J. Appl. Probab. **60,** 921–944 (2023)
*doi:*10.1017/jpr.2022.107

ROBUSTNESS OF ITERATED FUNCTION SYSTEMS OF LIPSCHITZ MAPS

LOÏC HERVÉ,* ** AND
JAMES LEDOUX ,* *** *Université de Rennes, INSA Rennes, CNRS, IRMAR-UMR 6625*

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 $\mathfrak{p}$, $F(\cdot, \cdot)$ 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 $\mathfrak{p}$, 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.

Keywords: Markov chain; geometric ergodicity; perturbation; non-linear stochastic recursion; autoregressive process

2020 Mathematics Subject Classification: Primary 60J05
Secondary 47B07; 62F99

1. Introduction

Let $(\mathbb{X}, d)$ be a Polish space equipped with its Borel σ-algebra $\mathcal{X}$. The random variables are assumed to be defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Throughout the paper we are concerned with iterated function systems (IFSs) 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 independent and identically distributed (i.i.d.) random variables with common distribution denoted by $\mathfrak{p}$. Let X_0 be an $\mathbb{X}$-valued random variable that 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 IFS is the sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ recursively defined, starting from X_0, as follows:

$$\text{for all } n \geq 1, \quad X_n := F(X_{n-1}, \vartheta_n). \tag{1.1}$$

Received 10 November 2021; revision received 6 September 2022.
* Postal address: INSA Rennes, 20 Avenue des Buttes de Cöesmes, CS 70 839, 35708 Rennes Cedex 7, France.
** Email address: loic.herve@insa-rennes.fr
*** Email address: james.ledoux@insa-rennes.fr

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 $V_a(\cdot)$, i.e.

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

Note that $(\mathcal{B}_0, |\cdot|_0)$ is 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 μ_0 and μ_1 on $\mathbb{X}$ is defined by $\|\mu_0 - \mu_1\|_{\mathrm{TV}} := \sup_{|f|_0 \leq 1} |\mu_0(f) - \mu_1(f)|$, where $\mu_i(f) := \int_{\mathbb{X}} f(x)\, \mathrm{d}\mu_i(x)$, $i = 0, 1$.

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:

$$\text{for all } x \in \mathbb{X} \text{ and } A \in \mathcal{X}, \quad P(x, A) = \mathbb{E}[\mathbf{1}_A(F(x, \vartheta_1))] = \int_{\mathbb{V}} \mathbf{1}_A(F(x, v))\, \mathrm{d}\mathfrak{p}(v). \tag{1.2}$$

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

$$\text{for all } n \geq 1 \text{ and } f \in \mathcal{B}_a, \quad |P^n f - \pi(f)\mathbf{1}_{\mathbb{X}}|_a \leq C_a \rho_a^n |f|_a. \tag{1.3}$$

The V_a-geometric ergodicity of IFSs has been extensively studied (see, e.g., [2, 12, 18, 36, 47] 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 1.1 below (see, e.g., [13]), for which we introduce the following notation. If $\psi : (\mathbb{X}, d) \to (\mathbb{X}, d)$ is a Lipschitz continuous function, we define

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

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

Condition 1.1. The function $F(\cdot, \cdot)$ and the sequence $\{\vartheta_n\}_{n \geq 1}$ satisfy

$$\mathbb{E}\left[d\big(x_0, F(x_0, \vartheta_1)\big)^a\right] < \infty, \tag{1.4}$$

$$\mathbb{E}\left[L_F(\vartheta_1)^a\right] < 1. \tag{1.5}$$

The condition $a \geq 1$ in Condition 1.1 is just a technical assumption for applying the Hölder inequality, for instance. In fact, Condition 1.1 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 1.1 for the so-called vector autoregressive model.

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 $d(x, y) := \|x - y\|$ as 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 IFS $X_0 \in \mathbb{R}^q$, for all $n \geq 1$, $X_n := AX_{n-1} + \vartheta_n$. 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 $d(0, F(0, v)) = \|v\|$. Consequently, Condition 1.1 holds for some $a \in [1, +\infty)$ provided that $\mathbb{E}[\|\vartheta_1\|^a] < \infty$ and $\|A\| < 1$. Moreover, if ϑ_1 has a probability density function (PDF) on $\mathbb{R}^q$, then P is V_a-geometrically ergodic. More precisely, inequality in (1.3) holds for any $\rho_a \in (\|A\|, 1)$; see Remark 4.2.

The aim of this work is to use the results of [14, 23, 43] to investigate the robustness first of the V_a-geometrical ergodicity property (1.3), second of the stationary distribution π, and third of the probability distribution of X_n. This study is made with respect to parametric variations of both the function F and the probability distribution of the noise random variable ϑ_n in (1.1). 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 random variables, both satisfying the assumptions of Definition 1.1. The common parametric probability distribution of $\{\vartheta_n^{(\gamma)}\}_{n\geq 1}$ is denoted by $\mathfrak{p}_\gamma$. 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

$$X_0^{(\theta)} \in \mathbb{X}, \text{ for all } n \geq 1, \quad X_n^{(\theta)} := F_\xi\left(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}\right). \tag{1.6}$$

The transition kernel of the Markov chain $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ is denoted by P_θ, and μ_θ is the probability distribution 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.

Assumption 1.1. *There exists $a \geq 1$ such that P_{θ_0} is V_a-geometrically ergodic with stationary distribution denoted by π_{θ_0}; i.e. P_{θ_0} satisfies* (1.3) *for some $\rho_a \in (0, 1)$ and $C_a > 0$.*

Assumption 1.2. $M_a := \sup_{\theta\in\Theta} \mathbb{E}\left[d\left(x_0, F_\xi\left(x_0, \vartheta_1^{(\gamma)}\right)\right)^a\right]^{1/a} < \infty.$

Assumption 1.3. $\kappa_a := \sup_{\theta\in\Theta} \mathbb{E}\left[L_{F_\xi}\left(\vartheta_1^{(\gamma)}\right)^a\right]^{1/a} < 1.$

Assumption 1.4. $\Delta_\theta := \|P_\theta - P_{\theta_0}\|_{0,a} \xrightarrow[\theta\to\theta_0]{} 0$, *where*

$$\|P_\theta - P_{\theta_0}\|_{0,a} := \sup_{f\in\mathcal{B}_0,\ |f|_0\leq 1} |P_\theta f - P_{\theta_0} f|_a.$$

Assumption 1.1 is the natural starting point for our perturbation issues. Note that Assumptions 1.2 and 1.3 are nothing but the uniform version with respect to θ of Condition 1.1. As a by-product, it follows from Assumptions 1.2 and 1.3 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 drift condition, uniform in $\theta \in \Theta$, holds true (see Appendix A):

$$\text{for all } \theta \in \Theta, \quad P_\theta V_a \leq \delta_a V_a + K_a, \text{ with } \delta_a := \kappa^a \text{ and } K_a := \frac{(1 + \kappa_a + M_a)^a (1 + M_a)^a}{(\kappa - \kappa_a)^a}. \tag{1.7}$$

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 $\theta \mapsto \pi_\theta$

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

continuous with respect to the total variation distance? 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 [14, 23, 43].

Theorem 1.1. *Under Assumptions 1.1–1.4, 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,*

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

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

$$\textit{for all } \theta \in \Theta, \quad \|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} \leq \frac{\exp(1) K_a D_a^{[\ln(\Delta_\theta^{-1})]^{-1}}}{(1-\delta_a)(1-\rho_a)} \Delta_\theta \ln(\Delta_\theta^{-1}), \tag{1.8}$$

provided that $\Delta_\theta \in (0, \exp(-1))$, where the constants ρ_a, C_a, δ_a, and K_a are given in Assumption 1.1 and (1.7), and $D_a = 2C_a(K_a + 1)$.

- P_3: *We have, for every $n \geq 1$ and for every $\theta \in \Theta$,*

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

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

In the general framework of V-geometrically ergodic Markov chains, Property P_1 and the first statement in P_2 are proved in [14, Theorem 1] by using the Keller–Liverani perturbation theorem [29]. The real-valued parameter ε in [14] may be replaced with the Θ-valued parameter θ. The inequality (1.8) in P_2 follows from [23, Proposition 2.1] or [43, (3.19)]. The formulation of [43, (3.19)] has been preferred to that in [23, Proposition 2.1] in connection with Property P_3. Property P_3 is proved in [43, Theorem 3.2] by using the Wasserstein distance associated with a suitable metric on $\mathbb{X}$ defined from the Lyapunov function V, as introduced in [19].

The goal of this work is to present various applications when both the function F and the probability distribution $\mathfrak{p}$ of the noise in Definition 1.1 are perturbed, and to show that the weak continuity assumption, Assumption 1.4, is well suited to such a study. This last claim is highlighted by the following first simple application, where only the probability distribution 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 $\mathfrak{p}_0$. Its transition kernel $P_{\mathfrak{p}_0}$ is given, for all $f \in \mathcal{B}_0$, by $(P_{\mathfrak{p}_0} f)(x) = \int f(F(x, y))\, d\mathfrak{p}_0(y)$. Let us consider the specific perturbation scheme $X_n^{(\theta)} := F\big(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}\big)$ for $X_0^{(\theta)} \in \mathbb{X}$ and all $n \geq 1$, where $\{\vartheta_n^{(\gamma)}\}_{n\geq 1}$ is a sequence of $\mathbb{V}$-valued i.i.d. random variables with a common parametric probability distribution denoted by $\mathfrak{p}_\gamma$. 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, for all $x \in \mathbb{X}$, $|(P_{\mathfrak{p}_\gamma}f)(x) - (P_{\mathfrak{p}_0}f)(x)| \leq \|\mathfrak{p}_\gamma - \mathfrak{p}_0\|_{\mathrm{TV}}$. It follows that

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

Hence, Assumption 1.4 is satisfied provided that $\lim \|\mathfrak{p}_\gamma - \mathfrak{p}_0\|_{\mathrm{TV}} = 0$.

In Section 2, a second application of Theorem 1.1, which again illustrates the interest of Assumption 1.4, 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 probability distribution of the noise ϑ_1 are perturbed. This Markov chain is called an autoregressive process of order one with autoregressive conditional heteroscedastic errors of order one (AR(1)-ARCH(1)). Such autoregressive models with conditional heteroscedastic errors were introduced to allow the conditional variance of time-series models to depend on past information. It turns out that such processes fit many types of econometrics and financial data very well where stochastic volatility must be taken into account (see, e.g., [46]). Note that the perturbation results of Section 2 can be extended to multivariate AR(p)-ARCH(q) processes with any order (p, q) [35] 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 a computer to get practical answers. This concerns simulation, approximation, numerical schemes, and so on. Thus, when a Markov model is implemented on a 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 [7, 42]. 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, Assumption 1.4, is still proved to be well adapted, as illustrated in Proposition 3.1, for the VAR model defined in Example 1.1. Note that the function F_ξ in (1.6) is fixed in Example 1.2 so that we do not have to divide by $V(x)$ to prove Assumption 1.4. Indeed, the inequality $\|P_{\mathfrak{p}_\gamma} - P_{\mathfrak{p}_0}\|_{0,0} \leq \|\mathfrak{p}_\gamma - \mathfrak{p}_0\|_{\mathrm{TV}}$ in Example 1.2 is directly obtained and it automatically gives Assumption 1.4. When F_ξ in (1.6) is not fixed as in Sections 2 and 3, the division by $V(x)$ in the definition of $\|\cdot\|_{0,a}$ must be done to investigate Assumption 1.4.

In Section 4 we propose a new approach to prove the V_a-geometrical ergodicity of IFSs of Lipschitz maps under Condition 1.1, together with a bound on the spectral gap of P (i.e. the infimum bound of the positive real numbers ρ_a satisfying (1.3)). In Section 5, further applications of Theorem 1.1 are presented. The goal of Section 5.1 is to show that the arguments developed in Sections 2 and 3 for specific IFSs of Lipschitz maps naturally extend to more general IFSs. In Section 5.2 we apply Theorem 1.1 to the case where the function F and the PDF $\mathfrak{p}$ of $\{\vartheta_n\}_{n\geq 1}$ in (1.1) are perturbed, by thresholding and by truncation respectively.

Perturbation theory for Markov chains is a natural issue which has been widely developed in recent decades. As mentioned in [45, p. 1126] (see also [14]), the strong continuity assumption introduced in [26, 27], 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 [3]) and for uniformly ergodic Markov chains (i.e. when (1.3) holds with $a = 0$); see [1, 25, 39, 40]. Similar questions arise for dynamical systems, and [28, p. 316] seems to be the first work which introduced a weaker continuity assumption using two norms as in Assumption 1.4 (instead of a single

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

one in the standard theory). Then, the Keller–Liverani perturbation theorem [4, 29, 31] has proved to be very powerful for studying the behaviour of the Sinai–Ruelle–Bowen measures of certain perturbed dynamical systems (see, e.g., [4, Theorem 2.10] and [16, Theorem 2.8]). In the context of V-geometrically ergodic Markov chains, Keller's approach is used in [45] and the Keller–Liverani theorem is applied in [14, 23]. The recent works of [34, 43] and references therein combine Keller's approach and the elegant idea of [19] using the Wasserstein distance associated with a suitable metric on $\mathbb{X}$ defined from the Lyapunov function V. Perturbation issues have also been investigated in the framework of roundoff errors [7, 42] (see Section 3) and in the special case of reversible transition kernels as in Markov chain Monte Carlo methods; see, e.g., [33, 41] and the references therein. The purpose of this paper is to show that the material developed in [14, 23, 43] is very well suited to the perturbation of general IFSs. In the IFS context, Assumption 1.4 has so far only been investigated in [14, 43] 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 1.4 allows us to deal with perturbation schemes of the general IFS (1.1) with respect to both function F and the probability distribution 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 (see, e.g., [11, 17] in the Markov context).

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

According to Definition 1.2, we consider the perturbed AR(1)-ARCH(1) real-valued $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ defined, for $X_0^{(\theta)}$ a given real-valued random variable, by

$$\text{for all } n \geq 1, \quad X_n^{(\theta)} := F_\xi\big(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}\big), \tag{2.1}$$

where $F_\xi(x, v) = \alpha x + v(\beta + \lambda x^2)^{1/2}$ with constants $\alpha \in \mathbb{R}$, $\beta > 0$, $\lambda > 0$, and $\{\vartheta_n^{(\gamma)}\}_{n\in\mathbb{N}}$ has a common PDF $\mathfrak{p}_\gamma$ 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 $\big\{X_n^{(\theta)}\big\}_{n\in\mathbb{N}}$ is given by $P_\theta(x, A) := \int_{\mathbb{R}} \mathbf{1}_A(y) p_\theta(x, y)\, \mathrm{d}y$ ($A \in \mathcal{X}$), with

$$p_\theta(x, y) := \big(\beta + \lambda x^2\big)^{-1/2} \mathfrak{p}_\gamma\left(\frac{y - \alpha x}{(\beta + \lambda x^2)^{1/2}}\right). \tag{2.2}$$

Next, we report the following observations with respect to basic quantities required in Assumptions 1.2 and 1.3. First, it can be checked (see Lemma B.1) that

$$L_{F_\xi}(\vartheta_1) = \max\big(|\alpha - \sqrt{\lambda}\vartheta_1|,\ |\alpha + \sqrt{\lambda}\vartheta_1|\big). \tag{2.3}$$

Hence, the real number κ_a in Assumption 1.3 is

$$\kappa_a = \sup_{\theta\in\Theta} \left(\int_{\mathbb{R}} \max\big(|\alpha - \sqrt{\lambda}v|,\ |\alpha + \sqrt{\lambda}v|\big)^a \mathfrak{p}_\gamma(v)\, \mathrm{d}v\right)^{1/a}. \tag{2.4}$$

Second, the real number M_a in Assumption 1.2 is given by

$$M_a := \sup_{\theta\in\Theta} \sqrt{\beta}\, \mathbb{E}\big[\big|\vartheta_1^{(\gamma)}\big|^a\big]^{1/a} = \sup_{\theta\in\Theta} \sqrt{\beta}\left(\int_{\mathbb{R}} |v|^a \mathfrak{p}_\gamma(v)\, \mathrm{d}v\right)^{1/a}. \tag{2.5}$$

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

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

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.

Assumption 2.1. *There exists $a \geq 1$ such that:*

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

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

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

(b) $M_a < \infty$, where M_a is given in (2.5).

(c) $\kappa_a < 1$, where κ_a is given in (2.4).

Assumption 2.2. $\lim_{\gamma\to\gamma_0} \|\mathfrak{p}_\gamma - \mathfrak{p}_{\gamma_0}\|_{\mathbb{L}^1(\mathbb{R})} = 0$.

Proposition 2.1. *Under Assumptions 2.1 and 2.2 for the AR(1)-ARCH(1) process given in (2.1), the assertions P1–P3 of Theorem 1.1 hold.*

Proof. Let $\theta_0 = (\alpha_0, \beta_0, \lambda_0, \gamma_0) \in \mathring{\Theta}$, and let $a \geq 1$ in Assumption 2.1. As already discussed, the conditions in Assumptions 2.1(b) and (c) imply that Assumptions 1.2 and 1.3 of Theorem 1.1 hold. Moreover, we can use (A.1) to state that there exist $\delta_a < 1$, $K_a > 0$. and $r_a > 0$ such that $P_{\theta_0} V_a \leq \delta_a V_a + K_a \mathbf{1}_{[-r_a,r_a]}$.

Next, Assumption 2.1(a) ensures that, for all $x \in [-r_a, r_a]$ and $A \in \mathcal{X}$, $P_{\theta_0}(x, A) \geq \varphi_{r_a,\theta_0}(A)$ 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 Assumption 2.1(a). Then, Assumption 1.1 holds true; see [5, Theorem 1.1] and [36].

The following lemma asserts that Assumption 1.4 holds true under Assumption 2.2, so the proof is complete. □

Lemma 2.1. *If* $\lim_{\gamma\to\gamma_0} \|\mathfrak{p}_\gamma - \mathfrak{p}_{\gamma_0}\|_{\mathbb{L}^1(\mathbb{R})} = 0$ *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, for all $x \in \mathbb{X}$,

$$\frac{|(P_\theta f)(x) - (P_{\theta_0} f)(x)|}{V_a(x)} = \frac{\left|\int_{\mathbb{R}} (p_\theta(x, y) - p_{\theta_0}(x, y)) f(y)\, dy\right|}{V_a(x)} \leq \frac{\int_{\mathbb{R}} \left|p_\theta(x, y) - p_{\theta_0}(x, y)\right| dy}{V_a(x)}.$$

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

$$|x| > B \Rightarrow \text{for all } \theta \in \Theta, \quad \frac{|(P_\theta f)(x) - (P_{\theta_0}) f(x)|}{V_a(x)} < \frac{\varepsilon}{2}. \tag{2.6}$$

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

$$\text{for all } A > 0, \quad \lim_{\theta\to\theta_0} \sup_{|x|\leq A} \frac{\int_{\mathbb{R}} |p_\theta(x, y) - p_{\theta_0}(x, y)|\, dy}{V_a(x)} = 0. \tag{2.7}$$

Indeed, (2.6) and (2.7) with $A = B$ ensure that $\|P_\theta - P_{\theta_0}\|_{0,a} < \varepsilon$ when θ is sufficiently close to θ_0.

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

Let us prove (2.7). It follows from (2.2) that

$$\int_{\mathbb{R}} |p_\theta(x, y) - p_{\theta_0}(x, y))| \, \mathrm{d}y$$

$$\leq \int_{\mathbb{R}} (\beta + \lambda x^2)^{-1/2} \left| \mathfrak{p}_\gamma\left(\frac{y - \alpha x}{(\beta + \lambda x^2)^{1/2}}\right) - \mathfrak{p}_{\gamma_0}\left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}}\right)\right| \mathrm{d}y \tag{2.8}$$

$$+ \int_{\mathbb{R}} \mathfrak{p}_{\gamma_0}\left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}}\right) \left|(\beta + \lambda x^2)^{-1/2} - (\beta_0 + \lambda_0 x^2)^{-1/2}\right| \mathrm{d}y. \tag{2.9}$$

First, using the change of variables $z = (y - \alpha x)/(\beta + \lambda x^2)^{1/2}$ in the integral (2.8) and the triangle inequality, we obtain

$$(2.8) = v \int_{\mathbb{R}} \left| \mathfrak{p}_\gamma(z) - \mathfrak{p}_{\gamma_0}\left(\left(\frac{\beta + \lambda x^2}{\beta_0 + \lambda_0 x^2}\right)^{1/2} z + x \frac{\alpha - \alpha_0}{(\beta_0 + \lambda_0 x^2)^{1/2}}\right)\right| \mathrm{d}z$$

$$\leq \int_{\mathbb{R}} |\mathfrak{p}_\gamma(z) - \mathfrak{p}_{\gamma_0}(z)| \, \mathrm{d}z + \int_{\mathbb{R}} |\mathfrak{p}_{\gamma_0}(z) - \mathfrak{p}_{\gamma_0}(b_{\beta,\lambda}(x) z + a_\alpha(x))| \, \mathrm{d}z, \tag{2.10}$$

where

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

The first integral in (2.10) does not depend on x and is equal to $\|\mathfrak{p}_\gamma - \mathfrak{p}_{\gamma_0}\|_{\mathbb{L}^1(\mathbb{R})}$, 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 Assumption 2.2, Lemma B.3 allows us to conclude that the second integral in (2.10) is such that

$$\lim_{(\alpha,\beta,\lambda)\to(\alpha_0,\beta_0,\lambda_0)} \sup_{|x|\leq A} \int_{\mathbb{R}} |\mathfrak{p}_{\gamma_0}(z) - \mathfrak{p}_{\gamma_0}(b_{\beta,\lambda}(x) z + a_\alpha(x))| \, \mathrm{d}z = 0.$$

Second, let us consider the integral (2.9). 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 $|x| \leq A$,

$$(2.9) = \left|(\beta + \lambda x^2)^{-1/2} - (\beta_0 + \lambda_0 x^2)^{-1/2}\right| \times \int_{\mathbb{R}} \mathfrak{p}_{\gamma_0}\left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}}\right) \mathrm{d}y$$

$$= (\beta_0 + \lambda_0 x^2)^{-1/2} \left|\frac{1}{b_{\beta,\lambda}(x)} - 1\right| \times \int_{\mathbb{R}} \mathfrak{p}_{\gamma_0}\left(\frac{y - \alpha_0 x}{(\beta_0 + \lambda_0 x^2)^{1/2}}\right) \mathrm{d}y$$

$$= \left|\frac{1 - b_{\beta,\lambda}(x)}{b_{\beta,\lambda}(x)}\right| \int_{\mathbb{R}} \mathfrak{p}_{\gamma_0}(z) \, \mathrm{d}y \quad \text{(change of variables } z = (y - \alpha_0 x)/(\beta_0 + \lambda_0 x^2)^{1/2})$$

$$\leq \frac{|1 - b_{\beta,\lambda}(x)|}{b_\beta(A)},$$

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

Remark 2.1. If the PDF $\mathfrak{p}_{\gamma_0}$ for the unperturbed AR(1)-ARCH(1) process is continuous on $\mathbb{R}$, then Assumption 2.1(a) (stated to prove Assumption 1.1) can be omitted in Proposition 2.1. Actually, under the condition $\int_{\mathbb{R}} L_{F_{\xi_0}}(v)^a \mathfrak{p}_{\gamma_0}(v)\,\mathrm{d}v < 1$, which is contained in Assumption 2.1(c), Assumption 1.1 holds with any real number ρ_a (and the associated constant C_a) such that

$$\left(\int_{\mathbb{R}} L_{F_{\xi_0}}(v)^a \mathfrak{p}_{\gamma_0}(v)\,\mathrm{d}v\right)^{1/a} < \rho_a < 1.$$

Indeed, the kernel $p_{\theta_0}(x, y)$ given by (2.2) is continuous, so Remark 4.3 and Proposition 4.2 ensure that, under the conditions in Assumption 2.1(b) and (c), P_{θ_0} satisfies (1.3) for any ρ_a satisfying the above condition. In other words, if the PDF $\mathfrak{p}_{\gamma_0}$ is continuous on $\mathbb{R}$, then only the conditions in Assumption 2.1(b) and (c) with $\Theta = \{\theta_0\}$ are useful in obtaining Assumption 1.1.

Remark 2.2. It is well known from Scheffé's lemma [44] that the almost-everywhere pointwise convergence of the PDF $\mathfrak{p}_\gamma$ to the PDF $\mathfrak{p}_{\gamma_0}$ when $\gamma \to \gamma_0$ provides the $\mathbb{L}^1(\mathbb{R})$-convergence required in Assumption 2.2.

3. Robustness of IFS under roundoff error

From [7, 42], the effect of roundoff errors using a Markov chain with transition kernel P means considering a Markov chain with a (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 $h_\theta \to \mathrm{id}$ 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 as

$$\text{for } X_0^{(\theta)} \in \mathbb{X} \text{ and all } n \geq 1,\quad X_n^{(\theta)} = F_\theta\big(X_{n-1}^{(\theta)}, \vartheta_n\big),$$

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

$$\text{for all } f \in \mathcal{B}_0 \text{ and } x \in \mathbb{X},\quad (P_\theta f)(x) = P(f \circ h_\theta)(x) = \int_{\mathbb{V}} f((h_\theta \circ F)(x, v))\,\mathrm{d}\mathfrak{p}(v). \tag{3.1}$$

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 [42] to study this question is to assume that $h_\theta \to \mathrm{id}$ uniformly on $\mathbb{R}^q$ when $\theta \to \theta_0$ (i.e., for all $x \in \mathbb{R}^q$, $\|h(x) - x\| \leq \varepsilon(\theta)$ with $\lim_{\theta\to\theta_0} \varepsilon(\theta) = 0$). However, as mentioned in [7], this assumption is too restrictive in practice since the roundoff errors for some $x \in \mathbb{R}^q$ are obviously proportional to x. In [7], the authors introduced the weaker assumption $\|h(x) - x\| \leq \varepsilon(\theta)\|x\|$ 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 again find 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 probability distribution 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 [7].

Let $\{X_n\}_{n\in\mathbb{N}}$ be an $\mathbb{R}^q$-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

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

we consider a family $(h_\theta)_{\theta\in\Theta}$ of functions on $\mathbb{X} := \mathbb{R}^q$ such that $h_0 = \mathrm{id}$. 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 (3.1)),

$$\text{for all } f \in \mathcal{B}_0 \text{ and } x \in \mathbb{X}, \quad (P_\theta f)(x) = \int_{\mathbb{R}^q} f(h_\theta(Ax + v))\, \mathrm{d}\mathfrak{p}(v). \tag{3.2}$$

If $g : \mathbb{R}^q \to \mathbb{R}^q$ is differentiable and $z \in \mathbb{R}^q$, 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 $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$.

Assumption 3.1. $\|A\| < 1$ *and there exists* $a \geq 1$ *such that* $\mathbb{E}[\|\vartheta_1\|^a] < \infty$.

Assumption 3.2. $\sup_{\theta\in\Theta} \int_{\mathbb{R}^q} \|h_\theta(v)\|^a \mathfrak{p}(v)\, \mathrm{d}v < \infty$.

Assumption 3.3. *For any* $\theta \in \Theta$, h_θ *is differentiable on* $\mathbb{R}^q$ *and* $\sup_{\theta\in\Theta} \|\nabla h_\theta\|_\infty < \|A\|^{-1}$.

Assumption 3.4.

(a) The probability distribution of ϑ_1 *admits a bounded continuous PDF* $\mathfrak{p}$ *satisfying the following monotonicity-type condition: there exists* $M > 0$ *such that, for every* $z_1, z_2 \in \mathbb{R}^q$, $M \leq \|z_1\| \leq \|z_2\| \Rightarrow \mathfrak{p}(z_2) \leq \mathfrak{p}(z_1)$.

(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) There exists* $c \in (0, 1)$ *such that, for all* $\theta \in \Theta$ *and* $z \in \mathbb{R}^q$, $\|g_\theta(z) - z\| \leq c\|z\|$.
- *(ii) For all* $z \in \mathbb{R}^q$, $\lim_{\theta\to 0} g_\theta(z) = z$.
- *(iii)* $\sup_{\theta\in\Theta} \|\nabla g_\theta\|_\infty < \infty$, *and* $\lim_{\theta\to 0} \nabla g_\theta = \mathrm{id}$ *uniformly on each ball of* $\mathbb{R}^q$ *centred at* 0, *i.e. for all* $A > 0$ *and* $\eta > 0$ *there exists* $\alpha > 0$ *such that, for all* $\theta \in \Theta$ *with* $\|\theta\| < \alpha$, $\sup_{\|z\|\leq A} \|\nabla g_\theta(z) - \mathrm{id}\| < \eta$.

Proposition 3.1. *Under Assumptions 3.1–3.4 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. The conditions in Assumption 3.4(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 [7], the uniform convergence $\lim_{\theta\to 0} g_\theta = \mathrm{id}$ (or $\lim_{\theta\to 0} h_\theta = \mathrm{id}$) 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 $\lim_{\theta\to 0} g_\theta = \mathrm{id}$ nor the convergence $\lim_{\theta\to 0} h_\theta = \mathrm{id}$ are uniform on $\mathbb{R}$.

Proof. Recall that $\theta_0 = 0$ here. We know that Assumption 1.1 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 (3.2), P_θ has the form

$$\text{for all } f \in \mathcal{B}_0 \text{ and } x \in \mathbb{R}^q, \ (P_\theta f)(x) = \int_{\mathbb{R}^q} f(z)\mathfrak{p}(g_\theta(z) - Ax)\Gamma_\theta(z)\, \mathrm{d}z$$

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, Assumption 3.2 is just Assumption 1.2 (here $x_0 = 0$), while Assumption 1.3 is implied by Assumption 3.3 from Taylor's inequality applied to h_θ.

Next, we prove Assumption 1.4. 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

$$\text{for all } x \in \mathbb{R}^q \setminus B(0, K), \quad \frac{|(P_\theta f)(x) - (P_0 f)(x)|}{V(x)} \le \frac{2}{V(x)} < \varepsilon. \tag{3.3}$$

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

$$|(P_\theta f)(x) - (P_0 f)(x)| \le \int_{\mathbb{R}^q} |\mathfrak{p}(g_\theta(z) - Ax)\Gamma_\theta(z) - \mathfrak{p}(z - Ax)| \, \mathrm{d}z \tag{3.4}$$

since $g_0 = \mathrm{id}$. Set $d := 2/(1 - c)$, where c is given in Assumption 3.4(b)(i). Note that $\|Ax\| \le K$ and that Assumption 3.4(b)(i) provides, for all $z \in \mathbb{R}^q$, $\|g_\theta(z)\| \ge (1 - c)\|z\|$. 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 Assumption 3.4(a) that we have, for every $\theta \in \Theta$,

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

Since the function $z \mapsto \mathfrak{p}(d^{-1}z)$ is Lebesgue integrable on $\mathbb{R}^q$, we can choose $C \equiv C(\varepsilon) > 0$ such that $\int_{\|z\|\ge C} \mathfrak{p}(d^{-1}z) \, \mathrm{d}z \le \varepsilon/2(\gamma + 1)$, where $\gamma := \sup_{\theta\in\Theta} \sup_{z\in\mathbb{R}^q} \Gamma_\theta(z)$. Note that $\gamma < \infty$ from the first condition of Assumption 3.4(b)(iii) and from the continuity of the function $\det(\cdot)$. Set $D = \max(B, C)$. We deduce from the triangle inequality that, for every $\theta \in \Theta$,

$$\int_{\|z\|\ge D} |\mathfrak{p}(g_\theta(z) - Ax)\Gamma_\theta(z) - \mathfrak{p}(z - Ax)| \, \mathrm{d}z \le (\gamma + 1) \int_{\|z\|\ge C} \mathfrak{p}(d^{-1}z) \, \mathrm{d}z \le \frac{\varepsilon}{2}. \tag{3.5}$$

Now we investigate the integrand in (3.4) for $z \in B(0, D)$ (recall that $x \in B(0, K)$). First, setting $m := \sup_{u\in\mathbb{R}^q} \mathfrak{p}(u)$, we have, for every $z \in B(0, D)$ and $x \in B(0, K)$,

$$|\mathfrak{p}(g_\theta(z) - Ax)\Gamma_\theta(z) - \mathfrak{p}(z - Ax)| \le \gamma\,|\mathfrak{p}(g_\theta(z) - Ax) - \mathfrak{p}(z - Ax)| + m|\Gamma_\theta(z) - 1|. \tag{3.6}$$

We have, for all $z \in B(0, D)$, $\|g_\theta(z)\| \le (1 + c)D$ (using Assumption 3.4(b)(i)). From the standard statement for uniform convergence of differentiable functions, we deduce from the conditions in Assumption 3.4(b)(ii) and (b)(iii) that $\lim_{\theta\to 0} g_\theta = \mathrm{id}$ 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 $\mathfrak{p}$ on $B(0, (1 + c)D + K)$, there exists an open neighbourhood $\mathcal{V}_0$ of $\theta = 0$ in $\mathbb{R}^p$ such that

$$\text{for all } \theta \in \mathcal{V}_0,\ z \in B(0, D),\ \text{and } x \in B(0, K), \quad |\mathfrak{p}(g_\theta(z) - Ax) - \mathfrak{p}(z - Ax)| < \frac{\varepsilon}{4\gamma\ell_D}.$$

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

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

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

from Assumption 3.4(b)(iii) and from the uniform continuity of the function det(·) on every compact subset of the set $\mathcal{M}_q(\mathbb{R})$ of real $q \times q$ matrices. It then follows from (3.6) that

$$\text{for all } \theta \in \mathcal{V}_0',\ z \in B(0, D),\ \text{and } x \in B(0, K), \quad |\mathfrak{p}(g_\theta(z) - Ax)\Gamma_\theta(z) - \mathfrak{p}(z - Ax)| \le \frac{\varepsilon}{2\ell_D}.$$

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

$$\text{for all } \theta \in \mathcal{V}_0' \text{ and } x \in B(0, K), \quad \int_{\|z\| \le D} |\mathfrak{p}(g_\theta(z) - Ax)\Gamma_\theta(z) - \mathfrak{p}(z - Ax)|\, dz \le \frac{\varepsilon}{2}. \tag{3.7}$$

We deduce from (3.4), (3.5), and (3.7) that

$$\text{for all } \theta \in \mathcal{V}_0' \text{ and } x \in B(0, K), \quad \frac{|(P_\theta f)(x) - (P_0 f)(x)|}{V(x)} \le |(P_\theta f)(x) - (P_0 f)(x)| \le \varepsilon.$$

This inequality, combined with (3.3), gives Assumption 1.4. □

4. V_a-geometric ergodicity of IFS

For $a \ge 1$, 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 $\mathcal{L}_a$:

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

Such Lipschitz-weighted spaces were introduced in [30] to obtain the quasi-compactness of Lipschitz kernels (see also [6, 13, 20, 21, 38]). Note that, for $f \in \mathcal{L}_a$, we have, for all $x \in \mathbb{X}$, $|f(x)| \le |f(x_0)| + 2^{a-1}\, m_a(f)\, V_a(x)$, 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, for all $f \in \mathcal{L}_a$, $\|f\|_a := m_a(f) + |f|_a$.

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 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 $(v_1, \ldots, v_n) \in \mathbb{V}^n$ ($n \in \mathbb{N}^*$), define

$$F_{v_n:v_1} := F_{v_n} \circ \cdots \circ F_{v_1}, \qquad L(v_n : v_1) := L(F_{v_n:v_1}). \tag{4.1}$$

By hypothesis we have $L(v) < \infty$, and thus $L(v_n : v_1) < \infty$. Note that, for each $a \ge 1$, the limit $\hat{\kappa}_a := \lim_{n \to +\infty} \mathbb{E}[L(\vartheta_n : \vartheta_1)^a]^{1/(na)}$ 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 a standard contraction/moment condition, Condition 4.1 (counterpart of Condition 1.1 in Section 1), for P given in (1.2) to have a geometric rate of convergence on $\mathcal{L}_a$ (see Proposition 4.1). Then the passage to V_a-geometric ergodicity is addressed in Proposition 4.2.

Condition 4.1. For some $a \in [1, +\infty)$,

$$\mathbb{E}[d(x_0, F(x_0, \vartheta_1)^a)] < \infty, \tag{4.2}$$

$$\widehat{\kappa}_a < 1. \tag{4.3}$$

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

Note that (4.3) is equivalent to

$$\text{there exists } N \in \mathbb{N}^* \text{ such that } \mathbb{E}[L(\vartheta_N : \vartheta_1)^a] < 1, \tag{4.4}$$

and Condition 1.1 in Section 1 corresponds to (4.2) and to (4.4) with $N = 1$.

The properties in the next proposition can be derived from the results of [13, Chapter 6]; see also [6] for the existence and uniqueness of the invariant distribution. For convenience, in Appendix C the properties (4.5) and (4.6) are proved with explicit constants under the assumptions (4.2) and (4.4) with $N = 1$ (i.e. $\mathbb{E}[L(\vartheta_1)^a] < 1$).

Proposition 4.1. ([13, Chapter 6].) *Under Condition 4.1, 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 exist positive constants $c \equiv c_\kappa$ and $c' \equiv c'_\kappa$ such that:*

$$\text{for all } f \in \mathcal{L}_a \text{ and } n \geq 1, \quad |P^n f - \pi(f)\mathbf{1}_{\mathbb{X}}|_a \leq c\kappa^n m_a(f); \tag{4.5}$$

$$\text{for all } f \in \mathcal{L}_a \text{ and } n \geq 1, \quad \|P^n f - \pi(f)\mathbf{1}_{\mathbb{X}}\|_a \leq c'\kappa^n \|f\|_a. \tag{4.6}$$

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

$$\text{for all } f \in \mathcal{L}_a \text{ and } n \geq 1, \quad |P^n f - \pi(f)\mathbf{1}_{\mathbb{X}}|_a \leq c_1 {\kappa_{1,a}}^n m_a(f), \tag{4.7}$$

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

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

Under Condition 4.1, Property (4.5) 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 (4.5) (or (4.7)) does not provide the V_a-geometric ergodicity (1.3) since (4.5) (or (4.7)) is only established for $f \in \mathcal{L}_a$. Under Condition 4.1, it was proved in [2, Proposition 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 4.1, the Markov chain $\{X_n\}_{n\in\mathbb{N}}$ is shown to be V_a-geometrically ergodic in [47, Proposition 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 (4.8) is the same as in [47, Proposition 7.2], but the Feller-type assumptions of [47] are replaced with the following: $P^\ell : \mathcal{B}_0 \to \mathcal{B}_a$ for some $\ell \geq 1$ is compact (see Remark 4.2 for comparisons).

Proposition 4.2. *Let us assume that Condition 4.1 holds, 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 (1.3) holds) satisfies the following bound:*

$$\rho_{V_a}(P) \leq \widehat{\kappa}_a. \tag{4.8}$$

Proof. To avoid confusion, we simply denote by P the action of $P(x, \mathrm{d}y)$ on $\mathcal{B}_a$, and we denote by $P_{|\mathcal{L}_a}$ the restriction of P to $\mathcal{L}_a$. 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 (4.5). Then, Property (4.5)

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

applied to $f := V_a$ gives $P^N V_a \leq \delta^N V_a + \pi(V_a)$. We deduce from [23, Proposition 5.4 and Remark 5.5] that P is a power-bounded quasi-compact operator on $\mathcal{B}_a$, and that its essential spectral radius $r_{\text{ess}}(P)$ satisfies $r_{\text{ess}}(P) \leq \widehat{\kappa}_a$ since δ is arbitrarily close to $\widehat{\kappa}_a$ (see, e.g., [20] 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 $r_{\text{ess}}(P^*) \leq \widehat{\kappa}_a$.

Next, let us establish that P is V_a-geometrically ergodic from [24, Proposition 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 $r_{\text{ess}}(P^*) \leq \widehat{\kappa}_a < |\lambda|$. 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, for all $f \in \mathcal{L}_a$, $|\langle f', f\rangle| \leq \|f'\|_{\mathcal{B}_a'} |f|_a \leq \|f'\|_{\mathcal{B}_a'} \|f\|_a$. 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 (4.6), $P_{|\mathcal{L}_a}$ is quasi-compact on $\mathcal{L}_a$ and $r_{\text{ess}}(P_{|\mathcal{L}_a}) \leq \widehat{\kappa}_a < r_0 \leq |\lambda|$. Finally, we have $\lambda = 1$. Indeed, if $\lambda \neq 1$, then any $f \in \mathcal{L}_a$ satisfying $Pf = \lambda f$ is such that $\pi(f) = 0$, and thus $f = 0$ from (4.6) (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 eigenprojection on $\ker(P - I) = \ker(P - I)^2$. The last equality holds since P is power bounded on $\mathcal{B}_a$. Set $m := \dim \ker(P - I)$. From [47, Proposition 4.6] (see also [22, Theorem 1]), there exist m linearly independent non-negative functions $f_1, \ldots, f_m \in \ker(P - I)$ and probability measures $\mu_1, \ldots, \mu_m \in \ker(P^* - I)$ satisfying $\mu_k(V_a) < \infty$ such that, for all $f \in \mathcal{B}_a$, $\Pi f = \sum_{k=1}^m \mu_k(f) f_k$. That 1 is a simple eigenvalue of P on $\mathcal{B}_a$ then follows from the first assertion of Proposition 4.1.

From [24, Proposition 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 (4.8) means that, for any real number $\rho \in (\widehat{\kappa}_a, 1)$, there exists a constant $C \equiv C_\rho$ such that, for all $n \geq 1$ and $f \in \mathcal{B}_a$, $|P^n f - \pi(f)\mathbf{1}_{\mathbb{X}}|_a \leq C\rho^n |f|_a$. Unfortunately, neither the proof of Proposition 4.1 nor that of [47, Proposition 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 (see, e.g., [5, 23, 24, 32, 37] and the references therein). It is worth mentioning that explicit bounds on ρ and C are also provided in [15] 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 hypothesis that there exist a positive measure η on $(\mathbb{X}, \mathcal{X})$ and a measurable function $K : \mathbb{X}^2 \to [0, +\infty)$ such that

$$\text{for all } x \in \mathbb{X}, \qquad P(x, \mathrm{d}y) = K(x, y)\, \mathrm{d}\eta(y). \tag{4.9}$$

If P^ℓ is strongly Feller for some $\ell \geq 1$, then $P^{2\ell}$ is compact from $\mathcal{B}_0$ to $\mathcal{B}_a$ (see, e.g., [18, Lemma 3]). Hence, if P admits a kernel as in (4.9), then assuming that P^N is strongly Feller for some N in [47, Proposition 7.2] is more restrictive than the compactness hypothesis of Proposition 4.2. A detailed comparison with the approach in [47, Proposition 7.2] is provided in [18] for general Markov kernels. Finally, note that the transition kernel P of a VAR process (see Example 1.1) is always strongly Feller. Indeed, let $f \in \mathcal{B}_0$ be such that $\|f\|_0 \leq 1$. Then we have

$$\text{for all } (x, x') \in \mathbb{R}^q \times \mathbb{R}^q, \qquad |(Pf)(x') - (Pf)(x)| \leq \int_{\mathbb{R}^q} |\mathfrak{p}(y - A(x' - x)) - \mathfrak{p}(y)|\, \mathrm{d}y.$$

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

Since $t \mapsto \mathfrak{p}(\cdot - t)$ 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 [47, Section 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 the assumption in (4.9) 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 4.1. Indeed, we have, for all $(x, x') \in \mathbb{X}^2$ and for any $f \in \mathcal{B}_0$,

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

Since $K(\cdot, \cdot) \ge 0$, $\int K(\cdot, y)\, d\eta(y) = 1$, 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 $\{Pf, \ |f|_0 \le 1\}$ 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 $P^N V_a \le \delta^N V_a + \pi(V_a)$ was written with any $\delta \in (\widehat{\kappa}_a, 1)$ by using Property (4.5) of Proposition 4.1 in order to deduce the bound $r_{\mathrm{ess}}(P) \le \widehat{\kappa}_a$ 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 (4.6), from which we deduce the bound $r_{\mathrm{ess}}(P_{|\mathcal{L}_a}) \le \widehat{\kappa}_a$. Actually, for any $\delta \in (\widehat{\kappa}_a^a, 1)$, the drift inequality $P^N V_a \le \delta^N V_a + K$ with some $N \ge 1$ and $K > 0$ can be derived from Condition 4.1 by adapting the proof in Appendix A (here with $P_{\theta_0} = P$ and $\Theta = \{\theta_0\}$). Then, the more accurate bound $r_{\mathrm{ess}}(P) \le \widehat{\kappa}_a^a$ can be derived from [23, Proposition 5.4 and Remark 5.5] under the compactness assumption of Proposition 4.2. See also [47, Proposition 7.2], which provides the same bound under Feller-type assumptions.

5. Further applications

Theorem 1.1 was applied in Section 2 to real-valued AR(1) models with ARCH(1) errors (see Proposition 2.1), and in Section 3 to the roundoff errors of a VAR model (see Proposition 3.1). Although these applications have been presented for specific IFSs, it is worth noticing that they give a general road map for investigating the issues in P_1–P_3 of Section 1 for other instances of $\mathbb{R}^q$-valued IFSs, provided that the probability distribution of the noise $\mathfrak{p}_\gamma$ in Definition 1.2 admits a PDF 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 valid for every $x \in \mathbb{R}^q$, where $F_\xi(\cdot, \cdot)$ is the perturbed function involved in Definition 1.2. In Section 5.1 we propose two examples to support this claim. Finally, in Section 5.2 we discuss the robustness of IFSs of Lipschitz maps under perturbation caused 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

$$\text{for all } n \ge 1, \quad X_n = \psi(X_{n-1}) + B(X_{n-1})\vartheta_n, \tag{5.1}$$

where $\psi : \mathbb{R}^q \to \mathbb{R}^q$, $B : \mathbb{R}^q \to \mathrm{GL}_q(\mathbb{R})$, and the random variables $\{\vartheta_n\}_{n\ge 1}$ have common PDF $\mathfrak{p}$. 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 (5.1) encompasses a very large class of

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

non-linear time series models (see, e.g., [36, Chapter 2], [46, Chapter 4]), and [8, 9, 10, 35, 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 (5.1):

$$\text{for all } n \geq 1,\ X_n^{(\theta)} = \psi_\xi\big(X_{n-1}^{(\theta)}\big) + B_\xi\big(X_{n-1}^{(\theta)}\big)\vartheta_n^{(\gamma)},$$

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 random variables with common parametric PDF denoted by $\mathfrak{p}_\gamma$ (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 $P_\theta(x, A) := \int_{\mathbb{R}} \mathbf{1}_A(z) p_\theta(x, z)\,\mathrm{d}z$ ($A \in \mathcal{X}$), with

$$p_\theta(x, z) := |\det B_\xi(x)|^{-1}\mathfrak{p}_\gamma(B_\xi(x)^{-1}(z - \psi_\xi(x))), \tag{5.2}$$

the following remarks are relevant to investigating Assumptions 1.1–1.4 of Theorem 1.1.

Remark 5.1. If the PDF $\mathfrak{p}_{\gamma_0}$ 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 1.1. More precisely, in this case, Assumption 1.1 holds with any real number ρ_a (and the associated constant C_a) such that $\mathbb{E}\big[L_{F_{\xi_0}}\big(\vartheta_1^{(\gamma_0)}\big)^a\big]^{1/a} < \rho_a < 1$, where $F_{\xi_0}(x, v) = \psi_{\xi_0}(x) + B_{\xi_0}(x)v$.

Remark 5.2. The moment/contractive conditions (1.4) and (1.5) related to $\theta_0 = (\xi_0, \gamma_0)$ in Remark 5.1 involve expectations which depend on the above function F_{ξ_0} and on the PDF $\mathfrak{p}_{\gamma_0}$. Hence, the conditions in Assumptions 1.2 and 1.3 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).

Remark 5.3. Thanks to (5.2), Assumption 1.4 holds provided that, for every $A > 0$,

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

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 complement in order to use the 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 (5.1) under roundoff error. If $(h_\theta)_{\theta\in\Theta}$ is the roundoff family with h_θ close to $h_0 = \mathrm{id}$ when $\theta \to 0$, then the roundoff transition kernel $P_\theta(x, A) = P\big(x, h_\theta^{-1}(A)\big)$ is written as $P_\theta(x, A) := \int_{\mathbb{R}} \mathbf{1}_A(z) p_\theta(x, z)\,\mathrm{d}z$ with

$$p_\theta(x, z) := \Gamma_\theta(z)\mathfrak{p}(B(x)^{-1}(g_\theta(z) - \psi(x))) \tag{5.3}$$

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 $\Gamma_\theta(z) := |\det B_\xi(x)|^{-1}|\det \nabla g_\theta(z)|$. Using the kernels in (5.3), Remarks 5.1–5.3 then apply.

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

5.2. Robustness of IFS under thresholding/truncation

Here we consider $\mathbb{X} := \mathbb{R}^d$ ($d \geq 1$) equipped with the Euclidean norm $\|\cdot\|$, and $\mathbb{V} := \mathbb{R}^q$ ($q \geq 1$) 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,

$$\text{for } X_0 \in \mathbb{R}^d \text{ and all } n \geq 1, \quad X_n := F(X_{n-1}, \vartheta_n), \tag{5.4}$$

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 probability distribution of ϑ_1 is absolutely continuous with respect to Lebesgue's measure on $\mathbb{R}^q$, with PDF denoted by $\mathfrak{p}$. 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 (5.4) caused by some thresholding and/or truncation? Such an issue may arise as soon as a numerical implementation of the model is considered. Thus, let us investigate the robustness of the IFS (5.4) when thresholding the function F on the infinite set $\mathbb{X}$ and truncating the PDF $\mathfrak{p}$ 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:

$$\text{for all } x \in \mathbb{R}^d, \quad \Phi_\xi(x) = \min\left(\frac{\xi}{\|x\|}, 1\right)x = \begin{cases} x & \text{if } \|x\| \leq \xi, \\ \xi \dfrac{x}{\|x\|} & \text{if } \|x\| > \xi. \end{cases}$$

Moreover, for any $\gamma \in (0, +\infty)$, define the truncated PDF $\mathfrak{p}_\gamma$ at level γ, for all $v \in \mathbb{R}^q$, by

$$\mathfrak{p}_\gamma(v) = c_\gamma \mathfrak{p}(v)\mathbf{1}_{B(0,\gamma)}(v) \text{ with } c_\gamma := \left(\int_{B(0,\gamma)} \mathfrak{p}(v)\,\mathrm{d}v\right)^{-1},$$

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 perturbed IFS $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ defined by $X_0^{(\theta)} \in \mathbb{X}$ and

$$\text{for all } n \geq 1, \quad X_n^{(\theta)} := F_\xi\big(X_{n-1}^{(\theta)}, \vartheta_n^{(\gamma)}\big), \quad \text{with } F_\xi(x, v) := \Phi_\xi(F(x, v)), \tag{5.5}$$

where the sequence $\big\{\vartheta_n^{(\gamma)}\big\}_{n\geq 1}$ of $\mathbb{R}^q$-valued i.i.d. random variables is assumed to admit the common PDF $\mathfrak{p}_\gamma$. Note that the stability of quantitative bounds for Markov chains via truncation rather than thresholding is studied in [34]. 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 [34, Section 3.2, Theorem 9], since the resulting perturbed process is no longer an IFS of Lipschitz maps. Morevover, note that the study of $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ does not fit into 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, Proposition 5.1 is stated in the general framework of Definition 1.1 up to the condition of absolute continuity of the probability distribution of ϑ_1. 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 (5.4) satisfies Definition 1.1, with ϑ_1 having a PDF on $\mathbb{R}^q$. Moreover, suppose that Assumption 1.1 holds for some $a \geq 1$, and that $\widetilde{M}_a := \mathbb{E}\big[\|F(0, \vartheta_1)\|^a\big]^{1/a} < \infty$, $\widetilde{\kappa}_a := \mathbb{E}[L_F(\vartheta_1)^a]^{1/a} < 1$, and $\mathbb{E}[\|\vartheta_1\|^a] < \infty$. Let $\kappa_a \in (\widetilde{\kappa}_a, 1)$, and let $\Theta := (0, +\infty) \times (\gamma_0, +\infty)$, with $\gamma_0 > 0$ defined by*

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

the condition, for all $\gamma > \gamma_0$, $c_\gamma \le (\kappa_a/\widetilde{\kappa}_a)^a$. Then, the perturbed IFS $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ defined by (5.5) with $\theta \in \Theta$ satisfies P1–P3 of Theorem 1.1 with $\Delta_\theta \to 0$ when $\xi \to +\infty$ and $\gamma \to +\infty$. More precisely, for every $\varepsilon \in (0, 2)$ define $A_\varepsilon = 2^a\varepsilon^{-a} - 1$. Then we have $\Delta_\theta \le \varepsilon$, provided that $\theta := (\xi, \gamma) \in \Theta$ is such that

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

Appendix A. Proof of (1.7)

Suppose that Assumptions 1.2 and 1.3 are fulfilled. Then we prove the drift inequality (1.7) in Section 1. In fact, for any $\kappa \in (\kappa_a, 1)$, we prove that the following strengthened inequality holds:

$$\text{for all } \theta \in \Theta, \quad P_\theta V_a \le \delta_a V_a + K_a \mathbf{1}_{[-r_a, r_a]}, \tag{A.1}$$

where the constants $\delta_a < 1$ and $K_a > 0$ are given in (1.7), and where $r_a := (1 + M_a + \kappa_a - \kappa)/(\kappa - \kappa_a)$. We have, for any $\theta \in \Theta$ and any $x \in \mathbb{X}$,

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

using Holder's inequality. It follows from Assumptions 1.2 and 1.3 that

$$\text{for all } \theta \in \Theta \text{ and } x \in \mathbb{X}, \quad \left(\frac{(P_\theta V_a)(x)}{V_a(x)}\right)^{1/a} \le \frac{1}{1 + d(x; x_0)} + \kappa_a + \frac{M_a}{1 + d(x; x_0)}. \tag{A.2}$$

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 $d(x; x_0) > r_a$,

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

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

$$(P_\theta V_a)(x) \le \kappa^a V_a(x). \tag{A.3}$$

Moreover, for every $\theta \in \Theta$ and for every $x \in \mathbb{X}$ such that $d(x; x_0) \le r_a$, we deduce from (A.2) that

$$(P_\theta V_a)(x) \le F_a V_a(x) \le F_a(1 + r_a)^a, \tag{A.4}$$

where $F_a := (1 + \kappa_a + M_a)^a$. Finally, combining (A.3) and (A.4) provides (A.1), and then (1.7) with $\delta_a := \kappa_a{}^a < 1$ and $K_a := F_a(1 + r_a)^a > 0$.

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

Appendix B. Complements to Proposition 2.1

First, we prove (2.3).

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

$$L(v) := \sup_{(x,y)\in\mathbb{R}^2,\, x\neq y} \frac{|F(x, v) - F(y, v)|}{|x - y|} = \max\big(|\alpha - \sqrt{\lambda}v|; |\alpha + \sqrt{\lambda}v|\big). \tag{B.1}$$

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

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

If $v = 0$, (B.1) is obvious. Assume that $v > 0$. Then F_v' is strictly increasing, so that

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

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

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

and the same conclusion holds. That $L(v) \ge \max(|\alpha - \sqrt{\lambda}v|; |\alpha + \sqrt{\lambda}v|)$ follows from the inequality $L(v) \ge |F_v'(x)|$ for any $x \in \mathbb{R}$, which is easily deduced from the definition of $L(v)$ in (B.1). Hence, we obtain $L(v) \ge \lim_{x\pm\infty} |F_v'(x)|$. The proof of (B.1) 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*

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

Then, for any $A > 0$,

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

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

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

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

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

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* $f \in \mathbb{L}^1(\mathbb{R})$, $\lim_{(a,b)\to(0,1)} \int_{\mathbb{R}} |f(a+bz) - f(z)|\, \mathrm{d}z = 0$.

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 $b \geq \frac{1}{2}$,

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

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

Appendix C. Proof of (4.5) and (4.6) under the assumptions (4.2) and (4.4) with $N = 1$

Thoughout this section, the conditions (4.2) and (4.4) with $N = 1$ are assumed to hold. Note that (4.4) with $N = 1$ is $\kappa_{1,a} = \mathbb{E}[L(\vartheta_1)^a]^{1/a} < 1$. We prove (4.5) and (4.6) of Proposition 4.1 with explicit constants. Under the general assumption $\widehat{\kappa}_a < 1$ of Assumption 4.1, the proofs of (4.5) and (4.6) are similar (replace P with P^N where N is 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 (A.1) 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 notation. 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 $\|\mu\|_a := \left(\int_{\mathbb{X}} V_a(y)\, \mathrm{d}\mu(y)\right)^{1/a} < \infty$. Finally, for $n \in \mathbb{N}$ and for any probability measures μ_1 and μ_2 on $\mathbb{X}$, define $\Delta_n(\mu_1, \mu_2) := d\big(X_n^{\mu_1}, X_n^{\mu_2}\big)\, \big(p(X_n^{\mu_1}) + p(X_n^{\mu_2})\big)^{a-1}$.

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

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

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

$$\mathbb{E}\big[|f(X_n^{\mu_1}) - f(X_n^{\mu_2})|\big] \leq \xi^{(a-1)/a} m_a(f) \kappa_{1,a}^n \mathbb{E}\big[d\big(X_0^{\mu_1}, X_0^{\mu_2}\big)\big](\|\mu_1\|_a + \|\mu_2\|_a)^{a-1}. \tag{C.2}$$

Proof. Note that $X_n^\mu = F_{\vartheta_n:\vartheta_1} X_0^\mu$ from Definition 1.1 and the notation introduced in (4.1). If $a := 1$, then (C.1) follows from the independence of the ϑ_n and from the definitions of $L(v)$ and $\kappa_{1,a}$. Now assume that $a \in (1, +\infty)$. Without loss of generality, we can suppose that the sequence $\{\vartheta_n\}_{n\geq 1}$ is independent of $(X_0^{\mu_1}, X_0^{\mu_2})$. Also note that, if $\mu \in \mathcal{M}_a$, then we have $\mathbb{E}\big[p(X_n^\mu)^a\big] = \int_{\mathbb{X}} (P^n V_a)(x)\, \mathrm{d}\mu(x) \leq \xi \|\mu\|_a^a$. From Holder's inequality (use $1 =$

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

$1/a + (a-1)/a$), we obtain

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

This proves (C.1); (C.2) follows from (C.1) and the definition of $m_a(f)$. □

Now recall that we consider the case $N = 1$ in (4.4). Let us prove the inequality (4.5) in this case (that is, (4.7)). Applying (C.2) to $\mu_1 := \delta_x$ and $\mu_2 := \pi$ gives

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

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

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

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

Finally, to prove (4.6), it remains to study $m_a(P^n f)$ for $f \in \mathcal{L}_a$. Applying (C.2) to $\mu_1 := \delta_x$ and $\mu_2 := \delta_y$ for any $(x, y) \in \mathbb{X}^2$ gives

$$\text{for all } f \in \mathcal{L}_a, \quad |P^n f(x) - P^n f(y)| \le \xi^{(a-1)/a} m_a(f)\kappa_{1,a}^n d(x, y)(p(x) + p(y))^{a-1}.$$

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

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

Combining the last inequality with (4.7) gives (4.6).

Appendix D. Proof of Proposition 5.1

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

$$\mathbb{E}\big[\big\|F_\xi\big(0, \vartheta_1^{(\gamma)}\big)\big\|^a\big] = \int_{\mathbb{R}^d} \|\Phi_\xi(F(0, v))\|^a \mathfrak{p}_\gamma(v)\,\mathrm{d}v \le \left(\frac{\kappa_a}{\widetilde{\kappa}_a}\right)^a \int_{\mathbb{R}^q} \|F(0, v)\|^a \mathfrak{p}(v)\,\mathrm{d}v$$

is finite, so Assumption 1.2 holds with $x_0 = 0$ and $M_a := \widetilde{M}_a\kappa_a/\widetilde{\kappa}_a$. Moreover, note that $L_{\Phi_\xi} \le 1$, so that, for all $v \in \mathbb{V}$, $L_{F_\xi}(v) \le L_F(v)$. Hence, we have, for every $\theta = (\xi, \gamma) \in \Theta$,

$$\mathbb{E}\big[L_{F_\xi}\big(\vartheta_1^{(\gamma)}\big)^a\big] = \int_{\mathbb{R}^d} L_{F_\xi}(v)^a \mathfrak{p}_\gamma(v)\,\mathrm{d}v \le c_\gamma \int_{\mathbb{R}^d} L_F(v)^a \mathfrak{p}(v)\,\mathrm{d}v \le c_\gamma\widetilde{\kappa}_a^a \le \kappa_a^a.$$

Thus, Assumption 1.3 holds. It remains to check Assumption 1.4 and to specify the error term Δ_θ. Let P (respectively P_θ) denote the transition kernel of the unperturbed IFS $\{X_n\}_{n\in\mathbb{N}}$

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

(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 $|f|_0 \le 1$. First, note that we have, for every $x \in \mathbb{R}^d$ satisfying $\|x\| > A_\varepsilon$,

$$\frac{|(P_\theta f)(x) - (Pf)(x)|}{V_a(x)} \le \frac{2}{V_a(x)} \le \varepsilon \tag{D.1}$$

by the 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 subsets $E_{\theta,x}$ and $G_{\theta,x}$ of $\mathbb{R}^q$:

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

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

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

$$\begin{aligned}(P_\theta f)(x) &= \int_{\mathbb{R}^q} f(F_\xi(x, v))\mathfrak{p}_\gamma(v)\,\mathrm{d}v \\ &= c_\gamma \int_{E_{\theta,x}} f(F(x, v))\mathfrak{p}(v)\,\mathrm{d}v + c_\gamma \int_{G_{\theta,x}} f(\eta_{x,v})\mathfrak{p}(v)\,\mathrm{d}v,\end{aligned}$$

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

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

from the definition of $\mathbb{R}^q \setminus E_{\theta,x}$ and the condition $c_\gamma \le (\kappa_a/\widetilde{\kappa}_a)^a$. Now let $x \in \mathbb{R}^d$ be such that $\|x\| \le A_\varepsilon$. Then, for all $v \in \mathbb{R}^q$, $\|F(x, v)\| \le L_F(v)A_\varepsilon + \|F(0, v)\|$ from $F(x, v) = (F(x, v) - F(0, v)) + F(0, v)$ and the triangle inequality. Therefore, $\big[L_F(v)A_\varepsilon \le \xi/2$ and $\|F(0, v)\| \le \xi/2\big] \Rightarrow \|F(x, v)\| \le \xi$, from which we deduce that

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

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

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

The conclusion of Proposition 5.1 follows from this and (D.1).

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

Acknowledgement

We thank the reviewers for their valuable comments, which have improved the manuscript. In particular the material of Section 5.2 addresses an issue raised by one of the reviewers.

Funding information

There are no funding bodies to thank relating to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ALQUIER, P., FRIEL, N., EVERITT, R. AND BOLAND, A. (2016). Noisy Monte Carlo: Convergence of Markov chains with approximate transition kernels. *Statist. Comput.* **26**, 29–47.
[2] ALSMEYER, G. (2003). On the Harris recurrence of iterated random Lipschitz functions and related convergence rate results. *J. Theoret. Prob.* **16**, 217–247.
[3] ALTMAN, E., AVRACHENKOV, K. E. AND NÚÑEZ-QUEIJA, R. (2004). Perturbation analysis for denumerable Markov chains with application to queueing models. *Adv. Appl. Prob.* **36**, 839–853.
[4] BALADI, V. (2000). *Positive Transfer Operators and Decay of Correlations* (Adv. Ser. Nonlinear Dynamics **16**). World Scientific, River Edge, NJ.
[5] BAXENDALE, P. H. (2005). Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Prob.* **15**, 700–738.
[6] BENDA, M. (1998). A central limit theorem for contractive stochastic dynamical systems. *J. Appl. Prob.* **35**, 200–205.
[7] BREYER, L., ROBERTS, G. O. AND ROSENTHAL, J. S. (2001). A note on geometric ergodicity and floating-point roundoff error. *Statist. Prob. Lett.* **53**, 123–127.
[8] CLINE, D. B. H. (2007). Regular variation of order 1 nonlinear AR-ARCH models. *Stoch. Process. Appl.* **117**, 840–861.
[9] CLINE, D. B. H. (2007). Stability of nonlinear stochastic recursions with application to nonlinear AR-GARCH models. *Adv. Appl. Prob.* **39**, 462–491.
[10] CLINE, D. B. H. AND PU, H. H. (2002). A note on a simple Markov bilinear stochastic process. *Statist. Prob. Lett.* **56**, 283–288.
[11] DOUC, R. AND MOULINES, E. (2012). Asymptotic properties of the maximum likelihood estimation in misspecified hidden Markov models. *Ann. Statist.* **40**, 2697–2732.
[12] DOUC, R., MOULINES, E., PRIOURET, P. AND SOULIER, P. (2018). *Markov Chains*. Springer, New York.
[13] DUFLO, M. (1997). *Random Iterative Models*. Springer, Berlin.
[14] FERRÉ, D., HERVÉ, L. AND LEDOUX, J. (2013). Regular perturbation of V-geometrically ergodic Markov chains. *J. Appl. Prob.* **50**, 184–194.
[15] GALTCHOUK, L. AND PERGAMENSHCHIKOV, S. (2014). Geometric ergodicity for classes of homogeneous Markov chains. *Stoch. Process. Appl.* **124**, 3362–3391.
[16] GOUËZEL, S. AND LIVERANI, C. (2006). Banach spaces adapted to Anosov systems. *Ergodic Theory Dynam. Systems* **26**, 189–217.
[17] GREENWOOD, P. E. AND WEFELMEYER, W. (1998). Maximum likelihood estimator and Kullback–Leibler information in misspecified Markov chain models. *Theory Prob. Appl.* **42**, 103–111.
[18] GUIBOURG, D., HERVÉ, L. AND LEDOUX, J. (2011). Quasi-compactness of Markov kernels on weighted-supremum spaces and geometrical ergodicity. Preprint, arXiv:1110.3240.
[19] HAIRER, M. AND MATTINGLY, J. C. (2011). Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, ed. R. Dalang, M. Dozzi and F. Russo. Springer, Basel, pp. 109–117.
[20] HENNION, H. (1993). Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc*. **118**, 627–634.
[21] HENNION, H. AND HERVÉ, L. (2001). *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness* (Lect. Notes Math. **1766**). Springer, New York.

https://doi.org/10.1017/jpr.2022.107 Published online by Cambridge University Press

[22] HERVÉ, L. (2008). Quasi-compactness and mean ergodicity for Markov kernels acting on weighted supremum normed spaces. *Ann. Inst. H. Poincaré Prob. Statist.* **44**, 1090–1095.
[23] HERVÉ, L. AND LEDOUX, J. (2014). Approximating Markov chains and V-geometric ergodicity via weak perturbation theory. *Stoch. Process. Appl.* **124**, 613–638.
[24] HERVÉ, L. AND LEDOUX, J. (2014). Spectral analysis of Markov kernels and aplication to the convergence rate of discrete random walks. *Adv. Appl. Prob.* **46**, 1036–1058.
[25] JOHNDROW, J. E., MATTINGLY, J. C., MUKHERJEE, S. AND DUNSON, D. Optimal approximating Markov chains for Bayesian inference. Preprint, arXiv:1508.03387v3.
[26] KARTASHOV, N. V. (1986). Inequalities in theorems of ergodicity and stability for Markov chains with common phase space I. *Theory Prob. Appl.* **30**, 247–259.
[27] KARTASHOV, N. V. (1986). Inequalities in theorems of ergodicity and stability for Markov chains with common phase space II. *Theory Prob. Appl.* **30**, 505–515.
[28] KELLER, G. (1982). Stochastic stability in some chaotic dynamical systems. *Monatshefte Math.* **94**, 313–333.
[29] KELLER, G. AND LIVERANI, C. (1999). Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa – Classe di Scienze (4)* **XXVIII**, 141–152.
[30] LE PAGE, É. (1983). Théorèmes de renouvellement pour les produits de matrices aléatoires. Équations aux différences aléatoires. In *Séminaires de probabilités Rennes 1983*. Publ. Sém. Math. Univ. Rennes I, Rennes, p. 116.
[31] LIVERANI, C. (2004). Invariant measures and their properties. A functional analytic point of view. In *Dynamical Systems. Part II: Topological Geometrical and Ergodic Properties of Dynamical Systems*, ed. S. Marmi (Publ. Scuola Normale Superiore). Springer, New York.
[32] LUND, R. B. AND TWEEDIE, R. L. (1996). Geometric convergence rates for stochastically ordered Markov chains. *Math. Operat. Res.* **21**, 182–194.
[33] MEDINA-AGUAYO, F. J., LEE, A. AND ROBERTS, G. O. (2016). Stability of noisy Metropolis–Hastings. *Statist. Comput.* **26**, 1187–1211.
[34] MEDINA-AGUAYO, F. J., RUDOLF, D. AND SCHWEIZER, N. (2020). Perturbation bounds for Monte Carlo within Metropolis via restricted approximations. *Stoch. Process. Appl.* **130**, 2200–2227.
[35] MEITZ, M. AND SAIKKONEN, P. (2010). A note on the geometric ergodicity of a nonlinear AR-ARCH model. *Statist. Prob. Lett.* **80**, 631–638.
[36] MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
[37] MEYN, S. P. AND TWEEDIE, R. L. (1994). Computable bounds for geometric convergence rates of Markov chains. *Ann. Prob.* **4**, 981–1011.
[38] MILHAUD, X. AND RAUGI, A. (1989). Etude de l'estimateur du maximum de vraisemblance dans le cas d'un processus autorégressif : convergence, normalité asymptotique, vitesse de convergence. *Ann. Inst. H. Poincaré Prob. Statist.* **25**, 383–428.
[39] MITROPHANOV, A. Y. (2005). Sensitivity and convergence of uniformly ergodic Markov chains. *J. Appl. Prob.* **42**, 1003–1014.
[40] MOUHOUBI, Z. AND ASSANI, D. (2010). New perturbation bounds for denumerable Markov chains. *Linear Algebra Appl.* **432**, 1627–1649.
[41] NEGREA, J. AND ROSENTHAL, J. S. Approximations of geometrically ergodic reversible Markov chains. *Adv. Appl. Prob.* **53**, 981–1022.
[42] ROBERTS, G. O., ROSENTHAL, J. S. AND SCHWARTZ, P. O. (1998). Convergence properties of perturbed Markov chains. *J. Appl. Prob.* **35**, 1–11.
[43] RUDOLF, D. AND SCHWEIZER, N. (2018). Perturbation theory for Markov chains via Wasserstein distance. *Bernoulli* **24**, 2610–2639.
[44] SCHEFFÉ, H. (1947). A useful convergence theorem for probability distributions. *Ann. Math. Statist.* **18**, 434–438.
[45] SHARDLOW, T. AND STUART, A. M. (2000). A perturbation theory for ergodic Markov chains and application to numerical approximations. *SIAM J. Numer. Anal.* **37**, 1120–1137.
[46] TSAY, R. S. (2010). *Analysis of Financial Time Series*, 3rd edn. John Wiley, Hoboken, NJ.
[47] WU, L. (2004). Essential spectral radius for Markov semigroups. I. Discrete time case. *Prob. Theory Relat. Fields* **128**, 255–321.