Pf)(x)\big|}{V_a(x)} \leq \frac{2}{V_a(x)} \leq \varepsilon \tag{54}$$

by definition of A_ε in Proposition 5.1. Next, for every $\theta := (\xi,\gamma) \in \Theta$ and for every $x \in \mathbb{R}^d$ define the following subset $E_{\theta,x}$ and $G_{\theta,x}$ of $\mathbb{R}^q$:

$$E_{\theta,x} := \big\{v \in \mathbb{R}^q : v \in B(0,\gamma),\ \|F(x,v)\| \leq \xi\big\},\ G_{\theta,x} := \big\{v \in \mathbb{R}^q : v \in B(0,\gamma),\ \|F(x,v)\| > \xi\big\}.$$

From the definition of the thresholding function Φ_ξ, we have for every $x \in \mathbb{R}^d$

$$\begin{aligned}(P_\theta f)(x) &= \int_{\mathbb{R}^q} f\big(F_\xi(x,v)\big)\,\nu_\gamma(v)\,dv \\ &= c_\gamma \int_{E_{\theta,x}} f\big(F(x,v)\big)\,\nu(v)\,dv + c_\gamma \int_{G_{\theta,x}} f(\eta_{x,v})\,\nu(v)\,dv\end{aligned}$$

with $\eta_{x,v} := \xi\|F(x,v)\|^{-1}\,F(x,v)$. Hence

$$\begin{aligned}\big|(P_\theta f)(x) - (Pf)(x)\big| &\leq |c_\gamma - 1| \int_{E_{\theta,x}} \nu(v)\,dv + (1+c_\gamma) \int_{\mathbb{R}^q\setminus E_{\theta,x}} \nu(v)\,dv \\ &\leq |c_\gamma - 1| + \left(1 + \left(\frac{\kappa_a}{\widetilde{\kappa}_a}\right)^a\right)\Big(\mathbb{P}\big(\|\vartheta_1\| > \gamma\big) + \mathbb{P}\big(\|F(x,\vartheta_1)\| > \xi\big)\Big)\end{aligned}$$

from the definition of $\mathbb{R}^q \setminus E_{\theta,x}$ and the condition $c_\gamma \leq (\kappa_a/\widetilde{\kappa}_a)^a$. Now let $x \in \mathbb{R}^d$ be such that $\|x\| \leq A_\varepsilon$. Then

$$\forall v \in \mathbb{R}^q, \quad \|F(x,v)\| \leq L_F(v)\,A_\varepsilon + \|F(0,v)\|$$

from $F(x,v) = (F(x,v) - F(0,v)) + F(0,v)$ and from the triangular inequality. Therefore

$$\Big[\; L_F(v)\, A_\varepsilon \leq \frac{\xi}{2} \quad \text{and} \quad \|F(0,v)\| \leq \frac{\xi}{2} \;\Big] \implies \|F(x,v)\| \leq \xi,$$

from which we deduce that

$$\begin{aligned}
\mathbb{P}\big(\|F(x,\vartheta_1)\| > \xi\big) \;&\leq\; \mathbb{P}\big(L_F(\vartheta_1) > \frac{\xi}{2A_\varepsilon}\big) + \mathbb{P}\big(\|F(0,\vartheta_1)\| > \frac{\xi}{2}\big) \\
&\leq\; \frac{(2A_\varepsilon)^a}{\xi^a}\mathbb{E}\big[L_F(\vartheta_1)^a\big] + \frac{2^a}{\xi^a}\mathbb{E}\big[\|F(0,\vartheta_1)\|^a\big] \\
&\leq\; \frac{(2A_\varepsilon\widetilde{\kappa}_a)^a + (2\widetilde{M}_a)^a}{\xi^a}
\end{aligned}$$

from Markov inequality. Consequently we obtain that for every $x \in \mathbb{R}^d$ such that $\|x\| \leq A_\varepsilon$

$$\begin{aligned}
&\frac{\big|(P_\theta f)(x) - (Pf)(x)\big|}{V_a(x)} \\
&\leq\; \big|(P_\theta f)(x) - (Pf)(x)\big| \\
&\leq\; |c_\gamma - 1| + \left(1 + \left(\frac{\kappa_a}{\widetilde{\kappa}_a}\right)^a\right)\left(\frac{\mathbb{E}\big[\|\vartheta_1\|^a\big]}{\gamma^a} + \frac{(2A_\varepsilon\widetilde{\kappa}_a)^a + (2\widetilde{M}_a)^a}{\xi^a}\right). \qquad (55)
\end{aligned}$$

The conclusion of Proposition 5.1 follows from (54)-(55).

References

[AANQ04] E. Altman, K. E. Avrachenkov, and R. Núñez-Queija. Perturbation analysis for denumerable Markov chains with application to queueing models. *Adv. in Appl. Probab.*, 36(3):839–853, 2004.

[AFEB16] P. Alquier, N. Friel, R. Everitt, and A. Boland. Noisy Monte Carlo: convergence of Markov chains with approximate transition kernels. *Stat. Comput.*, 26(1-2):29–47, 2016.

[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.

[Bal00] V. Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[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.

[Cli07a] D. B. H. Cline. Regular variation of order 1 nonlinear AR-ARCH models. *Stochastic Process. Appl.*, 117(7):840–861, 2007.

[Cli07b] D. B. H. Cline. Stability of nonlinear stochastic recursions with application to nonlinear AR-GARCH models. *Adv. in Appl. Probab.*, 39(2):462–491, 2007.

[CP02] D. B. H. Cline and H. H. Pu. A note on a simple Markov bilinear stochastic process. *Statist. Probab. Lett.*, 56(3):283–288, 2002.

[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.

[GL06] S. Gouëzel and C. Liverani. Banach spaces adapted to Anosov systems. *Ergodic Theory Dynam. Systems*, 26(1):189–217, 2006.

[GP14] L. Galtchouk and S. Pergamenshchikov. 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.*, 42:103–111, 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 aplication 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.

[JMMD] J. E. Johndrow, J. C. Mattingly, S. Mukherjee, and D. Dunson. Optimal approximating markov chains for bayesian inference. arXiv:1508.03387v3.

[Kar86] N. V. Kartashov. Inequalities in theorems of ergodicity and stability for markov chains with common phase space, parts I and II. *Theor. Probab. Appl.*, 30:pp. 247–259 and pp.505–515, 1986.

[Kel82] G. Keller. Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.*, 94(4):313–333, 1982.

[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.

[Liv04] C. Liverani. Invariant measure and their properties. a functional analytic point of view. In Pisa Pubblicazioni della Classe di Scienze, Scuola Normale Superiore, editor, *Dynamical Systems. Part II: Topological Geometrical and Ergodic Properties of Dynamics*, Centro di Ricerca Matematica Ennio De Giorgi, 2004.

[LP83] É. Le Page. Théorèmes de renouvellement pour les produits de matrices aléatoires. Équations aux différences aléatoires. In *Séminaires de probabilités Rennes 1983*, Publ. Sém. 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.

[MA10] Z. Mouhoubi and D. Aïssani. New perturbation bounds for denumerable Markov chains. *Linear Algebra Appl.*, 432(7):1627–1649, 2010.

[MALR16] F. J. Medina-Aguayo, A. Lee, and G. O. Roberts. Stability of noisy Metropolis-Hastings. *Stat. Comput.*, 26(6):1187–1211, 2016.

[MARS20] F. J. Medina-Aguayo, D. Rudolf, and N. Schweizer. Perturbation bounds for Monte Carlo within Metropolis via restricted approximations. *Stochastic Process. Appl.*, 130(4):2200–2227, 2020.

[Mit05] A. Y. Mitrophanov. Sensitivity and convergence of uniformly ergodic Markov chains. *J. Appl. Probab.*, 42(4):1003–1014, 2005.

[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.

[NR] J. Negrea and J. S. Rosenthal. Error bounds for approximations of geometrically ergodic Markov chains. arXiv:1702.07441v5.

[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.

[SS00] T. Shardlow and A. M. Stuart. A perturbation theory for ergodic Markov chains and application to numerical approximations. *SIAM J. Numer. Anal.*, 37(4):1120–1137, 2000.

[Tsa10] R. S. Tsay. *Analysis of financial time series.* Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2010.

[Wu04] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probab. Theory Related Fields*, 128(2):255–321, 2004.