Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2015, Vol. 51, No. 2, 781–808
DOI: 10.1214/13-AIHP592

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Parametric first-order Edgeworth expansion for Markov additive functionals. Application to M-estimations

D. Ferré

L.A.R.E.M.A. (UMR-CNRS 6093), Université d'Angers, UFR sciences, 2 bd Lavoisier, 49045 Angers Cedex 01, France.
E-mail: deborah.ferre@univ-angers.fr

Received 17 February 2012; revised 25 September 2013; accepted 4 November 2013

Abstract. We give a spectral approach to prove a parametric first-order Edgeworth expansion for bivariate additive functionals of strongly ergodic Markov chains. In particular, given any V-geometrically ergodic Markov chain $(X_n)_{n\in\mathbb{N}}$ whose distribution depends on a parameter θ, we prove that $\{\xi_p(X_{n-1},X_n);\ p\in\mathcal{P}, n\geq 1\}$ satisfies a uniform (in (θ,p)) first-order Edgeworth expansion provided that $\{\xi_p(\cdot,\cdot);\ p\in\mathcal{P}\}$ satisfies some non-lattice condition and an almost optimal moment domination condition. Furthermore, the sequence $(X_n)_{n\in\mathbb{N}}$ need not be stationary. This result is applied to M-estimators of Markov chains and in particular of V-geometrically ergodic Markov chains. The M-estimators of some autoregressive processes are studied.

Résumé. Grâce à une approche spectrale, nous donnons des conditions assurant la validité du développement d'Edgeworth d'ordre 1 paramétrique, dans le cadre général des fonctionnelles bivariées et additives de chaînes de Markov fortement ergodiques. En particulier, soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov V-géométriquement ergodique dont la loi dépend d'un paramètre θ. Nous montrons alors que $\{\xi_p(X_{n-1},X_n);\ p\in\mathcal{P}, n\geq 1\}$ satisfait un développement d'Edgeworth d'ordre 1 uniforme (en (θ,p)) si $\{\xi_p(\cdot,\cdot);\ p\in\mathcal{P}\}$ satisfait une condition de type non-lattice ainsi qu'une condition quasi-optimale de moment-domination. De plus, ce résultat est établi dans le cas où les données $(X_n)_{n\in\mathbb{N}}$ ne sont pas nécessairement stationnaires. Ce résultat est appliqué en particulier aux M-estimateurs associés à des chaînes de Markov V-géométriquement ergodiques. Les M-estimateurs de processus autorégressifs sont étudiés.

MSC: 60F05; 60J05; 62F12; 62M05

Keywords: Edgeworth expansion; V-geometrically ergodic Markov chain; Non-arithmeticity condition; Perturbation operator

1. Introduction

Let $(E,\mathcal{E})$ be any measurable space, and let $(X_n)_{n\geq 0}$ be a Markov chain on a general state space E with transition kernel $(Q_\theta(x,\cdot);\ x\in E)$ where θ is a parameter in some set Θ. The initial distribution of the chain is denoted by μ_θ. The underlying probability measure is denoted by $\mathbb{P}_{\theta,\mu_\theta}$.

Let $\{\xi_p(\cdot,\cdot);\ p\in\mathcal{P}\}$ be a family of measurable functions from E^2 into $\mathbb{R}$, where $\mathcal{P}$ is any set. Let us define the following bivariate additive functionals

$$\forall n\geq 1, \forall p\in\mathcal{P},\qquad S_n(p):=\sum_{k=1}^{n}\xi_p(X_{k-1},X_k). \tag{1}$$

We are interested in appropriate conditions on the model, on the family $\{\xi_p(\cdot,\cdot);\ p\in\mathcal{P}\}$ and on the initial probability measure μ_θ, under which a first-order Edgeworth expansion exists (also called Esseen theorem), namely there exist a polynomial function $A_{\theta,p}(\cdot)$ and a positive real number $\sigma_{\theta,p}$ such that

$$\sup_{(\theta,p)\in\Theta\times\mathcal{P}}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\mu_\theta}\left\{\frac{S_n(p)}{\sigma_{\theta,p}\sqrt{n}}\le u\right\}-\mathcal{N}(u)-\eta(u)n^{-1/2}A_{\theta,p}(u)\right|=\mathrm{o}\big(n^{-1/2}\big),\tag{2}$$

where $\mathcal{N}$ is the standard normal distribution function and η is its density. Note that expansion (2) holds uniformly in $(\theta,p)\in\Theta\times\mathcal{P}$.

One of these appropriate conditions is the following non-arithmeticity condition:

Hypothesis (N-A) (Non-arithmeticity). *For any compact subset K_0 of $\mathbb{R}^*$, there exists $\rho\in[0,1)$ such that for all $n\ge 1$, $\sup\{|\mathbb{E}_{\theta,\mu_\theta}[\mathrm{e}^{\mathrm{i}tS_n(p)}]|;\ t\in K_0,(\theta,p)\in\Theta\times\mathcal{P}\}=\mathrm{O}(\rho^n)$.*

This non-arithmeticity condition may be satisfied under the following non-lattice condition:

Hypothesis (N-L)′. *For all $p\in\mathcal{P}$, there exist neither continuous function $\mathcal{A}_p(\cdot):E\to\mathbb{R}$ nor constant C_p such that we have for all $(x,y)\in E^2$, $\xi_p(x,y)=\mathcal{A}_p(y)-\mathcal{A}_p(x)+C_p$.*

As illustrated later in M-estimation, the bivariate and parametric form of (1), as well as the previous uniform control and the possible non-stationarity of μ_θ, are required for statistical applications.

Edgeworth expansions in the Markov setting can be established by the two following methods:

1. *The regeneration method*. This standard method, introduced by [29], was used by Bolthausen [4] to establish the Berry–Esseen theorem for univariate additive functionals of the form $S_n=\sum_{k=1}^n\xi(X_k)$, by splitting S_n into a sum of independent blocks. This method can be applied to the general class of Harris-recurrent chains $(X_n)_{n\ge0}$ which either possess an accessible atom or satisfy some minorization condition. Bolthausen's work was extended to Edgeworth expansions by Malinovskii [23] and next generalized to bivariate additive functionals $S_n=\sum_{k=1}^n\xi(X_{k-1},X_k)$ by Jensen [19].

   Note that in [4,19,23], neither the distribution of $(X_n)_{n\ge0}$ nor the function ξ depends on parameters. However a recent work due to Bertail and Clémençon [2] provides a Berry–Esseen theorem adapted to the above mentioned parametric setting (just mention that the regeneration method allows to establish the Berry–Esseen theorem for "studentized functionals"), but the extension to Edgeworth expansions would generate even more difficulties. Furthermore this statement only concerns univariate additive functionals and the extension of their proof to the bivariate case (1) induces dependence between the regeneration blocks and hence provides at least one more difficulty to handle with.
2. *The weak Nagaev–Guivarc'h spectral method*. This method, based on the Keller–Liverani perturbation theorem [20], enables the statement of limit theorems for additive functionals associated to strongly ergodic Markov chains (Harris recurrence is no more required). This method has been fully described in [18] in the case of univariate additive functionals. It is specially efficient for ρ-mixing and V-geometrically Markov chains, as well as for iterated function systems. In those models, the extension of Berry–Esseen type results of [18] to the case of bivariate additive functionals of the type (1) has already been obtained in [11,12,17] with in addition the desired control on the parameters (θ,p). The resulting moment conditions on $\{\xi_p;\ p\in\mathcal{P}\}$ are (almost) optimal with respect to the independent case. Let us note that they are explicit for these three models and do not depend on the initial probability (unlike the ones given by the regeneration method).

In this paper, we will state expansion (2) for the class of strongly ergodic Markov chains, and apply Fourier techniques via the perturbation operator theory of Nagaev–Guivarc'h.

Our work extends the Berry–Esseen type results of Hervé, Ledoux and Patilea [17] to the first-order Edgeworth expansion. As in the independent case, the gap from Berry–Esseen to Edgeworth type results induces at least a new difficulty: the requirement of the non-arithmeticity hypothesis.

In Section 2.1, we consider a family of random variables (r.v.) $S_n(p)$ (not necessarily derived from Markovian models) defined on a general parametric probability space $(\Omega,\mathcal{F},\{\mathbb{P}_\theta;\theta\in\Theta\})$, and we state hypotheses called $\mathcal{R}(m)$

and (N-A) under which expansion (2) holds true. These hypotheses concern the behavior of the characteristic function $t \mapsto \phi_{n,p}(t)$ of $S_n(p)$: Hypothesis $\mathcal{R}(m)$ focuses on the form and the regularity of $\phi_{n,p}$ near $t = 0$; whereas Hypothesis (N-A) focuses on the behavior of $\phi_{n,p}$ outside $t = 0$.

In Section 2.2, we specify the form of $S_n(p)$: from now on, $S_n(p)$ is defined by (1) where $(X_n)_{n\geq 0}$ is assumed to be a strongly ergodic Markov chain, and we give a brief review of the weak Nagaev–Guivarc'h spectral method to check Hypothesis $\mathcal{R}(m)$ and (N-A) in this Markov context. In fact, as already done in [17], Hypothesis $\mathcal{R}(m)$ can be investigated thanks to an easy extension of the results of [18].

By contrast, the method developed in [18] is not sufficient to study Hypothesis (N-A). Indeed, the non-arithmeticity condition has to be checked uniformly in both the parameter θ of the Markovian model and the parameter p of the family $\{\xi_p;\, p \in \mathcal{P}\}$ involved in (1). The study of (N-A) in this context is original and constitutes an important part of this work (actually, even in the independent case, this question is far from being obvious). In our Markov setting, this study is based on the operator perturbation theory, quasi-compactness arguments and Ascoli theorem. Specifically, in Section 3, we give three approaches to reduce Hypothesis (N-A) to some simple non-lattice conditions in the case of general strongly ergodic Markov chains.

Section 4 is devoted to V-geometrically ergodic Markov chains. For this instance and more specifically for dominated models,[1] we reduce (N-A) using one of the three approaches presented in Section 3.3. Combining this result together with the sufficient conditions of [17] to check Hypothesis $\mathcal{R}(m)$ and the general Edgeworth type statement of Section 2.1, provides expansion (2) under assumptions close to the ones of the independent case.

At last, statistical applications are studied in Section 5: a first-order Edgeworth expansion for M-estimators of dominated V-geometrically ergodic Markov chains is derived from the results of Section 4.

More precisely, let α_0 be the so-called true value of some real parameter of interest and $(\hat{\alpha}_n)_{n\in\mathbb{N}^*}$ its estimator of the form $\hat{\alpha}_n := \arg\min_{\alpha\in\mathcal{A}} \frac{1}{n}\sum_{k=1}^n F(\alpha, X_{k-1}, X_k)$, where $(X_n)_{n\geq 0}$ is some dominated V-geometrically ergodic Markov chain (whose underlying probability measure is still denoted by $\mathbb{P}_{\theta,\mu_\theta}$). We state in Section 5 the appropriate conditions under which there exist a polynomial function $A_\theta(\cdot)$ and a positive real number $\sigma(\theta)$ (both explicitly defined in Theorem 2) such that

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\mu_\theta}\left\{\frac{\sqrt{n}}{\sigma(\theta)}(\hat{\alpha}_n - \alpha_0) \leq u\right\} - \mathcal{N}(u) - \eta(u)n^{-1/2}A_\theta(u)\right| = \mathrm{o}\big(n^{-1/2}\big).$$

This theorem, which extends Pfanzagl theorem [25] obtained for independent and identically distributed (i.i.d.) data under some moment conditions of order 3, is valid under a natural adaptation of the statistical regularity conditions of [25], moment domination conditions of order $3 + \varepsilon$, and some simple non-lattice condition as well. To the best of our knowledge, this result is new. Notice that our moment domination conditions are not only almost optimal, but also take the same form as the ones used in [6] to prove the asymptotic normality of M-estimators under V-geometrically ergodicity.

Whereas statistical applications studied in Section 5 only concern dominated models, the results of Section 6 are much more general. Indeed the adaptation of Pfanzagl proof is developed in Section 6 for general statistical models under Hypotheses $\mathcal{R}(3)$ and (N-A). Note that this adaptation is not straightforward. Finally, the results of this section are applied in Section 6.4 to M-estimators of an example of non-dominated V-geometrically ergodic Markov chains: some AR(d) processes with $d > 1$.

2. Fourier techniques and first-order Edgeworth expansion

In this section, we present some results based on Fourier techniques. These results appeal to the next Hypotheses $\mathcal{R}(m)$ and (N-A) that are well-suited for the markovian case as explained in Section 2.2.

2.1. *Hypotheses* $\mathcal{R}(m)$ *and* (N-A) *and first-order Edgeworth expansion*

Let $(\Omega, \mathcal{F}, \{\mathbb{P}_\theta; \theta \in \Theta\})$ be any statistical model, where Θ is some parameter space. The underlying expectation is denoted by $\mathbb{E}_\theta$. Consider a family $\{S_n(p);\, n \in \mathbb{N}^*,\, p \in \mathcal{P}\}$ of real r.v. defined on $(\Omega, \mathcal{F}, \{\mathbb{P}_\theta; \theta \in \Theta\})$, where $\mathcal{P}$ is any set. Note that, since we will get results uniformly in (θ, p), they can be applied when the parameter p depends on θ.

[1] This choice is only for convenience but is not necessary, see, for example, Section 6.4.

Hypothesis $\mathcal{R}(m)$, $m \in \mathbb{N}^*$. *There exists a bounded open interval $I_0 \subset \mathbb{R}$ of $t = 0$ such that one has for all $(\theta, p) \in \Theta \times \mathcal{P}$, $n \geq 1$, $t \in I_0$*

$$\mathbb{E}_\theta\big[\mathrm{e}^{\mathrm{i}tS_n(p)}\big] = \lambda_{\theta,p}(t)^n\big(1 + l_{\theta,p}(t)\big) + r_{\theta,p,n}(t), \tag{3}$$

where $\lambda_{\theta,p}(\cdot)$, $l_{\theta,p}(\cdot)$ and $r_{\theta,p,n}(\cdot)$ are $\mathbb{C}$-valued functions of class $\mathcal{C}^m$ on I_0 satisfying the following properties:

$$\lambda_{\theta,p}(0) = 1, \qquad \lambda_{\theta,p}^{(1)}(0) = 0, \qquad l_{\theta,p}(0) = 0, \qquad r_{\theta,p,n}(0) = 0,$$

and for $\ell = 0, \dots, m$

$$\sup\big\{\big|\lambda_{\theta,p}^{(\ell)}(t)\big| + \big|l_{\theta,p}^{(\ell)}(t)\big|; t \in I_0, (\theta, p) \in \Theta \times \mathcal{P}\big\} < +\infty,$$

$$\exists\kappa \in [0,1), \exists G > 0, \forall n \geq 1, \quad \sup\big\{\big|r_{\theta,p,n}^{(\ell)}(t)\big|; t \in I_0, (\theta, p) \in \Theta \times \mathcal{P}\big\} \leq G\kappa^n.$$

Furthermore, the functions $\lambda_{\theta,p}^{(m)}(\cdot)$, $l_{\theta,p}^{(m)}(\cdot)$ and $r_{\theta,p,n}^{(m)}(\cdot)$ are continuous on I_0 uniformly in $(\theta, p) \in \Theta \times \mathcal{P}$.

Hypothesis (N-A) (Non-arithmeticity). *For any compact subset K_0 of $\mathbb{R}^*$, there exists $\rho \in [0, 1)$ such that*

$$\forall n \geq 1, \quad \sup\big\{\big|\mathbb{E}_\theta\big[\mathrm{e}^{\mathrm{i}tS_n(p)}\big]\big|; t \in K_0, (\theta, p) \in \Theta \times \mathcal{P}\big\} = \mathrm{O}\big(\rho^n\big).$$

Note that under Hypothesis $\mathcal{R}(2)$, the function $t \mapsto \mathbb{E}_\theta[\mathrm{e}^{\mathrm{i}tS_n(p)}]$ is of class $\mathcal{C}^2$ on I_0 for all $(\theta, p) \in \Theta \times \mathcal{P}$. Then by Fatou lemma, for all $(\theta, p) \in \Theta \times \mathcal{P}$, one has $\mathbb{E}_\theta[S_n(p)^2] < +\infty$. Therefore, when considering the derivative of Eq. (3), one easily obtains that for all $(\theta, p) \in \Theta \times \mathcal{P}$, $\lim \mathbb{E}_\theta[S_n(p)]/n = 0$ when $n \to +\infty$. Note that under Hypothesis $\mathcal{R}(2)$, when considering the second derivative of Eq. (3), one easily obtains as well

$$\forall n \geq 1, \quad \lim_{n\to+\infty} \sup_{(\theta,p)\in\Theta\times\mathcal{P}} \left|\frac{\mathbb{E}_\theta[S_n(p)^2]}{n}\right| < +\infty, \tag{4}$$

and in a similar way, under Hypothesis $\mathcal{R}(4)$,

$$\forall n \geq 1, \quad \lim_{n\to+\infty} \sup_{(\theta,p)\in\Theta\times\mathcal{P}} \left|\frac{\mathbb{E}_\theta[S_n(p)^4]}{n^2}\right| < +\infty. \tag{5}$$

Finally, under Hypothesis $\mathcal{R}(3)$, we obtain some of the assertions of Proposition 1 below. The other ones can be proved by borrowing the proof of [8], Chapter XVI.4, Theorem 4.1.

Proposition 1 (First-order Edgeworth expansion). *If $\{S_n(p); n \in \mathbb{N}^*, p \in \mathcal{P}\}$ satisfies Hypothesis $\mathcal{R}(3)$, then for all $(\theta, p) \in \Theta \times \mathcal{P}$, the following limits*

$$b_{\theta,p} := \lim_{n\to+\infty} \mathbb{E}_\theta\big[S_n(p)\big], \qquad \sigma_{\theta,p}^2 := \lim_{n\to+\infty} \frac{\mathbb{E}_\theta[S_n(p)^2]}{n},$$

are well-defined and bounded in $\theta \in \Theta$. The number $b_{\theta,p}$ is the bias of order $1/n$ of the statistics of interest $S_n(p)/n$. Furthermore if $\inf_{(\theta,p)\in\Theta\times\mathcal{P}} \sigma_{\theta,p} > 0$ and if the family $\{S_n(p); n \in \mathbb{N}^*, p \in \mathcal{P}\}$ satisfies Hypothesis (N-A) *as well, then there exists a polynomial function $A_{\theta,p}$ such that*

$$\sup_{(\theta,p)\in\Theta\times\mathcal{P}} \sup_{u\in\mathbb{R}} \left|\mathbb{P}_\theta\left\{\frac{S_n(p)}{\sigma_{\theta,p}\sqrt{n}} \leq u\right\} - \mathcal{N}(u) - \eta(u)n^{-1/2}A_{\theta,p}(u)\right| = \mathrm{o}\big(n^{-1/2}\big).$$

The polynomial function is of the type $A_{\theta,p}(u) = a_1(\theta, p) + a_2(\theta, p)u^2$ where the coefficients satisfy for $i = 1, 2$, $\sup_{(\theta,p)\in\Theta\times\mathcal{P}} |a_i(\theta, p)| < +\infty$. Furthermore, if $\mathbb{E}_\theta[|S_n(p)|^3] < +\infty$ for all $n \geq 1$ and $(\theta, p) \in \Theta \times \mathcal{P}$, then the limit

$$m_{\theta,p,3}^3 := \lim_{n\to+\infty} \frac{\mathbb{E}_\theta[S_n(p)^3]}{n} - 3\sigma_{\theta,p}^2 b_{\theta,p}$$

is well-defined and bounded in $\theta \in \Theta$, *and moreover* $A_{\theta,p}(u) := \frac{m^3_{\theta,p,3}}{6\sigma^3_{\theta,p}}(1-u^2) - \frac{b_{\theta,p}}{\sigma_{\theta,p}}$.

Remark 1. *In the i.i.d. case, Hypotheses* $\mathcal{R}(3)$ *and* (N-A) *are easily checked. Indeed consider* $(X_n)_{n\in\mathbb{N}^*}$ *a sequence of i.i.d.* E*-valued r.v. whose common distribution depends on* $\theta \in \Theta$, *and* $\{\xi_p(\cdot);\ p \in \mathcal{P}\}$ *a family of measurable functions from* E *into* $\mathbb{R}$. *The following assertions are obviously equivalent*:

(a) *The family* $\{\sum_{k=1}^n \xi_p(X_k);\ n \in \mathbb{N}^*,\ p \in \mathcal{P}\}$ *fulfills Hypothesis* $\mathcal{R}(m)$ *if and only if* $\mathbb{E}_\theta[\xi_p(X_1)] = 0$ *for all* $(\theta, p) \in \Theta \times \mathcal{P}$ *and* $\sup_{(\theta,p)\in\Theta\times\mathcal{P}} \mathbb{E}_\theta[|\xi_p(X_1)|^m] < +\infty$.
(b) *The family* $\{\sum_{k=1}^n \xi_p(X_k);\ n \in \mathbb{N}^*,\ p \in \mathcal{P}\}$ *fulfills Hypothesis* (N-A) *if and only if, for any compact subset* K_0 *of* $\mathbb{R}^*$, *one has*

$$\sup_{t\in K_0} \sup_{(\theta,p)\in\Theta\times\mathcal{P}} \left|\mathbb{E}_\theta\left[e^{it\xi_p(X_1)}\right]\right| < 1. \tag{6}$$

When (6) *is considered at* (θ, p) *fixed, it can be easily relaxed to the usual condition*: $\xi_p(X_1)$ *is non-lattice. By contrast, it is not easy to relax the uniform condition* (6). *Note that this condition is only discussed in* [25] *under the stronger Cramér condition*:

$$\limsup_{t\to+\infty} \sup_{(\theta,p)\in\Theta\times\mathcal{P}} \left|\mathbb{E}_\theta\left[e^{it\xi_p(X_1)}\right]\right| < 1.$$

Hypotheses $\mathcal{R}(m)$ *and* (N-A) *are the tailor-made assumptions to borrow the proof of the first-order Edgeworth expansion in the i.i.d. case,*[2] *and consequently to expand* $\mathbb{P}_\theta\{S_n(p)/(\sigma_{\theta,p}\sqrt{n}) \le u\}$ *with a polynomial function independent on* n. *Notice that, under less restrictive conditions, the results of* [7] *provide a first-order Edgeworth-type expansion but with a polynomial function depending on* n.

2.2. *The main lines of the weak spectral method for Markovian models*

Consider from now on the following general Markovian setting. Let $(E, \mathcal{E})$ be any measurable space, and let $(X_n)_{n\ge 0}$ be a Markov chain with state space E and transition kernel $(Q_\theta(x,\cdot);\ x \in E)$ where θ is a parameter in some set Θ. The initial distribution of the chain is denoted by μ_θ (i.e., $X_0 \sim \mu_\theta$). The underlying probability measure and the associated expectation are denoted by $\mathbb{P}_{\theta,\mu_\theta}$ and $\mathbb{E}_{\theta,\mu_\theta}$. We assume that $(X_n)_{n\in\mathbb{N}}$ admits an invariant probability measure denoted by π_θ (i.e., $\forall\theta \in \Theta$, $\pi_\theta \circ Q_\theta = \pi_\theta$). Notice that we do not require stationarity for $(X_n)_{n\in\mathbb{N}}$.

Let $\{\xi_p(\cdot,\cdot);\ p \in \mathcal{P}\}$ be a family of measurable functions from E^2 into $\mathbb{R}$, where $\mathcal{P}$ is any set. Let us define the following r.v.

$$\forall n \ge 1, \forall p \in \mathcal{P}, \quad S_n(p) := \sum_{k=1}^n \xi_p(X_{k-1}, X_k). \tag{7}$$

This kind of (parametric and bivariate) functionals is required when concerning with Markovian M-estimators, as detailed in Section 5.

Now we are going to study Hypotheses $\mathcal{R}(m)$ and (N-A) using the Nagaev–Guivarc'h spectral method. For all $t \in \mathbb{R}$, $(\theta, p) \in \Theta \times \mathcal{P}$ and $x \in E$, let us define the Fourier kernel of (Q_θ, ξ_p) by

$$Q_{\theta,p}(t)(x, dy) := e^{it\xi_p(x,y)} Q_\theta(x, dy). \tag{8}$$

As usual, for all bounded measurable $\mathbb{C}$-valued function f on E, we set

$$Q_{\theta,p}(t)f := \int_E f(y)e^{it\xi_p(\cdot,y)} Q_\theta(\cdot, dy).$$

[2] One difference is that $b_{\theta,p}$ is null in the i.i.d. case.

It is easy to see that we have from Markov property

$$\forall t \in \mathbb{R}, \forall(\theta, p) \in \Theta \times \mathcal{P}, \forall n \geq 1, \quad \mathbb{E}_{\theta,\mu_\theta}\big[e^{itS_n(p)} f(X_n)\big] = \mu_\theta\big[Q_{\theta,p}(t)^n f\big].$$

In particular, we obtain

$$\forall t \in \mathbb{R}, \forall(\theta, p) \in \Theta \times \mathcal{P}, \forall n \geq 1, \quad \mathbb{E}_{\theta,\mu_\theta}\big[e^{itS_n(p)}\big] = \mu_\theta\big[Q_{\theta,p}(t)^n \mathbf{1}_E\big], \tag{9}$$

where $\mathbf{1}_E$ stands for the function identically equal to 1 on E.

Equality (9) links the characteristic function of $S_n(p)$ to the iterated Fourier operator $Q_{\theta,p}(t)^n$. Thus, according to Eq. (9), Hypothesis $\mathcal{R}(m)$ requires the study of the behavior of $t \mapsto Q_{\theta,p}(t)^n$ near 0. A natural assumption to do it is to assume that there exists a Banach space $\mathcal{B}$ which contains the function $\mathbf{1}_E$ and on which $(Q_\theta)_{\theta\in\Theta}$ acts continuously (i.e., $\forall\theta \in \Theta$, $Q_\theta \in \mathcal{L}(\mathcal{B})$) and $(Q_\theta)_{\theta\in\Theta}$ satisfies the following uniform strong ergodicity properties (ERG.1) and (ERG.2):

(ERG.1) $\{\pi_\theta; \theta \in \Theta\}$ is bounded in $\mathcal{B}'$.

(ERG.2) The transition kernel $(Q_\theta)_{\theta\in\Theta}$ has a spectral gap on $\mathcal{B}$ (uniformly in θ), that is

$$\lim_{n\to+\infty} \sup_{\theta\in\Theta} \big\| Q_\theta^n - \Pi_\theta \big\|_{\mathcal{B}} = 0,$$

where Π_θ denotes the rank-one projection defined on $\mathcal{B}$ by $\Pi_\theta f := \pi_\theta(f)\mathbf{1}_E$.

More precisely, we use the following equivalent form (ERG.2)$'$ of (ERG.2):

(ERG.2)$'$ There exist $c_0 > 0$ and $0 \leq \kappa_0 < 1$ (independent on $\theta \in \Theta$) such that

$$\forall\theta \in \Theta, \forall n \in \mathbb{N}, \quad \big\| Q_\theta^n - \Pi_\theta \big\|_{\mathcal{B}} \leq c_0\kappa_0^n.$$

Note that under (ERG.2)$'$, for all $\theta \in \Theta$, the spectrum $\sigma(Q_{\theta|\mathcal{B}})$ of Q_θ acting on $\mathcal{B}$ belongs to the set $\{z \in \mathbb{C}; |z| \leq \kappa_0\} \cup \{1\}$.

Then, to derive the properties of $\mathcal{R}(m)$ from (9), we need some spectral perturbation method to control (uniformly in $(\theta, p) \in \Theta \times \mathcal{P}$) the spectrum of $Q_{\theta,p}(t)$ acting on $\mathcal{B}$ whenever $|t|$ is small enough. The usual method requires the continuity at $t = 0$ of the $\mathcal{L}(\mathcal{B})$-valued function $t \mapsto Q_{\theta,p}(t)$, but this continuity assumption involves too strong hypotheses (see [18], Section 3, for details). An alternative method consists in using the Keller–Liverani theorem [20,22] (see also [1,9]). Using this method, the regularity of $\lambda_{\theta,p}(\cdot)$, $l_{\theta,p}(\cdot)$ and $r_{\theta,p,n}(\cdot)$ is studied in [18] in the case of ρ-mixing Markov chains, V-geometrically ergodic Markov chains and for iterated function systems. More exactly, their results are only established for additive univariate functionals of $(X_n)_{n\in\mathbb{N}^*}$, but the extension to our parametric bivariate case (7) is quite natural. This work has already been done in [17] in the case of V-geometrically ergodic Markov chains (in Section 4, we will directly use their results).

By contrast, as already mentioned in Introduction, the method developed in [18] is not sufficient to investigate Hypothesis (N-A) in our parametric bivariate case. We can all the same easily state the following implication: thanks to (9), provided that the following condition is imposed on $(\mu_\theta)_{\theta\in\Theta}$:

$$\{\mu_\theta; \theta \in \Theta\} \text{ is bounded in } \mathcal{B}' \tag{10}$$

and that for all $t \in \mathbb{R}$, $(\theta, p) \in \Theta \times \mathcal{P}$, the operator $Q_{\theta,p}(t)$ belongs to $\mathcal{L}(\mathcal{B})$, the family $\{S_n(p) := \sum_{k=1}^n \xi_p(X_{k-1}, X_k); n \in \mathbb{N}^*, p \in \mathcal{P}\}$ fulfills Hypothesis (N-A) whenever the following condition holds:

Hypothesis (N-A)$'$ (Operator-type non-arithmeticity). *For any compact subset $K_0 \subset \mathbb{R}^*$, there exists $\rho < 1$ such that*

$$\forall n \geq 1, \quad \sup\big\{\big\| Q_{\theta,p}(t)^n \big\|_{\mathcal{B}}; t \in K_0, (\theta, p) \in \Theta \times \mathcal{P}\big\} = \mathrm{O}\big(\rho^n\big).$$

In the next section, we replace this condition by some more practical non-lattice conditions.

3. From non-lattice conditions to (N-A)′

We assume that the general Markovian assumptions of the previous Section 2.2 hold true. Furthermore, we also assume that $\mathcal{B}$ is a Banach space of complex measurable functions defined on E which contains the function $\mathbf{1}_E$, such that for all $\theta \in \Theta$, $\pi_\theta \in \mathcal{B}'$ and such that for all $t \in \mathbb{R}$, $(\theta, p) \in \Theta \times \mathcal{P}$, the Fourier operators $Q_{\theta,p}(t)$ defined in (8) belong to $\mathcal{L}(\mathcal{B})$.

Let us introduce the following non-lattice condition which will be proved (under some additional conditions) to imply the previous operator-type non-arithmetic condition (N-A)′.

Hypothesis (N-L) (Non-lattice). *There exist no $(\theta_0, p_0) \in \Theta \times \mathcal{P}$, no real $a = a(\theta_0, p_0)$, no closed subgroup $H = c\mathbb{Z}$ with $c = c(\theta_0, p_0) \in \mathbb{R}^*$, no π_{θ_0}-full Q_{θ_0}-absorbing set*[3] *$A = A(\theta_0, p_0) \in \mathcal{E}$, and finally no measurable bounded function $\alpha = \alpha(\theta_0, p_0) : E \to \mathbb{R}$ such that*

$$\forall x \in A, \quad \xi_{p_0}(x, y) + \alpha(y) - \alpha(x) \in a + H, \quad Q_{\theta_0}(x, \mathrm{d}y)\text{-a.s.} \tag{11}$$

3.1. *Intermediate conditions*

The link between (N-L) and (N-A)′ is based on the three following operator-type properties. The first one concerns a control of the spectral radius of $Q_{\theta,p}(t)$ acting on $\mathcal{B}$ denoted by $r(Q_{\theta,p}(t)_{|\mathcal{B}})$:

$$\forall t \neq 0, \forall(\theta, p) \in \Theta \times \mathcal{P}, \quad r\big(Q_{\theta,p}(t)_{|\mathcal{B}}\big) < 1. \tag{i}$$

The second property consists in assuming that one has for any compact subset $K_0 \subset \mathbb{R}^*$

$$r_{K_0} := \sup\big\{r\big(Q_{\theta,p}(t)_{|\mathcal{B}}\big); t \in K_0, (\theta, p) \in \Theta \times \mathcal{P}\big\} < 1. \tag{ii}$$

Notice that, whenever (ii) holds true, for all $z \in \mathbb{C}$, $|z| > r_{K_0}$ and for all $t \in K_0$, for all $(\theta, p) \in \Theta \times \mathcal{P}$, the resolvent operator $(z - Q_{\theta,p}(t))^{-1}$ is well-defined in $\mathcal{L}(\mathcal{B})$. Then the last property consists in assuming that there exists $\rho_0 \in [r_{K_0}, 1)$ such that, for all $\rho \in (\rho_0, 1)$,

$$\sup\big\{\big\|\big(z - Q_{\theta,p}(t)\big)^{-1}\big\|_{\mathcal{B}}; t \in K_0, (\theta, p) \in \Theta \times \mathcal{P}, |z| = \rho\big\} < +\infty. \tag{iii}$$

Below we study the following implications:

(a) (N-L) ⇒ (i) under some conditions (and even better: (N-L) ⇔ (i) under some additional conditions)
(b) (i) ⇒ (ii) and (iii) under some conditions;
(c) (ii) and (iii) ⇒ (N-A)′.

The main difficulty is the proof of the statement (b). For this part, three methods are proposed in Section 3.3. Notice that the operator-type non-arithmetic condition (N-A)′ obviously implies property (i).

3.2. *From the non-lattice condition* (N-L) *to property* (i)

The following lemma is an easy extension of [18], Lemma 12.1, to our parametric bivariate case.

Lemma 1. *Assume that the following assumptions hold true*:

1. *For all $\theta \in \Theta$, $\lambda \in \mathbb{C}$ such that $|\lambda| \geq 1$, and for all $f \in \mathcal{B}$, $f \neq 0$, we have*

$$\big[\forall n \geq 1, |\lambda|^n |f| \leq Q_\theta^n |f|\big] \quad \Rightarrow \quad \big[|\lambda| = 1 \text{ and } |f| = \pi_\theta\big(|f|\big) > 0 \ \pi_\theta\text{-a.s.}\big].$$

2. *For all $(\theta, p) \in \Theta \times \mathcal{P}$, $t \in \mathbb{R}^*$, there exists $0 \leq \gamma = \gamma(\theta, p, t) < 1$ such that the elements of the spectrum of $Q_{\theta,p}(t)$ acting on $\mathcal{B}$ with modulus greater than γ are isolated eigenvalues of finite multiplicity.*

[3] A set $A \in \mathcal{E}$ is said to be π_{θ_0}-full if $\pi_{\theta_0}(A) = 1$ and Q_{θ_0}-absorbing if $Q_{\theta_0}(a, A) = 1$ for all $a \in A$.

If (N-L) *holds true as well, then* (i) *is fulfilled. Moreover property* (i) *is equivalent to the following condition: there exist no* $t_0 \in \mathbb{R}^*$, *no* $(\theta_0, p_0) \in \Theta \times \mathcal{P}$, *no* $\lambda = \lambda(\theta_0, p_0, t_0) \in \mathbb{C}$ *such that* $|\lambda| = 1$, *no* π_{θ_0}*-full* Q_{θ_0}*-absorbing set* $A = A(\theta_0, p_0, t_0) \in \mathcal{E}$ *and finally no bounded* $w = w(\theta_0, p_0, t_0) \in \mathcal{B}$ *such that* $|w|_{|A}$ *is non-null constant, satisfying*

$$\forall x \in A, \quad e^{it_0 \xi_{p_0}(x,y)} w(y) = \lambda w(x), \quad Q_{\theta_0}(x, dy)\text{-}a.s.$$

The last property of Lemma 1 will not be used later, it is only recalled here for a better understanding.

Remark 2. *In fact, property* (i) *is equivalent to* (N-L) *whenever* $e^{i\psi} \in \mathcal{B}$ *for all bounded real measurable function* ψ *on* E. *Notice that this assumption is obviously fulfilled in the* V*-geometrically ergodic Markovian model to be studied.*

Assumption 1 of Lemma 1 is always satisfied for strongly ergodic models (cf. (ERG.2)) such that $\mathcal{B}$ is stable under complex moduli and such that for all $x \in E$, the Dirac distribution δ_x at x belongs to $\mathcal{B}'$. In particular, this assumption is satisfied by the V-geometrically ergodic Markovian model to be studied (other conditions are given in [18] to check assumption 1).

Assumption 2 is much more difficult to be checked. For now, we only mention that it is equivalent to the following condition: for all $(\theta, p) \in \Theta \times \mathcal{P}$, $t \in \mathbb{R}^*$, the essential spectral radius $r_{ess}(Q_{\theta,p}(t)_{|\mathcal{B}})$ of $Q_{\theta,p}(t)$ acting on $\mathcal{B}$ is such that $r_{ess}(Q_{\theta,p}(t)_{|\mathcal{B}}) \le \gamma < 1$. Recall that $Q_{\theta,p}(t)$ is said to be quasi-compact on $\mathcal{B}$ whenever $r_{ess}(Q_{\theta,p}(t)_{|\mathcal{B}}) < r(Q_{\theta,p}(t)_{|\mathcal{B}})$.

3.3. *Three methods for condition* (i) *to imply* (ii) *and* (iii)

To obtain the implication (i) $\Rightarrow$ (ii) and (iii), we can use one of the following three approaches, in which the sets Θ and $\mathcal{P}$ are assumed to be compact.

- *First approach*. Using the standard operator perturbation theory, specifically the upper-semi-continuity of the function “spectral radius” (see, e.g., [16], p. 19), one can prove the following statement:
  Assume that $\|Q_{\theta,p}(t) - Q_{\theta_0,p_0}(t_0)\|_{\mathcal{B}} \to 0$ *when* $(t, \theta, p) \to (t_0, \theta_0, p_0)$. *Then the implications* (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) *are true*.

  However, as already mentioned in Section 2.2, the last assumption of continuity of $t \mapsto Q_{\theta,p}(t)$ is too restrictive. That is why we will not apply this approach in this work.

- *Second approach*. It consists in using the perturbation Keller–Liverani theorem instead of the standard perturbation theory. The proof of the following proposition is not provided in this paper since it is an easy extension of [18], Lemma 12.3.

Proposition 2. *Assume that there exists some semi normed space* $\widetilde{\mathcal{B}}$ *such that for all* $t \in \mathbb{R}$ *and* $(\theta, p) \in \Theta \times \mathcal{P}$, $Q_{\theta,p}(t)$ *belongs to* $\mathcal{L}(\widetilde{\mathcal{B}})$ *and* $\mathcal{B} \hookrightarrow \widetilde{\mathcal{B}}$ (*i.e.*, $\mathcal{B} \subset \widetilde{\mathcal{B}}$ *and the identity map is continuous from* $\mathcal{B}$ *into* $\widetilde{\mathcal{B}}$). *Furthermore assume that for all* $t_0 \in \mathbb{R}^*$ *and* $(\theta_0, p_0) \in \Theta \times \mathcal{P}$, *there exists a neighborhood* $\widetilde{I_0}$ *of* (t_0, θ_0, p_0) *such that*

(C1) *there exist* $c > 0$, $0 \le \kappa < 1$ *and* $M > 0$ *such that for all* $(t, \theta, p) \in \widetilde{I_0}$, $f \in \mathcal{B}$, $n \in \mathbb{N}$, *one has* $\|Q_{\theta,p}(t)^n f\|_{\mathcal{B}} \le c\kappa^n \|f\|_{\mathcal{B}} + cM^n \|f\|_{\widetilde{\mathcal{B}}}$.
(C2) $\|Q_{\theta,p}(t) - Q_{\theta_0,p_0}(t_0)\|_{\mathcal{B},\widetilde{\mathcal{B}}} \to 0$ *when* $(t, \theta, p) \to (t_0, \theta_0, p_0)$.

Then the implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) *are true*.

This second approach is applied in Section 6.4 to some AR(d) processes with $d \ge 2$.

Note that condition (C2) may be difficult to be checked because of the continuity with respect to θ. However, under the standard dominated model assumption, condition (C2) can be dodged using the following approach.

- *Third approach*. It consists in using quasi-compactness and Ascoli-type arguments. For instance, let us give a brief account for the implication (i) $\Rightarrow$ (ii). We assume by absurd that (ii) does not hold and (i) holds true, namely on the one hand there exists a compact subset K_0 of $\mathbb{R}^*$ such that $r_{K_0} := \sup\{r(Q_{\theta,p}(t)_{|\mathcal{B}}); t \in K_0, (\theta, p) \in \Theta \times \mathcal{P}\} \ge 1$, and on the other hand $r_{K_0} \le 1 < +\infty$. Then there exist some sequences $(t_k)_{k\in\mathbb{N}} \in K_0^{\mathbb{N}}$ and $(\theta_k, p_k)_{k\in\mathbb{N}} \in (\Theta \times \mathcal{P})^{\mathbb{N}}$

such that $\lim r(Q_{\theta_k,p_k}(t_k)) \geq 1$ when $k \to +\infty$. Under the quasi-compactness assumption 2 of Lemma 1, the previous property implies the existence of $(\lambda_k)_{k\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ and $(w_k)_{k\in\mathbb{N}} \in \mathcal{B}^{\mathbb{N}}$ such that $Q_{\theta_k,p_k}(t_k)w_k = z_k w_k$ and $|\lambda_k| = r(Q_{\theta_k,p_k}(t_k))$. Finally, from compactness arguments (in particular by using Ascoli theorem), there exist some $\tilde{t} \in K_0$, $(\tilde{\theta}, \tilde{p}) \in \Theta \times \mathcal{P}$, $\tilde{\lambda} \in \mathbb{C}$, $|\tilde{\lambda}| \geq 1$, and $\tilde{w} \in \mathcal{B}$ such that $Q_{\tilde{\theta},\tilde{p}}(\tilde{t})\tilde{w} = \tilde{\lambda}\tilde{w}$, which is in contradiction with (i). Similar arguments can be used to prove (ii) ⇒ (iii). In practice, it is easier to use Ascoli theorem when the model is dominated (i.e., $\forall x \in E$, $\forall \theta \in \Theta$, $Q_\theta(x, \mathrm{d}y) = q_\theta(x, y)\mu(\mathrm{d}y)$) with suitable conditions on the function $(q_\theta)_{\theta\in\Theta}$ and on the dominating positive measure μ.

This approach is detailed in Section 4.2 for V-geometrically ergodic Markov chains.

3.4. *From properties* (ii) *and* (iii) *to the operator-type non-arithmetic condition* (N-A)′

Lemma 2. *Assume that properties* (ii) *and* (iii) *hold true. Then* (N-A)′ *is fulfilled.*

Proof. Let $K_0 \subset \mathbb{R}^*$ be any compact set and let Γ denote the oriented circle defined by $\{z \in \mathbb{C}; |z| = \rho\}$ where $\rho \in (\rho_0, 1)$. From the Von Neumann series, we have for all $t \in K_0$ and $(\theta, p) \in \Theta \times \mathcal{P}$,

$$z \in \mathbb{C},\ |z| = \rho \quad \Rightarrow \quad \big(z - Q_{\theta,p}(t)\big)^{-1} = \sum_{n=0}^{+\infty} z^{-n-1} Q_{\theta,p}(t)^n$$

and hence, we obtain

$$\forall t \in K_0, \forall(\theta, p) \in \Theta \times \mathcal{P}, \forall n \geq 1, \quad Q_{\theta,p}(t)^n = \frac{1}{2\mathrm{i}\pi}\int_\Gamma z^n \big(z - Q_{\theta,p}(t)\big)^{-1}\,\mathrm{d}z.$$

Then (N-A)′ can easily be derived thanks to (iii). □

3.5. *Conclusion*

Sections 3.2, 3.3 and 3.4 give a procedure to derive (N-A)′ (and so (N-A)) from the non-lattice condition (N-L). In some cases, we may need some even more simple condition than (N-L) to check (N-A). However, notice that this new condition, denoted by (N-L)′, is not equivalent to (N-L).

Assume that E is a topological connex set and let $\mathcal{E} := \mathcal{B}(E)$ be the associated Borel algebra.

Hypothesis (N-L)′. *For all $p \in \mathcal{P}$, there exist neither continuous function $\mathcal{A}_p(\cdot): E \to \mathbb{R}$ nor constant C_p such that we have for all $(x, y) \in E^2$, $\xi_p(x, y) = \mathcal{A}_p(y) - \mathcal{A}_p(x) + C_p$.*

To connect (N-L)′ with (N-L), we need the following hypotheses on both the model and $(\xi_p)_{p\in\mathcal{P}}$.

Hypothesis ($\mathcal{S}$). *There exists positive measures $\{\mu_\theta; \theta \in \Theta\}$ on E satisfying* $\mathrm{Supp}(\mu_\theta) = E$ *for all $\theta \in \Theta$ and such that we have for any $B \in \mathcal{E}$ and $\theta \in \Theta$:*

$$\big[\exists x \in E, Q_\theta(x, B) = 0\big] \quad \Longrightarrow \quad \big[\mu_\theta(B) = 0\big].$$

Hypothesis ($\mathcal{C}$). *For all $p \in \mathcal{P}$, the application ξ_p is continuous from E^2 into $\mathbb{R}$.*

Lemma 3. *Assume that Hypothesis* ($\mathcal{S}$) *and* ($\mathcal{C}$) *hold true. If the family $(\xi_p)_{p\in\mathcal{P}}$ fulfills* (N-L)′, *then* (N-L) *is fulfilled.*

Proof. Assume that (N-L) is not fulfilled, that is we have (11) with some $(\theta_0, p_0) \in \Theta \times \mathcal{P}$, $a \in \mathbb{R}$, some closed subgroup $H = c\mathbb{Z}$ ($c \in \mathbb{R}^*$), some π_{θ_0}-full Q_{θ_0}-absorbing set $A \in \mathcal{E}$, and finally some bounded measurable function $\alpha: E \to \mathbb{R}$. For the sake of simplicity, let us omit the dependence on (θ_0, p_0). For all $x \in A$, there exists $E_x \in \mathcal{E}$ such that $Q(x, E_x) = 1$ and $\xi(x, y) + \alpha(y) - \alpha(x) \in a + H$ for all $y \in E_x$. Let $x_0 \in A$. One has

$$\forall y \in E_{x_0}, \quad \xi(x_0, y) + \alpha(y) - \alpha(x_0) \in a + H.$$

Thanks to Hypothesis $(\mathcal{S})$, one has $\mu(E\backslash E_{x_0}) = 0$ and $\mu(E\backslash A) = 0$ (recall that A is Q-absorbing), and hence $\mu(E\backslash\{A \cap E_{x_0}\}) = 0$, $\overline{A \cap E_{x_0}} \supset \mathrm{Supp}(\mu) = E$ where $\overline{A \cap E_{x_0}}$ denotes the closure of $A \cap E_{x_0}$. In particular, $A \cap E_{x_0}$ is not empty. Let $x \in A \cap E_{x_0}$, then

$$\forall y \in E_{x_0} \cap E_x, \quad \xi(x, y) - \big(\xi(x_0, y) - \xi(x_0, x)\big) \in a + H.$$

Let us define $\mathcal{A}(x) := \xi(x_0, x)$ and $f(x, y) := \xi(x, y) + \mathcal{A}(x) - \mathcal{A}(y)$. Then $f(x, E_{x_0} \cap E_x) \subset a + H$ for all $x \in A \cap E_{x_0}$. Then, by continuity arguments and since $E = \mathrm{Supp}(\mu) = \overline{E_{x_0} \cap E_x}$, one can easily show that $f(x, E) \subset a + H$. In the same way, $f(A \cap E_{x_0}, E) \subset a + H$, and finally $f(E, E) \subset a + H$. Since $f(E, E)$ is connex and $a + H$ is discrete, f is constant on E^2. □

Remark 3. *Let μ be a positive measure on E satisfying* $\mathrm{Supp}(\mu) = E$. *Assume that the following dominated model condition holds*: *for all $\theta \in \Theta$, there exists a non-negative measurable application $q_\theta(\cdot,\cdot)$ on $(E \times E, \mathcal{E} \otimes \mathcal{E})$ such that for all $x \in E$, $B \in \mathcal{E}$, $Q_\theta(x, B) = \int_B q_\theta(x, y)\,\mathrm{d}\mu(y)$ and for all $x \in E$ and for μ-almost all $y \in E$, $q_\theta(x, y) > 0$. Then one can show that Hypothesis $(\mathcal{S})$ holds true.*

Remark 4. *In the i.i.d. case, Hypothesis* (N-L)$'$ *is never checked under Hypothesis* $(\mathcal{C})$. *However, we can state another assumption adapted to the i.i.d. case. Indeed consider $(X_n)_{n\in\mathbb{N}^*}$ a sequence of i.i.d. E-valued r.v. whose common distribution π_θ depends on $\theta \in \Theta$, and $\{\xi_p(\cdot); p \in \mathcal{P}\}$ a family of measurable functions from E into $\mathbb{R}$. Let us first recall* (N-L):

There exist no $(\theta_0, p_0) \in \Theta \times \mathcal{P}$, no real $a = a(\theta_0, p_0)$, no closed subgroup $H = c\mathbb{Z}$ with $c = c(\theta_0, p_0) \in \mathbb{R}^*$, no π_{θ_0}-full set $A = A(\theta_0, p_0) \in \mathcal{E}$, and finally no measurable bounded function $\alpha = \alpha(\theta_0, p_0) : E \to \mathbb{R}$ such that for all $x \in A$, $\xi_{p_0}(y) + \alpha(y) - \alpha(x) \in a + H\ \pi_{\theta_0}(\mathrm{d}y)$-a.s.

Then the following result is obvious using similar arguments as in Lemma 3. Assume that E is a topological set and that the family $(\xi_p)_{p\in\mathcal{P}}$ fulfills the following non-lattice condition:

For all $p \in \mathcal{P}$ and $\theta \in \Theta$, there exist neither closed subgroup $H = c\mathbb{Z}$ with $c = c(\theta, p) \in \mathbb{R}^*$ nor real $a = a(\theta, p)$ such that we have $\xi_p(x) \in a + H\ \pi_\theta(\mathrm{d}x)$-a.s.

Then (N-L) *is fulfilled.*

4. The case of uniform V-geometrically ergodic Markov chains

In this section, we illustrate the previous results for uniform V-geometrically ergodic Markov chains. From now on, for the sake of simplicity, we consider that $E := \mathbb{R}^d$ (with $d \in \mathbb{N}^*$), equipped with any norm $\|\cdot\|$, and μ_d^{Leb} denotes the Lebesgue-measure on E. Let us assume that Θ is a compact set. We introduce the uniform (in $\theta \in \Theta$) V-geometrically ergodic Markovian model:

Model $(\boldsymbol{\mathcal{M}})$. *For all $\theta \in \Theta$, there exist both a Q_θ-invariant probability measure denoted by π_θ and an unbounded function $V : E \to [1, +\infty)$ such that*

(VG1) $\sup_{\theta\in\Theta} \pi_\theta(V) < +\infty$,
(VG2) $\lim_{n\to+\infty} \sup\{|Q_\theta^n f(x) - \pi_\theta(f)|/V(x);\ f : E \to \mathbb{C}$ *measurable*, $|f| \le V, x \in E, \theta \in \Theta\} = 0$.

Model $(\mathcal{M})$ satisfies properties (ERG.1) and (ERG.2) on the weighted-supremum normed space associated with V. This space, denoted by $\mathcal{B}_1$, is the Banach space composed of measurable functions $f : E \to \mathbb{C}$ such that

$$\|f\|_{\mathcal{B}_1} := \sup_{x\in E} \frac{|f(x)|}{V(x)} < +\infty. \tag{12}$$

Model $(\mathcal{M})$ has already been considered for statistical investigation, see, for instance, [6,13,17]. When θ is fixed and when the Markov chain is irreducible and aperiodic, (VG1) and (VG2) can be checked using the so-called drift-criterion, we refer to [24], Chapter 15.2.2 (V4), for details. Notice that condition (10) on the initial distribution for $\mathcal{B} = \mathcal{B}_1$ is equivalent to the following one:

$$\sup_{\theta\in\Theta} \mu_\theta(V) < +\infty. \tag{13}$$

In the next Sections 4.1 and 4.2, we consider a family $(\xi_p)_{p\in\mathcal{P}}$ of measurable functions from E^2 into $\mathbb{R}$, with $\mathcal{P}$ assumed to be a compact set, and we successively study Hypotheses $\mathcal{R}(m)$ and (N-A) for Model $(\mathcal{M})$, before applying these results in Section 5 to M-estimators.

4.1. *Study of Hypothesis* $\mathcal{R}(m)$

Let us recall the following proposition which has already been proven in [17], Lemma 1.

Proposition 3. *Let us consider a Model* $(\mathcal{M})$. *Assume that, for all* $p\in\mathcal{P}$, ξ_p *is centered with respect to the invariant measure family* $(\pi_\theta)_{\theta\in\Theta}$,[4] *and that assume that* $(\xi_p)_{p\in\mathcal{P}}$ *fulfills the following moment domination condition for some* $m\in\mathbb{N}$:

$$\exists\varepsilon>0,\quad \sup\left\{\frac{|\xi_p(x,y)|^{m+\varepsilon}}{V(x)+V(y)};(x,y)\in E^2,p\in\mathcal{P}\right\}<+\infty. \tag{D_m}$$

Finally assume that the initial distribution family $(\mu_\theta)_{\theta\in\Theta}$ *satisfies* (13). *Then the family* $\{S_n(p):=\sum_{k=1}^n\xi_p(X_{k-1},X_k);n\in\mathbb{N}^*,p\in\mathcal{P}\}$ *satisfies Hypothesis* $\mathcal{R}(m)$.

Up to the arbitrarily small real number $\varepsilon>0$, condition (D_m) is the expected (with respect to the i.i.d. case) assumption to obtain Hypothesis $\mathcal{R}(m)$ in our model. Indeed, in [6], condition (D_2) is the key assumption to prove the asymptotic normality whereas in [17], condition (D_3) is the key assumption to prove Berry–Esseen bounds. Here one also needs to investigate Hypothesis (N-A).

4.2. *Study of Hypothesis* (N-A) *for dominated Models* $(\mathcal{M})$

Further assumptions are required to apply what we called the third approach in Section 3.3. Some of them concern the dominated model and the other ones involve the regularity of the applications $(\xi_p)_{p\in\mathcal{P}}$.

Assumption $(\mathcal{S}')$. *For all* $\theta\in\Theta$, *there exists a measurable application* $q_\theta(\cdot,\cdot)$ *on* E^2 *such that*

$$\forall x\in E,\quad Q_\theta(x,\mathrm{d}y)=q_\theta(x,y)\mu_d^{\mathrm{Leb}}(\mathrm{d}y).$$

Furthermore for all $x\in E$ *and for* μ_d^{Leb}*-almost all* $y\in E$, *the application* $\theta\mapsto q_\theta(x,y)$ *is continuous and there exists* $\beta>0$ *such that*

- *for all* $\theta_0\in\Theta$, *there exists a neighborhood* $\mathcal{V}_1=\mathcal{V}_1(\theta_0)$ *of* θ_0 *such that*

$$\forall x_0\in E,\quad \lim_{x\to x_0}\sup_{\theta\in\mathcal{V}_1}\int_E V(y)^\beta\left|q_\theta(x,y)-q_\theta(x_0,y)\right|\mu_d^{\mathrm{Leb}}(\mathrm{d}y)=0.$$

- *for all* $x_0\in E$ *and* $\theta_0\in\Theta$, *there exists a neighborhood* $\mathcal{V}_2=\mathcal{V}_2(x_0,\theta_0)$ *of* θ_0 *such that*

$$\int_E V(y)^\beta\sup_{\theta\in\mathcal{V}_2}\left|q_\theta(x_0,y)\right|\mu_d^{\mathrm{Leb}}(\mathrm{d}y)<+\infty.$$

Assumption $(\mathcal{C}')$. *The family* $(\xi_p)_{p\in\mathcal{P}}$ *satisfies*

- *for all* $x\in E$ *and for* μ_d^{Leb}*-almost all* $y\in E$, *the function* $p\mapsto\xi_p(x,y)$ *is continuous.*
- *for all* $x_0\in E$ *and* $p_0\in\mathcal{P}$, *there exist neighborhoods* $\mathcal{V}_3=\mathcal{V}_3(x_0,p_0)$ *of* x_0 *and* $\mathcal{V}_4=\mathcal{V}_4(p_0)$ *of* p_0, *some positive numbers* C, υ_1 *and* υ_2 *such that we have*

$$\forall p\in\mathcal{V}_4,\forall x\in\mathcal{V}_3,\forall y\in E,\quad \left|\xi_p(x,y)-\xi_p(x_0,y)\right|\le C\|x-x_0\|^{\upsilon_1}V(y)^{\upsilon_2}.$$

[4] I.e., for all $(\theta,p)\in\Theta\times\mathcal{P}$, $\mathbb{E}_{\theta,\pi_\theta}[\xi_p(X_0,X_1)]=0$.

Theorem 1. *Let us consider a Model* $(\mathcal{M})$, *and assume that the preceding Assumptions* $(\mathcal{S}')$ *and* $(\mathcal{C}')$ *hold true. If the non-lattice condition* (N-L) *of Section 3 holds true and if the family of initial distributions* $(\mu_\theta)_{\theta\in\Theta}$ *satisfies* (13), *then* $\{S_n(p) := \sum_{k=1}^{n} \xi_p(X_{k-1}, X_k); n \in \mathbb{N}^*, p \in \mathcal{P}\}$ *satisfies Hypothesis* (N-A).

In the introduction of this section, we defined the natural Banach space $\mathcal{B}_1$ composed of measurable functions $f : E \to \mathbb{C}$ such that (12) holds true. Actually, for a technical reason arising in Lemma 5 below, we need to work with another space. Let β be given in Assumption $(\mathcal{S}')$. Without loss of generality, one can suppose that $\beta \in (0, 1)$. Then we consider the Banach space $\mathcal{B}_\beta$ composed of measurable functions $f : E \to \mathbb{C}$ such that

$$\|f\|_{\mathcal{B}_\beta} := \sup_{x\in E} \frac{|f(x)|}{V(x)^\beta} < +\infty. \tag{14}$$

Notice that for any Model $(\mathcal{M})$, using the drift-criterion (cf. [24], Chapter 15.2.2 (V4)) and Jensen inequality, we can prove that (see [18], Section 10)

$$\lim_{n\to+\infty} \sup_{\theta\in\Theta} \left\| Q_\theta^n - \Pi_\theta \right\|_{\mathcal{B}_\beta} = 0. \tag{15}$$

Then, assumption (ERG.2) of Section 2.2 holds true with $\mathcal{B} := \mathcal{B}_\beta$. More precisely, we will use the equivalent form (ERG.2)′ of (ERG.2): there exist $\widetilde{c}_\beta > 0$ and $0 \le \kappa_\beta < 1$ (independent on $\theta \in \Theta$) such that

$$\forall\theta \in \Theta, \forall n \in \mathbb{N}, \quad \left\| Q_\theta^n - \Pi_\theta \right\|_{\mathcal{B}_\beta} \le \widetilde{c}_\beta \kappa_\beta^n. \tag{16}$$

Proof of Theorem 1. Let $\hbar$ denote $\hbar := (t, \theta, p) \in \mathbb{R} \times \Theta \times \mathcal{P}$ and $Q(\hbar) := Q_{\theta,p}(t)$. First of all, notice that, since $\mathcal{B}_\beta$ is a Banach lattice (i.e., for all $(f, g) \in \mathcal{B}_\beta \times \mathcal{B}_\beta$, $|f| \le |g| \Rightarrow \|f\|_{\mathcal{B}_\beta} \le \|g\|_{\mathcal{B}_\beta}$) and using (15), we can apply [27], Corollary 1.6, to prove that the essential spectral radius of $Q(\hbar)$ satisfies

$$\exists 0 \le \kappa < 1 \text{ such that } \forall \hbar_0 \in \mathbb{R}^* \times \Theta \times \mathcal{P}, \quad r_{\mathrm{ess}}\big(Q(\hbar_0)_{|\mathcal{B}_\beta}\big) \le \kappa. \tag{17}$$

Next, let us sum up the gap from (N-L) to (N-A), specifying their link with all the intermediate conditions introduced in Section 3.1:

- thanks to the previous Eq. (17) on the essential spectral radius of $Q(\hbar)$, assumption 2 of Lemma 1 holds true (assumption 1 of Lemma 1 also holds true: see the comments after Lemma 1). Thus the conclusions of Lemma 1 are satisfied: (N-L) $\Rightarrow$ (i);
- thanks to Lemma 2: (ii) and (iii) $\Rightarrow$ (N-A)′ with $\mathcal{B} := \mathcal{B}_\beta$;
- from condition (13): (N-A)′ $\Rightarrow$ (N-A).

Next, it only remains to prove that (i) $\Rightarrow$ (ii) and (iii). In fact, we show that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii), using quasi-compactness and Ascoli-type arguments, as announced in the third approach of Section 3.3. The proof of (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) involves the two following Lemmas 4 and 5.

Lemma 4 (Doeblin–Fortet inequality). *For any Model* $(\mathcal{M})$ *and every* $\beta \in (0, 1)$, *there exist* $c_\beta > 0$, $0 \le \kappa_\beta < 1$, *such that*

$$\forall\hbar \in \mathbb{R} \times \Theta \times \mathcal{P}, \forall f \in \mathcal{B}_\beta, \forall n \in \mathbb{N}, \quad \left\| Q(\hbar)^n f \right\|_{\mathcal{B}_\beta} \le c_\beta \kappa_\beta^n \|f\|_{\mathcal{B}_\beta} + c_\beta \|f\|_{\mathcal{B}_1}. \tag{D–F}$$

Proof. Doeblin–Fortet inequality (D–F) is a consequence of $\|Q_{\theta,p}(t)^n(f)\|_{\mathcal{B}_\beta} \le \|Q_\theta^n(|f|)\|_{\mathcal{B}_\beta}$ (since $\mathcal{B}_\beta$ is a Banach lattice) and (16) and (VG1). Indeed, for all $f \in \mathcal{B}_\beta$, $|f| \in \mathcal{B}_\beta$, and hence one has $\|Q_\theta^n(|f|) - \pi_\theta(|f|)\|_{\mathcal{B}_\beta} \le \widetilde{c}_\beta \kappa_\beta^n \|f\|_{\mathcal{B}_\beta}$, from which we easily deduce the desired inequality, since $\{\pi_\theta; \theta \in \Theta\}$ is bounded in $\mathcal{B}_1'$ (thanks to (VG1)). □

Lemma 5. *Let* $(w_k)_{k\in\mathbb{N}} \in (\mathcal{B}_\beta)^{\mathbb{N}}$ *be such that* $\|w_k\|_{\mathcal{B}_\beta} = 1$ *for all* $k \ge 1$. *If* $(w_k)_{k\in\mathbb{N}}$ *uniformly converges to* $\widetilde{w} \equiv 0$ *on any compact subset of* E, *then* $\sup_{\theta\in\Theta} \|w_k\|_{\mathcal{B}_1} \to 0$ *when* $k \to +\infty$.

Proof. Let $\widetilde{\varepsilon} > 0$, $\varepsilon := 1 - \beta$, and let $K = K_{\widetilde{\varepsilon},\varepsilon}$ be a compact subset of E such that $\sup_{x\in E\setminus K} V(x)^{-\varepsilon} \le \widetilde{\varepsilon}$. Since $|w_k(x)| \le \|w_k\|_{\mathcal{B}_\beta} V(x)^\beta = V(x)^\beta$, one has

$$\forall k \in \mathbb{N}, \quad \|w_k \mathbf{1}_{|E\setminus K}\|_{\mathcal{B}_1} \le \|V^\beta \mathbf{1}_{|E\setminus K}\|_{\mathcal{B}_1} = \sup_{x\in E\setminus K} \frac{V(x)^\beta}{V(x)} \le \widetilde{\varepsilon}. \tag{18}$$

Furthermore $\sup_{x\in K} |w_k(x)|/V(x) \le \sup_{x\in K} |w_k(x)| \to_k 0$, thus there exists $k_0 \in \mathbb{N}$ such that

$$\forall k \ge k_0, \quad \sup_{x\in K} \frac{|w_k(x)|}{V(x)} = \|w_k \mathbf{1}_{|K}\|_{\mathcal{B}_1} \le \widetilde{\varepsilon}. \tag{19}$$

By combining (18) and (19), $\|w_k\|_{\mathcal{B}_1} = \max(\|w_k\mathbf{1}_{|E\setminus K}\|_{\mathcal{B}_1}, \|w_k\mathbf{1}_{|K}\|_{\mathcal{B}_1}) \le \widetilde{\varepsilon}$. □

We are now ready to complete the proof of Theorem 1.

Lemma 6. *We have* (i) $\Rightarrow$ (ii).

Proof. We assume by absurd that (ii) does not hold and (i) holds true, namely on the one hand there exists a compact subset K_0 of $\mathbb{R}^*$ such that $r_{K_0} := \sup\{r(Q(\hbar)_{|\mathcal{B}_\beta}), \hbar \in K_0 \times \Theta \times \mathcal{P}\} \ge 1$, and on the other hand $r_{K_0} \le 1 < +\infty$. Thus there exists $(\hbar_k)_{k\in\mathbb{N}} \in (K_0 \times \Theta \times \mathcal{P})^{\mathbb{N}}$ such that $\lim r(Q(\hbar_k)_{|\mathcal{B}_\beta}) = r_{K_0}$ when $k \to +\infty$, and for all $k \ge 0$, $r(Q(\hbar_k)_{|\mathcal{B}_\beta}) > \kappa$, where κ is defined in Eq. (17) on the essential spectral radius of $Q(\hbar)$. Then for all $k \ge 0$, there exists an eigenvalue λ_k such that $|\lambda_k| = r(Q(\hbar_k)_{|\mathcal{B}_\beta})$. Let $w_k \in \mathcal{B}_\beta$, $w_k \ne 0$, $\|w_k\|_\beta = 1$, such that

$$Q(\hbar_k) w_k = \lambda_k w_k. \tag{20}$$

By compacity argument, we can suppose $\lim \hbar_k := \widetilde{\hbar} = (\widetilde{t}, \widetilde{\theta}, \widetilde{p})$ and $\lim \lambda_k := \widetilde{\lambda}$ when $k \to +\infty$, with $\widetilde{\hbar} \in K_0 \times \Theta \times \mathcal{P}$ and $|\widetilde{\lambda}| = r_{K_0} \ge 1$.

(a) $(w_k)_k$ *converges on* E *to some* $\widetilde{w} \in \mathcal{B}_\beta$: Under the first point of $(\mathcal{S}')$ and the second one of $(\mathcal{C}')$, and using the Ascoli theorem, it is easy to see that $(Q(\hbar_k)w_k)_{k\ge k_0}$ is relatively compact in $(\mathcal{C}(K,\mathbb{R}), \|\cdot\|_\infty)$ for any compact subset K of E. By diagonal extraction, we can suppose that $(Q(\hbar_k)w_k)_{k\in\mathbb{N}}$ converges pointwise on E and uniformly on any compact subset of E, and so does the sequence $(w_k)_{k\in\mathbb{N}}$ thanks to Eq. (20). Its limit is denoted by $\widetilde{w} \in \mathcal{B}_\beta$.

(b) $\widetilde{w} \ne 0$: From Doeblin–Fortet inequality (D–F), from Eq. (20) which implies $Q(\hbar_k)^n w_k = \lambda_k^n w_k$ for all $n \in \mathbb{N}^*$, and from $\|w_k\|_{\mathcal{B}_\beta} = 1$, one obtains $|\lambda_k|^n \le c_\beta \kappa_\beta^n + \|w_k\|_{\mathcal{B}_1}$. Suppose that $\widetilde{w} = 0$. Then $\|w_k\|_{\mathcal{B}_1} \to 0$ when $k \to +\infty$ thanks to Lemma 5. Since $|\lambda_k| \to |\widetilde{\lambda}| = r_{K_0}$ when $k \to +\infty$, one has for all $n \in \mathbb{N}$, $r_{K_0}^n \le c_\beta \kappa_\beta^n$, which is in contradiction with the fact that $\kappa_\beta < 1 \le r_{K_0}$. Consequently $\widetilde{w} \ne 0$.

(c) *Conclusion*: Let $x_0 \in E$. From Assumption $(\mathcal{S}')$, we have

$$Q(\hbar_k) w_k(x_0) := \int_E w_k(y) e^{i t_k \xi_{p_k}(x_0,y)} q_{\theta_k}(x_0, y) \mu_d^{\mathrm{Leb}}(dy).$$

Then, under the second point of $(\mathcal{S}')$ and the first one of $(\mathcal{C}')$, and using Lebesgue dominated convergence theorem, one has $Q(\hbar_k)w_k(x_0) \to_k Q(\widetilde{\hbar})\widetilde{w}(x_0)$. We have just proven that there exist $\widetilde{\lambda} \in \mathbb{C}$, $|\widetilde{\lambda}| = r_{K_0} \ge 1$, a non-null function $\widetilde{w} \in \mathcal{B}_\beta$ and finally a parameter $\widetilde{\hbar} \in K_0 \times \Theta \times \mathcal{P}$ such that $Q(\widetilde{\hbar})\widetilde{w} = \widetilde{\lambda}\widetilde{w}$. This fact implies that $\widetilde{\lambda} \in \sigma(Q(\widetilde{\hbar})_{|\mathcal{B}_\beta})$, which is in contradiction with (i). □

Lemma 7. *We have* (ii) $\Rightarrow$ (iii).

Proof. We prove (ii) $\Rightarrow$ (iii) with $\rho_0 := \max(r_{K_0}, \kappa_\beta)$ where κ_β is defined in (D–F).

Let $K_0 \subset \mathbb{R}^*$ be compact. From (ii), we have $r_{K_0} := \sup\{r(Q(\hbar)_{|\mathcal{B}_\beta}), \hbar \in K_0 \times \Theta \times \mathcal{P}\} < 1$. By absurd, we assume that there exists ρ such that $\rho_0 < \rho < 1$ and $\sup_{|z|=\rho} \sup_{\hbar\in K_0\times\Theta\times\mathcal{P}}\{\|(z - Q(\hbar))^{-1}\|_{\mathcal{B}_\beta}\} = +\infty$. Thus there exist $(\hbar_k, z_k)_{k\in\mathbb{N}} \in (K_0 \times \Theta \times \mathcal{P})^{\mathbb{N}} \times \mathbb{C}^{\mathbb{N}}$, $|z_k| = \rho$, such that $\alpha_k := \|(z_k - Q(\hbar_k))^{-1}\|_{\mathcal{B}_\beta} \to +\infty$ when $k \to +\infty$, which implies by the Banach–Steinhaus theorem that there exists $f \in \mathcal{B}_\beta$ satisfying $\|(z_k - Q(\hbar_k))^{-1} f\|_{\mathcal{B}_\beta} \to +\infty$. Let $w_k := (z_k - Q(\hbar_k))^{-1} f/\alpha_k$ and $\varepsilon_k := f/\alpha_k \in \mathcal{B}_\beta$. Then one has

$$Q(\hbar_k) w_k = z_k w_k - \varepsilon_k. \tag{21}$$

By compacity argument, we can suppose that $\lim_{k\to+\infty} \hbar_k := \widetilde{\hbar} = (\widetilde{t}, \widetilde{\theta}, \widetilde{p})$ and $\lim_{k\to+\infty} z_k := \widetilde{z}$, with $\widetilde{\hbar} \in K_0 \times \Theta \times \mathcal{P}$, and $|\widetilde{z}| = \rho$.

(a) *$(w_k)_k$ converges on E to some $\widetilde{w} \in \mathcal{B}_\beta$*: Again from $(\mathcal{S}')$, $(\mathcal{C}')$, the Ascoli theorem, diagonal extraction and (21), we can suppose that $(w_k)_k$ converges pointwise on E and uniformly on any compact set of E, and we denote its limit by $\widetilde{w} \in \mathcal{B}_\beta$.

(b) $\widetilde{w} \neq 0$: From (21), one can easily show

$$\forall n \in \mathbb{N}, \quad z_k^n w_k = Q(\hbar_k)^n w_k + \sum_{i=0}^{n-1} z_k^i Q(\hbar_k)^{n-1-i} \varepsilon_k. \tag{22}$$

From (D–F), one has for all $\hbar_k = (t_k, \theta_k, p_k) \in \mathbb{R} \times \Theta \times \mathcal{P}$ and $n \in \mathbb{N}$: $\|Q(\hbar_k)^n \varepsilon_k\|_{\mathcal{B}_\beta} \le C_n \|\varepsilon_k\|_{\mathcal{B}_\beta}$ where $C_n := c_\beta(\kappa_\beta^n + 1)$. Recall that $|z_k| = \rho$. Thus considering again Eq. (22) and (D–F), we obtain

$$\rho^n \|w_k\|_{\mathcal{B}_\beta} \le c_\beta \kappa_\beta^n \|w_k\|_{\mathcal{B}_\beta} + c_\beta \|w_k\|_{\mathcal{B}_1} + \sum_{i=0}^{n-1} \rho^i C_{n-1-i} \|\varepsilon_k\|_{\mathcal{B}_\beta}.$$

Suppose that $\widetilde{w} = 0$, then $\|w_k\|_{\mathcal{B}_1} \to_k 0$ using Lemma 5. Since $\|w_k\|_{\mathcal{B}_\beta} = 1$ and $\|\varepsilon_k\|_{\mathcal{B}_\beta} = \|f\|_{\mathcal{B}_\beta}/\alpha_k \to_k 0$, one has for all $n \in \mathbb{N}$: $\rho^n \le c_\beta \kappa_\beta^n$, which is in contradiction with the fact that $\rho > \kappa_\beta$. Thus we have just proven that $\widetilde{w} \neq 0$.

(c) *Conclusion*: Using Lebesgue dominated convergence theorem, one has $Q(\hbar_k) w_k(x) \to_k Q(\widetilde{\hbar})\widetilde{w}(x)$ for all $x \in E$. We have just proven that there exist $\widetilde{z} \in \mathbb{C}$, $|\widetilde{z}| = \rho$, a non-null function $\widetilde{w} \in \mathcal{B}_\beta$ and a parameter $\widetilde{\hbar} \in K_0 \times \Theta \times \mathcal{P}$ such that $Q(\widetilde{\hbar}) w = \widetilde{z}\widetilde{w}$. This fact implies that $r(Q(\widetilde{\hbar})_{|\mathcal{B}_\beta}) \ge \rho$, which is in contradiction with the fact that $\rho > r_{K_0}$. Thus we have just proven by absurd that (ii) ⇒ (iii). □

□

5. *M*-estimators associated with dominated *V*-geometrically ergodic Markov chains

Let $(X_n)_{n\ge0}$ be a Markov chain with state space $E := \mathbb{R}^d$ and transition kernel $(Q_\theta(x, \cdot); x \in E)$, where θ is a parameter in some compact set Θ. The probability distribution of X_0 is denoted by μ_θ. As before, the underlying probability measure and the associated expectation are denoted by $\mathbb{P}_{\theta,\mu_\theta}$ and $\mathbb{E}_{\theta,\mu_\theta}$.

Let us introduce the parameter of interest $\alpha = \alpha(\theta) \in \mathcal{A}$ where $\mathcal{A}$ is an open interval of $\mathbb{R}$. To define the so-called true value of the parameter of interest $\alpha_0 = \alpha_0(\theta) \in \mathcal{A}$, we introduce the empirical mean functional

$$\forall \alpha \in \mathcal{A}, \forall n \in \mathbb{N}^*, \quad M_n(\alpha) := \frac{1}{n} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k), \tag{23}$$

where $F(\cdot,\cdot,\cdot)$ is a real-valued measurable function on $\mathcal{A} \times E^2$. For instance, $-M_n$ may be the log-likelihood of data $(X_0, \dots, X_n)$. We define α_0 as follows

$$\forall \theta \in \Theta, \quad \alpha_0(\theta) := \arg\min_{\alpha\in\mathcal{A}} \lim_{n\to+\infty} \mathbb{E}_{\theta,\mu_\theta}\big[M_n(\alpha)\big], \tag{24}$$

and its M-estimator is supposed to be well-defined by

$$\forall n \in \mathbb{N}^*, \quad \hat{\alpha}_n := \arg\min_{\alpha\in\mathcal{A}} M_n(\alpha). \tag{25}$$

Our goal is to provide an asymptotic expansion of $\mathbb{P}_{\theta,\mu_\theta}\{\sqrt{n}(\hat{\alpha}_n - \alpha_0)/\sigma(\theta) \le u\}$ uniformly in $\theta \in \Theta$ and $u \in \mathbb{R}$, where σ is some suitable (asymptotic) standard deviation. As in the i.i.d. case (see, for example, [25]), we assume throughout this section that the following hypotheses on $(\hat{\alpha}_n)_{n\in\mathbb{N}^*}$ hold true:

$$\text{HYP.1.}\quad \forall n \ge 1, (\partial M_n/\partial\alpha)(\hat{\alpha}_n) = 0. \tag{26a}$$

$$\text{HYP.2.}\quad \forall d > 0, \sup_{\theta\in\Theta} \mathbb{P}_{\theta,\mu_\theta}\bigl\{|\hat{\alpha}_n - \alpha_0| \ge d\bigr\} = \mathrm{o}\bigl(n^{-1/2}\bigr). \tag{26b}$$

Notice that the uniform consistency property (26b) has already been studied in a Markovian context, see, for example, [3,6,14,26,28].

Throughout the sequel, we assume that $(X_n)_{n\in\mathbb{N}}$ belongs to the class of Models $(\mathcal{M})$ (namely $(X_n)_{n\in\mathbb{N}}$ is V-geometrically ergodic uniformly in θ) and that the family of initial distributions $(\mu_\theta)_{\theta\in\Theta}$ satisfies (13). In particular, this last condition will be satisfied if $\mu_\theta \equiv \pi_\theta$ (see (VG1)), or if $\mu_\theta \equiv \delta_x$, where δ_x is the Dirac distribution at any $x \in E$ (independent on θ). Then, under some further conditions on the model and on the function F, we prove[5] that there exists a polynomial function $A_\theta(\cdot)$ such that

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\mu_\theta}\left\{\frac{\sqrt{n}}{\sigma(\theta)}(\hat{\alpha}_n - \alpha_0) \le u\right\} - \mathcal{N}(u) - \eta(u)n^{-1/2}A_\theta(u)\right| = \mathrm{o}\bigl(n^{-1/2}\bigr). \tag{27}$$

Notice that the true value of the parameter of interest (see (24)) can also be defined by

$$\forall\theta \in \Theta, \forall\alpha \in \mathcal{A}, \alpha \ne \alpha_0, \quad \mathbb{E}_{\theta,\pi_\theta}\bigl[F(\alpha, X_0, X_1)\bigr] > \mathbb{E}_{\theta,\pi_\theta}\bigl[F(\alpha_0, X_0, X_1)\bigr].$$

Asymptotic expansions for M-estimators in the Markovian case have already been studied in several papers. Indeed maximum likelihood estimators are fully studied in [5] and [21] in the specific case of stationary Gaussian processes. Some M-estimators for general non-stationary models are also studied in [13] and [15], but each author needs some additional Cramér-type hypothesis. Here we only need the much weaker non-arithmeticity condition. Furthermore our moment conditions on F and its derivatives are almost optimal with respect to the i.i.d. case, see the comments after Theorem 2.

In addition to the previous assumptions (namely $(X_n)_{n\in\mathbb{N}}$ belongs to the class of Models $(\mathcal{M})$ and $(\hat{\alpha}_n)_{n\in\mathbb{N}^*}$ satisfies (26a) and (26b)), we assume that $(X_n)_{n\in\mathbb{N}}$ is dominated, i.e., that condition $(\mathcal{S}')$ holds true (see its definition in Section 4.2). Furthermore, we assume that for all $x \in E$ and for μ_d^{Leb}-almost all $y \in E$, we have $q_\theta(x, y) > 0$.

Let us introduce the assumptions concerning the real-valued measurable function F involved in (23). Assume that the map $\alpha \mapsto F(\alpha, \cdot, \cdot)$ is 3-time-differentiable on $\mathcal{A}$ and let $F^{(j)} := \partial^j F/\partial\alpha^j$ denote the derivatives for $j = 1, 2, 3$. Assume that $F^{(1)}$, $F^{(2)}$, $F^{(3)}$ satisfy the following moment domination condition (D_3):

$$\exists\varepsilon > 0 \text{ such that } \forall j = 1, 2, 3, \quad \sup\left\{\frac{|F^{(j)}(\alpha, x, y)|^{3+\varepsilon}}{V(x) + V(y)}; (x, y) \in E^2, \alpha \in \mathcal{A}\right\} < +\infty. \tag{28}$$

We introduce for $j = 1, 2, 3$

$$\forall\theta \in \Theta, \quad m_j(\theta) := \mathbb{E}_{\theta,\pi_\theta}\bigl[F^{(j)}(\alpha_0, X_0, X_1)\bigr], \qquad \sigma_j(\theta)^2 := \lim_{n\to+\infty} \mathbb{E}_{\theta,\pi_\theta}\Bigl[n\bigl(M_n^{(j)}(\alpha_0) - m_j(\theta)\bigr)^2\Bigr],$$

where $M_n(\cdot)$ is given in (23) and $M_n^{(j)} := \partial^j M_n/\partial\alpha^j$, and where π_θ is the Q_θ-invariant probability measure given in $(\mathcal{M})$. Then, from (28) and using Proposition 1 and Proposition 3, the functions $\sigma_j(\cdot)$ for $j = 1, 2, 3$ are well-defined and bounded in $\theta \in \Theta$.

We consider the following additional assumptions:

C.1. $m_1 \equiv 0$ and $\inf_{\theta\in\Theta} m_2(\theta) > 0$.
C.2. $\inf_{\theta\in\Theta} \sigma_j(\theta) > 0$ for $j = 1, 2$.

[5] A small part of this work has been announced in [10], note without proof.

C.3. There exists a measurable function $W: E \to [0,+\infty)$ of the type $W = CV^\eta$ for some $\eta \in (0,1/2)$ and $C > 0$ such that

$$\forall(\alpha,\alpha') \in \mathcal{A}^2, \forall(x,y) \in E^2, \quad \left|F^{(3)}(\alpha',x,y) - F^{(3)}(\alpha,x,y)\right| \le \left|\alpha'-\alpha\right|\left(W(x)+W(y)\right).$$

Let us introduce some assumptions similar to $(\mathcal{C})$ and $(\mathcal{C}')$ (see definitions in Sections 3.5 and 4.2) concerning the regularity of $(F^{(j)})_{j=1,2,3}$. The function F is supposed to satisfy

C.4. For all $j = 1,2,3$ and $\alpha \in \mathcal{A}$, $F^{(j)}(\alpha,\cdot,\cdot)$ is continuous from E^2 into $\mathbb{R}$.
C.5. For all $x_0 \in E$ and $\alpha \in \mathcal{A}$, there exist neighborhoods $\mathcal{V}_3 = \mathcal{V}_3(x_0,\alpha)$ of x_0 and $\mathcal{V}_4 = \mathcal{V}_4(\alpha)$ of α, positive real numbers C, υ_1 and υ_2 such that for all $\alpha' \in \mathcal{V}_4$, $x \in \mathcal{V}_3$ and $y \in E$:

$$\forall j = 1,2,3, \quad \left|F^{(j)}(\alpha',x,y) - F^{(j)}(\alpha',x_0,y)\right| \le C\|x-x_0\|^{\upsilon_1} V(y)^{\upsilon_2}.$$

Theorem 2. *Assume that all the preceding assumptions hold true, that Hypothesis* (N-L)$'$ (*see definition page* 789) *is satisfied by the following functions*

(a) $\forall p = (\alpha,j) \in \mathcal{A}\times\{1,2\}$, $\xi_p(x,y) := F^{(j)}(\alpha,x,y)$,
(b) $\forall p = (\alpha,\upsilon) \in \mathcal{A}\times\mathbb{R}$, $\xi_p(x,y) := F^{(1)}(\alpha,x,y) + \upsilon F^{(2)}(\alpha,x,y) + (\upsilon^2/2)F^{(3)}(\alpha,x,y)$

and that the initial probability measure satisfies (13), *namely* $\sup_{\theta\in\Theta}\mu_\theta(V) < +\infty$. *Then for* $j = 1,2,3$,

$$\forall\theta\in\Theta, \quad m_j(\theta) = \lim_{n\to+\infty}\mathbb{E}_{\theta,\mu_\theta}\left[M_n^{(j)}(\alpha_0)\right], \qquad \sigma_j(\theta)^2 = \lim_{n\to+\infty}\mathbb{E}_{\theta,\mu_\theta}\left[n\left(M_n^{(j)}(\alpha_0) - m_j(\theta)\right)^2\right]$$

and there exists a polynomial function denoted by A_θ *such that* $(\hat\alpha_n)_{n\in\mathbb{N}^*}$ *satisfies expansion* (27) *with* $\sigma := \sigma_1/m_2$. *Furthermore, the coefficients of* A_θ *are bounded, and*

$$A_\theta(u) := \left[-\frac{1}{6}\frac{m_{3,1}(\theta)^3}{\sigma_1(\theta)^3} + \frac{b_1(\theta)}{\sigma_1(\theta)}\right] + \left[\frac{1}{6}\frac{m_{3,1}(\theta)^3}{\sigma_1(\theta)^3} - \frac{\sigma_{12}(\theta)}{\sigma_1(\theta)m_2(\theta)} + \frac{\sigma_1(\theta)}{2m_2(\theta)^2}m_3(\theta)\right]u^2,$$

where

$$\begin{cases} b_1(\theta) := \lim_{n\to+\infty}\mathbb{E}_{\theta,\mu_\theta}[nM_n^{(1)}(\alpha_0)], \\ \sigma_{12}(\theta) := \lim_{n\to+\infty}\mathbb{E}_{\theta,\pi_\theta}[nM_n^{(1)}(\alpha_0)(M_n^{(2)}(\alpha_0) - m_2(\theta))] \\ \qquad\quad = \lim_{n\to+\infty}\mathbb{E}_{\theta,\mu_\theta}[nM_n^{(1)}(\alpha_0)(M_n^{(2)}(\alpha_0) - m_2(\theta))], \\ m_{3,1}(\theta)^3 := \lim_{n\to+\infty}\mathbb{E}_{\theta,\pi_\theta}[n^2(M_n^{(1)}(\alpha_0))^3] = \lim\mathbb{E}_{\theta,\mu_\theta}[n^2(M_n^{(1)}(\alpha_0))^3] - 3\sigma_1^2(\theta)b_1(\theta). \end{cases}$$

When comparing with Pfanzagl results [25] in the i.i.d. case, expansion (27) proven in Theorem 2 is a natural substitute of the i.i.d. expansion, with an additional term due to the asymptotic bias (namely $b_1(\cdot)$). To the best of our knowledge, Theorem 2 is the most precise statement concerning the first-order Edgeworth expansion of real-valued M-estimators associated with V-geometrically ergodic Markov chains: in fact, the dominated model condition $(\mathcal{S}')$ on the model is classical in Markovian statistics, the condition (D_3) on the derivatives of F is the expected one (up to the real number $\varepsilon > 0$), conditions (C.1)–(C.5) are the Markovian substitutes of the so-called regularity conditions of the i.i.d. case and finally the non-lattice-type conditions (a) and (b) in Theorem 2 are quite general and easy to check.

The proof of Theorem 2 is postponed to Section 6.

As a direct application of Theorem 2, see the illustration given in [9] to autoregressive models AR(1).

6. Pfanzagl method to prove Theorem 2

In Sections 6.1 and 6.2, we adapt Pfanzagl results to some general setting. More specifically, some probabilistic Edgeworth expansions are explicitly required in Section 6.1, whereas the general Assumptions $\mathcal{R}(m)$ and (N-A) are involved in Section 6.2. Thanks to this work, Theorem 2 is easily proved in Section 6.3.

In Sections 6.1 and 6.2, we denote by $(\Omega,\mathcal{F},\{\mathbb{P}_\theta;\theta\in\Theta\})$ a general statistical model, where Θ is some parameter space (not necessarily compact in these subsections). The underlying expectation is denoted by $\mathbb{E}_\theta$. We assume that the following general statistical assumptions hold true: let $(M_n(\alpha))_{n\in\mathbb{N}^*}$ be any general sequence of real observations where $\alpha \equiv \alpha(\theta) \in \mathcal{A}$ is the parameter of interest and $\mathcal{A}$ is an open interval on the real line. Assume that for all $n \geq 1$, the map $\alpha \mapsto M_n(\alpha)$ is 3-time-differentiable on $\mathcal{A}$ and that the derivatives define r.v. on $(\Omega, \mathcal{F})$. We denote them by $M_n^{(j)} := \partial^j M_n/\partial\alpha^j$ for $j = 1, 2, 3$. We consider some $\alpha_0 \in \mathcal{A}$ and assume that the $\mathcal{A}$-valued r.v. $\hat{\alpha}_n$ is specified by (26a) and fulfills the uniform consistency property (26b).

Note that, in Sections 6.1 and 6.2, $(M_n(\alpha))_{n\in\mathbb{N}^*}$ is not necessarily associated with a function F as in (23).

6.1. *The revisited Pfanzagl method*

We appeal to the following conditions:

A.1. For all $n \geq 1$, there exists a positive r.v. W_n, independent on α, such that

$$\forall j = 2, 3, \forall(\alpha, \alpha') \in \mathcal{A}^2, \quad \left|M_n^{(j)}(\alpha') - M_n^{(j)}(\alpha)\right| \leq \left|\alpha' - \alpha\right| W_n.$$

Furthermore there exists $l : \Theta \to (0, +\infty)$ bounded on Θ such that $\sup_{\theta\in\Theta} \mathbb{P}_\theta\{W_n \geq l(\theta)\} = \mathrm{o}(n^{-1/2})$.

A.2. For $j = 1, 2$, there exist $\sigma_j(\cdot) > 0$ satisfying $\sup_{\theta\in\Theta} \sigma_j(\theta) < +\infty$, $\inf_{\theta\in\Theta} \sigma_1(\theta) > 0$, $m_2(\cdot)$ satisfying $\inf_{\theta\in\Theta} m_2(\theta) > 0$, and polynomial functions denoted by $B_\theta(\cdot)$ and $C_\theta(\cdot)$, such that

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_1(\theta)} M_n^{(1)}(\alpha_0) \leq u\right\} - \mathcal{N}(u) - \eta(u)n^{-1/2}B_\theta(u)\right| = \mathrm{o}\left(n^{-1/2}\right),$$

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_2(\theta)}\left(M_n^{(2)}(\alpha_0) - m_2(\theta)\right) \leq u\right\} - \mathcal{N}(u) - \eta(u)n^{-1/2}C_\theta(u)\right| = \mathrm{o}\left(n^{-1/2}\right).$$

Furthermore the coefficients of $B_\theta(\cdot)$ and $C_\theta(\cdot)$ are assumed to be uniformly bounded with respect to $\theta \in \Theta$.

Let us define $\sigma(\theta) := \sigma_1(\theta)/m_2(\theta)$. Notice that $\sigma(\cdot)$ satisfies $\sup_{\theta\in\Theta} \sigma(\theta) < +\infty$.

A.3. For all $n \geq 1$, $u \in \mathbb{R}$ such that $|u| \leq 2\sqrt{\ln n}$, there exist $\sigma_{n,u}(\cdot)^2 > 0$, $m_3(\cdot)$ bounded on Θ, $D_\theta(\cdot)$ and $E_\theta(\cdot)$ some polynomial functions such that

$$\sup_{\theta\in\Theta}\sup_{|u|\leq 2\sqrt{\ln n}}\left|\sigma_{n,u}(\theta)^{-1} - \left(\sigma_1(\theta)^{-1} + D_\theta(u)n^{-1/2}\right)\right| = \mathrm{o}\left(n^{-1/2}\right),$$

$$\sup_{\theta\in\Theta}\sup_{v\in\mathbb{R}}\sup_{|u|\leq 2\sqrt{\ln n}}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_{n,u}(\theta)}\widetilde{M}_n(\theta, u) \leq v\right\} - \mathcal{N}(v) - \eta(v)E_\theta(v)n^{-1/2}\right| = \mathrm{o}\left(n^{-1/2}\right),$$

where $\widetilde{M}_n(\theta, u)$ denotes

$$\widetilde{M}_n(\theta, u) := M_n^{(1)}(\alpha_0) + \frac{\sigma(\theta)}{\sqrt{n}}u\left(M_n^{(2)}(\alpha_0) - m_2(\theta)\right) + \left(\frac{\sigma(\theta)}{\sqrt{n}}\right)^2\frac{u^2}{2}\left(M_n^{(3)}(\alpha_0) - m_3(\theta)\right). \tag{29}$$

Furthermore, the coefficients of $D_\theta(\cdot)$ and $E_\theta(\cdot)$ are assumed to be uniformly bounded with respect to $\theta \in \Theta$.

Theorem 3. *Under conditions* (A.1), (A.2) *and* (A.3), *there exists a polynomial function* A_θ *such that one has with* $\sigma := \sigma_1/m_2$

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma(\theta)}(\hat{\alpha}_n - \alpha_0) \leq u\right\} - \mathcal{N}(u) - \eta(u)n^{-1/2}A_\theta(u)\right| = \mathrm{o}\left(n^{-1/2}\right). \tag{30}$$

Furthermore

$$\forall\theta \in \Theta, \forall u \in \mathbb{R}, \quad A_\theta(u) := D_\theta(u)\sigma_1(\theta)u + \frac{\sigma(\theta)^2}{2\sigma_1(\theta)}m_3(\theta)u^2 - E_\theta(-u). \tag{31}$$

The proof of Theorem 3 is postponed to the Appendix. It consists in adapting the Pfanzagl method [25] introduced for i.i.d. observations. Just mention that the Pfanzagl method is not exactly the one developed in the Appendix, but for convenience this discussion is omitted.

6.2. *An alternative statement using Hypotheses $\mathcal{R}(m)$ and* (N-A)

Below we appeal to the following assumptions involving Hypotheses $\mathcal{R}(m)$ and (N-A) of Section 2.1:

B.1. For all $n \geq 1$, there exists a positive r.v. W_n, independent on α, such that

$$\forall j = 2, 3, \forall (\alpha, \alpha') \in \mathcal{A}^2, \quad \left| M_n^{(j)}(\alpha') - M_n^{(j)}(\alpha) \right| \leq \left| \alpha' - \alpha \right| W_n.$$

Furthermore there exists $\widetilde{l} : \Theta \to (-1, +\infty)$ bounded on Θ such that $\{n(W_n - \widetilde{l}(\theta)); n \geq 1, \theta \in \Theta\}$ fulfills Hypothesis $\mathcal{R}(2)$.

B.2. The family $\{n M_n^{(1)}(\alpha_0); n \geq 1, \theta \in \Theta\}$ fulfills Hypotheses $\mathcal{R}(3)$ and (N-A). Furthermore there exists $m_2(\cdot)$ on Θ satisfying $\inf_{\theta \in \Theta} m_2(\theta) > 0$ such that $\{n(M_n^{(2)}(\alpha_0) - m_2(\theta)); n \geq 1, \theta \in \Theta\}$ fulfills both Hypotheses $\mathcal{R}(3)$ and (N-A).

Thanks to the last assumption (B.2) and Proposition 1, we can define the asymptotic variances

$$\sigma_1(\theta)^2 := \lim_{n \to +\infty} \mathbb{E}_\theta\left[n\left(M_n^{(1)}(\alpha_0)\right)^2\right], \qquad \sigma_2(\theta)^2 := \lim_{n \to +\infty} \mathbb{E}_\theta\left[n\left(M_n^{(2)}(\alpha_0) - m_2(\theta)\right)^2\right].$$

Furthermore we assume that the following conditions on these asymptotic variances hold true

B.3. $\inf_{\theta \in \Theta} \sigma_1(\theta) > 0$,

B.4. $\inf_{\theta \in \Theta} \sigma_2(\theta) > 0$.

The following additional conditions are also required:

B.5. There exists $m_3(\cdot)$, bounded on Θ, such that $\{n(M_n^{(3)}(\alpha_0) - m_3(\theta)); n \geq 1, \theta \in \Theta\}$ fulfills Hypothesis $\mathcal{R}(2)$, and $\{n \widetilde{M}_n(\theta, u); n \geq 1, \theta \in \Theta, |u| \leq 2\sqrt{\ln n}\}$ fulfills both Hypotheses $\mathcal{R}(3)$ and (N-A) as well, where $\widetilde{M}_n(\theta, u)$ is defined by (29).

Theorem 4. *Under assumptions* (B.1) *to* (B.5), *there exists a polynomial function A_θ independent on n such that one has* (30) *with $\sigma := \sigma_1/m_2$. The polynomial function A_θ is of the type $A_\theta(u) = a_1(\theta) + a_2(\theta)u^2$ where, for $i = 1, 2$, $\sup_{\theta \in \Theta} |a_i(\theta)| < +\infty$. Furthermore if we suppose that the additional moment condition holds true*:

$$\forall j = 1, 2, 3, \forall \theta \in \Theta, \forall n \in \mathbb{N}^*, \quad \mathbb{E}_\theta\left[\left|M_n^{(j)}(\alpha_0)\right|^3\right] < +\infty, \tag{32}$$

then one has more precisely

$$a_1(\theta) := -\frac{1}{6}\frac{m_{3,1}(\theta)^3}{\sigma_1(\theta)^3} + \frac{b_1(\theta)}{\sigma_1(\theta)}, \qquad a_2(\theta) := \frac{1}{6}\frac{m_{3,1}(\theta)^3}{\sigma_1(\theta)^3} - \frac{\sigma_{12}(\theta)}{\sigma_1(\theta) m_2(\theta)} + \frac{\sigma_1(\theta)}{2 m_2(\theta)^2} m_3(\theta)$$

with

$$\begin{cases} b_1(\theta) := \lim_{n \to +\infty} \mathbb{E}_\theta[n M_n^{(1)}(\alpha_0)], \\ \sigma_{12}(\theta) := \lim_{n \to +\infty} \mathbb{E}_\theta[n M_n^{(1)}(\alpha_0)(M_n^{(2)}(\alpha_0) - m_2(\theta))], \\ m_{3,1}(\theta)^3 := \lim_{n \to +\infty} \mathbb{E}_\theta[n^2 (M_n^{(1)}(\alpha_0))^3] - 3\sigma_1^2(\theta) b_1(\theta). \end{cases}$$

Proof. It is sufficient to show that the assumptions of Theorem 4 imply the previous ones of Theorem 3.

- From (B.1), (A.1) holds true with $l := \widetilde{l} + 1$. Indeed let $S_n(\theta) := n(W_n - \widetilde{l}(\theta))$. Thanks to (4) and Markov inequality, one easily obtains $\sup_{\theta \in \Theta} \mathbb{P}_\theta\{W_n \geq l(\theta)\} \leq (1/n) \sup_{\theta \in \Theta} (\mathbb{E}_\theta[S_n(\theta)^2]/n) = \mathrm{O}(n^{-1})$.

- (A.2) is directly implied by (B.2), (B.3), (B.4) using Proposition 1. Furthermore, under (32), we have

$$B_\theta(u) := \frac{m_{3,1}(\theta)^3}{6\sigma_1(\theta)^3}\left(1-u^2\right) - \frac{b_1(\theta)}{\sigma_1(\theta)}.$$

  Similar expression holds for $C_\theta(u)$.
- In the same way, to prove (A.3) under (B.2), (B.3), (B.5), let us define

$$S_n(\theta, p, v) := n\left[M_n^{(1)}(\alpha_0) + \varsigma_p(v,\theta)\left(M_n^{(2)}(\alpha_0) - m_2(\theta)\right) + \frac{\varsigma_p(v,\theta)^2}{2}\left(M_n^{(3)}(\alpha_0) - m_3(\theta)\right)\right],$$

  where $\varsigma_p(v,\theta) := v\sigma(\theta)/\sqrt{p}$, so that $S_n(\theta, n, u) = n\widetilde{M}_n(\theta, u)$ (cf. (29)). From (B.5) and using Proposition 1, we can define

$$\sigma_{p,v}(\theta)^2 := \lim_{n\to+\infty} \frac{\mathbb{E}_\theta[S_n(\theta,p,v)^2]}{n},$$

  and from (B.3) and using Proposition 1 again, we obtain (A.3). Moreover, under (32), we have

$$D_\theta(u) := -\frac{\sigma_{12}(\theta)}{\sigma_1(\theta)^3}\sigma(\theta)u \quad \text{and} \quad E_\theta(u) := \frac{m_{3,1}(\theta)^3}{6\sigma_1(\theta)^3}\left(1-u^2\right) - \frac{b_1(\theta)}{\sigma_1(\theta)}. \qquad \square$$

6.3. *Proof of Theorem 2*

Let us go back to our Markovian model $(\mathcal{M})$ and prove that the assumptions of Theorem 4 hold true whenever the assumptions of Theorem 2 are satisfied.

- Let us define $W_n := (1/n)\sum_{k=1}^n (W(X_{k-1}) + W(X_k))$ and $\widetilde{l}(\theta) := 2\mathbb{E}_{\theta,\pi_\theta}[W(X_1)]$, where W is defined in (C.3). Then, using Proposition 3, the Lipschitz condition (B.1) for $j=3$ is true. Indeed the family $\{\xi_\theta; \theta \in \Theta\}$ obviously fulfills the moment domination condition (D_2) with $\xi_\theta(x,y) := W(x) + W(y) - \widetilde{l}(\theta)$ and using (VG1). In the same way, the remaining part of (B.1) (for $j=2$) is checked under (28) (which means that the family $\{F^{(3)}(\alpha,\cdot,\cdot); \alpha \in \mathcal{A}\}$ fulfills (D_3) and a fortiori (D_2)).
- Firstly, we deduce from Proposition 3 that the part of (B.2) concerning Hypothesis $\mathcal{R}(3)$ is true under (28) since $m_1 \equiv 0$ (from (C.1)). Secondly, thanks to Assumption $(\mathcal{S}')$, we deduce from Lemma 3 and Theorem 1 that the part of (B.2) concerning Hypothesis (N-A) is true under condition (a) of Theorem 2 (see Remark 3 concerning the assumptions of Lemma 3).
- The conditions (B.3) and (B.4) are exactly (C.2).
- We deduce from Proposition 3 that the first point of (B.5) follows from (28). For the second point of (B.5), recall definition (29) of $\widetilde{M}_n(\theta,u)$, and notice that $\inf_{\theta\in\Theta} m_2(\theta) > 0$ from (C.1) and $\sup_{\theta\in\Theta}\sigma_1(\theta) < +\infty$, which imply that $\sup\{\sigma(\theta)u/\sqrt{n}; n \ge 1, \theta \in \Theta, |u| \le 2\sqrt{\ln n}\} < +\infty$. Thus the family

$$\left\{\sum_{j=1}^{3}\frac{1}{(j-1)!}\left(\frac{\sigma(\theta)}{\sqrt{n}}u\right)^{j-1}\left(F^{(j)}(\alpha_0,\cdot,\cdot) - m_j(\theta)\right); n \ge 1, \theta \in \Theta, |u| \le 2\sqrt{\ln n}\right\}$$

  obviously fulfills (D_3), and we conclude from Proposition 3 that $\{n\widetilde{M}_n(\theta,u); n \ge 1, \theta \in \Theta, |u| \le 2\sqrt{\ln n}\}$ fulfills Hypothesis $\mathcal{R}(3)$. Finally, thanks to Assumption $(\mathcal{S}')$, we easily check from Lemma 3 and Theorem 1 that the part of (B.5) concerning Hypothesis (N-A) is true under condition (b) of Theorem 2.

6.4. *Illustration of Theorem 4 in the case of some* AR(d) *processes*

In this subsection, we apply Theorem 4 to some linear autoregressive processes of order d, $d \ge 2$. In substance, such a model fulfills all the assumptions of Theorem 2, except the dominated model condition $(\mathcal{S}')$. Consequently, the non-arithmeticity conditions involved in the assumptions B.2 and B.5 of Theorem 4 cannot be checked using

Theorem 1 any more. Here, by reinforcing the assumptions on the density of the noise, we apply the second approach of Section 3.3 to study these non-arithmeticity conditions.

Let us consider the following autoregressive process of order $d \geq 1$ on $E := \mathbb{R}^d$:

$$\forall n \geq d, \quad Y_n := g_1(\theta)Y_{n-1} + \cdots + g_d(\theta)Y_{n-d} + Z_n, \tag{33}$$

where the probability distribution of $(Y_0, \ldots, Y_{d-1})$ is denoted by μ_θ and

- $\theta \in \mathbb{R}$ is a parameter,
- $(g_1, \ldots, g_d)$ are given real continuous functions,
- and $(Z_k)_{k\in\mathbb{N}^*}$ are i.i.d. real-valued r.v. independent on $(Y_0, \ldots, Y_{d-1})$ with common distribution which admits some probability density f_Z with respect to μ^{Leb}.

The parameter θ of the observed AR(d) process is assumed to be in a non-empty compact set $\Theta \subset \mathbb{R}$ such that for all $\theta \in \Theta$ the solutions of the equation

$$z^d - g_1(\theta)z^{d-1} - \cdots - g_{d-1}(\theta)z - g_d(\theta) = 0 \tag{34}$$

lie in $D(0,1) := \{z \in \mathbb{C}; |z| < 1\}$.

Introduce the column vector $X_n := (Y_n, \ldots, Y_{n-d+1})'$ for $n \geq d-1$. Then the process $(X_n)_{n\geq d-1}$ is a Markov chain with the following first-order autoregressive representation

$$\forall n \geq d \quad X_n = A(\theta)X_{n-1} + (Z_n, 0, \ldots, 0)', \tag{35}$$

where

$$A(\theta) := \begin{pmatrix} g_1(\theta) & & \cdots & & g_{d-1}(\theta) & g_d(\theta) \\ 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & & \ddots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 \end{pmatrix}.$$

Assuming that the solutions of Eq. (34) lie in $D(0,1)$ is equivalent to assume that the eigenvalues of $A(\theta)$ have moduli strictly less than unity, so that $\|A(\theta)\| < 1$ for all $\theta \in \Theta$ and $\sup_{\theta\in\Theta} \|A(\theta)\| < 1$.

The initial distribution of $(X_n)_{n\geq d-1}$ is μ_θ and its transition kernel Q_θ is given for all Borel set $B \in \mathcal{B}(\mathbb{R}^d)$ by

$$Q_\theta(x, B) = \int_{\mathbb{R}} \mathbf{1}_B\big(A(\theta)x + (z, 0, \ldots, 0)'\big) f_Z(z)\,\mathrm{d}z.$$

Note that the transition kernel Q_θ has the following representation:

$$Q_\theta(x, \mathrm{d}y) = f_Z\big(y_d - \langle g(\theta), x\rangle\big)\mu^{\mathrm{Leb}}(y_d)\delta_{x_d}(y_{d-1})\cdots\delta_{x_2}(y_1), \tag{36}$$

where x denotes the column vector $(x_d, \ldots, x_1)'$ and y denotes the column vector $(y_d, \ldots, y_1)'$. Then, as already mentioned, the dominated model condition $(\mathcal{S}')$ is not fulfilled in the multidimensional case $(d \neq 1)$.

Next, let us assume that the probability density f_Z of Z_1 fulfills the following conditions:

(A) $\forall z \in \mathbb{R}$, $f_Z(z) > 0$;
(B) $\mathbb{E}[Z_1] = 0$;
(C) $\mathbb{E}[|Z_1|^{10}] < +\infty$;
(D) f_Z is 4-time-differentiable on $\mathbb{R}$;
(E) for $j = 1, \ldots, 4$, $f_Z^{(j)}/f_Z$ is a bounded function;

(F) for all $9 < \gamma \leq 10$, there exists $0 < \beta \leq 1 - 1/\gamma$ such that, for all $x_0 \in \mathbb{R}$, there exists a neighborhood V_{x_0} of x_0 and a positive measurable function q_{x_0} satisfying

$$\int_{\mathbb{R}} \bigl(1+|y|\bigr)^{\beta} q_{x_0}(y)\,\mathrm{d}y < \infty \quad \text{and} \quad \forall y \in \mathbb{R}, \forall t \in V_{x_0}, f_Z(y+t) \leq q_{x_0}(y).$$

Actually, under (E), it is sufficient to assume in (C) that there exists some $\varepsilon > 0$ such that for all $\theta \in \Theta$, $\mathbb{E}_{\theta,\mu_\theta}[Z_1^{9+\varepsilon}] < +\infty$. Furthermore, note that assumption (E) can be relaxed provided that the order of the moment of Z_1 is increased. However, assumption (E) as above is satisfied in several interesting models and thus it does not need relaxing. Notice that Götze and Hipp [15], Theorem 1.5, assume that, under (E), Z_1 admits a moment of order 15.

Finally, the vector $g := (g_1, \dots, g_d)'$ is supposed to have the following properties:

(G) $\theta \mapsto (g_1(\theta), \dots, g_d(\theta)) \in \mathbb{R}^d$ is 4-time-continuously-differentiable on Θ;
(H) $\inf_{\theta\in\Theta} g_1^{(1)}(\theta) > 0$.

Let us define for all $9 < \gamma \leq 10$

$$\forall x \in \mathbb{R}^d, \quad V(x) := 1 + \|x\|^{\gamma}. \tag{37}$$

Lemma 8. *Under the previous conditions, $(X_n)_{n\geq d-1}$ belongs the class of Models $(\mathcal{M})$ defined at the beginning of Section 4 with the function V defined in (37).*

Proof. Under (C), one has for all $\theta \in \Theta$ and $x \in \mathbb{R}^d$,

$$\frac{Q_\theta V(x)}{V(x)} = \int_{\mathbb{R}} \frac{V(A(\theta)x + (z,0,\dots,0)')}{V(x)} f_Z(z)\,\mathrm{d}z \leq \int_{\mathbb{R}} \frac{1 + (\|A(\theta)\|\|x\| + |z|)^{\gamma}}{V(x)} f_Z(z)\,\mathrm{d}z$$

and thus we have

$$\limsup_{\|x\|\to\infty} \left(\sup_{\theta\in\Theta} \frac{Q_\theta V(x)}{V(x)}\right) \leq \sup_{\theta\in\Theta} \bigl\|A(\theta)\bigr\|^{\gamma} < 1.$$

Next, pick $\varrho \in (\sup_{\theta\in\Theta} \|A(\theta)\|^{\gamma}, 1)$. There exists $s > 0$ such that $Q_\theta V(x) \leq \varrho V(x)$ for all $\|x\| > s$ and $\theta \in \Theta$. Set $S := \{x \in \mathbb{R}^d; \|x\| \leq s\}$. Note that

$$\forall \theta \in \Theta, \forall x \in S, \quad Q_\theta V(x) \leq \varsigma := \sup_{\theta\in\Theta} \int_{\mathbb{R}} \bigl(1 + \bigl(\bigl\|A(\theta)\bigr\|\|s\| + |z|\bigr)^{\gamma}\bigr) f_Z(z)\,\mathrm{d}z < +\infty,$$

so that:

$$\forall \theta \in \Theta, \forall x \in \mathbb{R}^d, \qquad Q_\theta V(x) \leq \varrho V(x) + \varsigma.$$

Finally, since condition (A) holds true, it is easily checked that $(X_k)_{k\geq 0}$ is μ_d^{Leb}-irreducible, aperiodic and fulfills the drift-criterion [24], Chapter 15.2.2 (V4), uniformly in $\theta \in \Theta$. □

Set $e_1 := (1,0,\dots,0)' \in \mathbb{R}^d$. Then, let us consider the MLE $(\hat{\theta}_n)_{n\in\mathbb{N}^*}$ of the parameter θ ($\alpha_0(\theta) \equiv \theta$). We have

$$\forall n \geq d, \quad \langle e_1, X_n\rangle = \bigl\langle g(\theta), X_{n-1}\bigr\rangle + Z_n.$$

Maximum likelihood estimation requires to deal with the following function F

$$\forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^d, \forall \alpha \in \Theta, \quad F(\alpha, x, y) := -\ln f_Z\bigl(\langle e_1, y\rangle - \bigl\langle g(\alpha), x\bigr\rangle\bigr),$$

and the empirical mean functional

$$\forall n \in \mathbb{N}^*, \forall \alpha \in \Theta, \quad M_n(\alpha) := -\frac{1}{n}\sum_{k=1}^{n} \ln f_Z\bigl(\langle e_1, X_k\rangle - \bigl\langle g(\alpha), X_{k-1}\bigr\rangle\bigr).$$

Let us define

Hypothesis (N-L)$''$. *For all $p \in \mathcal{P}$, there exist neither continuous function $\mathcal{A}_p(\cdot) : \mathbb{R}^d \to \mathbb{R}$ nor constant C_p such that we have for all $x \in \mathbb{R}^d$ and for all $z \in \mathbb{R}$, $\xi_p(x, h_\alpha(x, z)) = \mathcal{A}_p(h_\alpha(x, z)) - \mathcal{A}_p(x) + C_p$, where*

$$\forall x := (x_{(d)}, \dots, x_{(1)})' \in \mathbb{R}^d, \forall z \in \mathbb{R}, \quad h_\alpha(x, z) := \big(z + \langle g(\alpha), x\rangle, x_{(d)}, \dots, x_{(2)}\big)'. \tag{38}$$

Proposition 4. *Assume that the previous assumptions on the model hold true and that the MLE $(\hat{\theta}_n)_{n\in\mathbb{N}^*}$ of the parameter θ associated with $(X_n)_{n\geq d-1}$ satisfies the uniform consistency property* (26b). *In addition, assume that the initial probability measure satisfies* $\sup_{\theta\in\Theta} \mu_\theta(V) < +\infty$ (*that is, property* (13)) *and that condition* (N-L)$''$ *as defined above is satisfied by the following functions*:

(a) $\forall p = (\alpha, j) \in \mathcal{A} \times \{1, 2\}, \xi_p(x, y) := F^{(j)}(\alpha, x, y)$,
(b) $\forall p = (\alpha, \upsilon) \in \mathcal{A} \times \mathbb{R}, \xi_p(x, y) := F^{(1)}(\alpha, x, y) + \upsilon F^{(2)}(\alpha, x, y) + (\upsilon^2/2) F^{(3)}(\alpha, x, y)$.

Then all the conclusions of Theorem 2 are true.

Proof. It is easily checked that the family $\{F^{(j)}(\alpha, x, y); \alpha \in \mathcal{A}, j = 1, 2, 3\}$ fulfills the moment domination condition (D_3) (i.e., (28)) mainly thanks to (E) and $\gamma > 3 \times 3$. Next, we claim that conditions (C.1) to (C.5) of Theorem 2 are satisfied. Indeed, concerning conditions (C.1) and (C.2), and recalling that $(\pi_\theta)_{\theta\in\Theta}$ denotes the invariant probability of the Markov chain $(X_n)_{n\geq d-1}$, we have

- $m_1(\theta) = \mathbb{E}_{\theta,\pi_\theta}[\langle g^{(1)}(\theta), X_0\rangle]\mathbb{E}[f_Z^{(1)}(Z_1)/f_Z(Z_1)] \equiv 0$;
- $m_2(\theta) = \mathbb{E}_{\theta,\pi_\theta}[\langle g^{(1)}(\theta), X_0\rangle^2]\mathbb{E}[f_Z^{(1)}(Z_1)^2/f_Z(Z_1)^2]$.

Furthermore $\mathbb{E}_{\theta,\pi_\theta}[\langle g^{(1)}(\theta), X_0\rangle^2] = \mathbb{E}_{\theta,\pi_\theta}[\langle g^{(1)}(\theta), A(\theta)X_0\rangle^2] + g_1^{(1)}(\theta)^2\mathbb{E}[Z_1^2] \geq g_1^{(1)}(\theta)^2\mathbb{E}[Z_1^2]$ thanks to definition (35) and condition (B). Then, thanks to condition (H), one has $\inf_{\theta\in\Theta} m_2(\theta) > 0$.

- $\sigma_1(\theta)^2 = \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}_{\theta,\pi_\theta}[\sum_{i=1}^n A_i^2 + \sum_{i\neq j} A_i A_j]$, where $A_i := \frac{f_Z^{(1)}(Z_i)}{f_Z(Z_i)}\langle g^{(1)}(\theta), X_{i-1}\rangle$. Hence one has $\sigma_1(\theta)^2 = \mathbb{E}_{\theta,\pi_\theta}[A_1^2] = m_2(\theta)$, $\inf_{\theta\in\Theta} \sigma_1(\theta) > 0$;
- In the same way, $\inf_{\theta\in\Theta} \sigma_2(\theta) > 0$.

Conditions (C.4) and (C.5) are obviously satisfied. Concerning (C.3), use the fact that the family $\{F^{(4)}(\alpha, x, y); \alpha \in \mathcal{A}\}$ fulfills (D_2) (this statement holds true mainly thanks to assumption (E) and $\gamma > 2 \times 4$).

By using the previous facts and proceeding as in the proof of Theorem 2 (see Section 6.3), one can see that all the assumptions of Theorem 4 but those concerning Hypothesis (N-A) are fulfilled. Consequently, to deduce Proposition 4 from Theorem 4, it only remains to establish that the characteristic functions of the following families $(\xi_p)_{p\in\mathcal{P}}$ (involved in assumptions B.2 and B.5 of Theorem 4) satisfy Hypothesis (N-A):

(a) $\forall p = (\alpha, j) \in \mathcal{A} \times \{1, 2\}, \xi_p(x, y) := F^{(j)}(\alpha, x, y)$,
(b) $\forall p = (\alpha, \upsilon) \in \mathcal{A} \times \mathbb{R}, \xi_p(x, y) := F^{(1)}(\alpha, x, y) + \upsilon F^{(2)}(\alpha, x, y) + (\upsilon^2/2) F^{(3)}(\alpha, x, y)$.

To that effect, we make use of the second approach of Section 3.3. Below, (i), (ii) and (iii) refer to the conditions introduced in Section 3.1:

Fact 1. *Families* (a) *and* (b) *satisfy condition* (N-L).

Indeed, thanks to condition (A), Fact 1 follows from assumptions 1 and 2 of this Proposition 4 (see Lemma 9 below and apply it to the case where $(\xi_p)_{p\in\mathcal{P}}$ is any of the above families (a) and (b)).

Lemma 9. *Assume that for all $p \in \mathcal{P}$, the application ξ_p is continuous from $\mathbb{R}^d \times \mathbb{R}^d$ into $\mathbb{R}$. If the family $(\xi_p)_{p\in\mathcal{P}}$ fulfills* (N-L)$''$, *then* (N-L) *is fulfilled.*

Proof. Assume that (N-L) is not fulfilled, that is we have (11) with some $(\theta_0, p_0) \in \Theta \times \mathcal{P}$, $a \in \mathbb{R}$, some closed subgroup $H = c\mathbb{Z}$, some π_{θ_0}-full Q_{θ_0}-absorbing set $A \in \mathcal{B}(\mathbb{R}^d)$, and finally some bounded measurable function

$\alpha:\mathbb{R}^d \to \mathbb{R}$. For the sake of simplicity, let us omit the dependence on (θ_0, p_0). Then, thanks to condition (A), for all $x \in A$, there exists $Z_x \subset \mathbb{R}$ such that $\mu_1^{\mathrm{Leb}}(\mathbb{R}\backslash Z_x) = 0$ and

$$\forall z \in Z_x, \quad h(x,z) \in A \quad \text{and} \quad \xi\big(x, h(x,z)\big) + \alpha\big(h(x,z)\big) - \alpha(x) \in a + H, \tag{39}$$

where $h(x,z)$ is the continuous function defined in (38).

Let $x_0 \in A$. Let us define $x_i := h(x_{i-1}, z_i)$ for all $i \in \{1,\dots,d\}$, for all $z_i \in \mathbb{R}$. Then iterating the previous property (39), one has:

$$\forall z_1 \in Z_{x_0}, \dots, z_d \in Z_{x_{d-1}}, \quad \sum_{i=1}^{d} \xi(x_{i-1}, x_i) + \alpha(x_d) - \alpha(x_0) \in ad + H. \tag{40}$$

The function $(z_1,\dots,z_d) \mapsto x_d$ is injective. Thus one can define the following continuous function $\mathcal{A}: x_d \mapsto \sum_{i=1}^d \xi(x_{i-1}, x_i)$.

Let us define

$$\begin{cases} E_{x_0} := \{x_d; z_1 \in Z_{x_0}, \dots, z_d \in Z_{x_{d-1}}\} \subset \mathbb{R}^d, \\ F_x := \{z \in Z_x; h(x,z) \in E_{x_0}\} \subset \mathbb{R}. \end{cases}$$

Then property (40) is equivalent to the following one:

$$\forall x_d \in E_{x_0}, \quad \mathcal{A}(x_d) + \alpha(x_d) - \alpha(x_0) \in ad + H. \tag{41}$$

Thanks to condition (A) and $Q^d(x_0, \mathbb{R}^d \backslash E_{x_0}) = 0$, one has $\mu_d^{\mathrm{Leb}}(\mathbb{R}^d \backslash E_{x_0}) = 0$.

Thanks to condition (A) and $Q^d(x_0, \mathbb{R}^d \backslash A) = 0$, one has $\mu_d^{\mathrm{Leb}}(\mathbb{R}^d \backslash \{A \cap E_{x_0}\}) = 0$ (recall that A is Q-absorbing, then $Q(a, A) = 1$ for all $a \in A$, and $Q^d(a, A) = 1$).

Let $x \in A \cap E_{x_0}$, then one has thanks to (39) and (41):

$$\forall z \in F_x, \quad \xi\big(x, h(x,z)\big) - \mathcal{A}\big(h(x,z)\big) + \mathcal{A}(x) \in a + H. \tag{42}$$

Let us define $f(x,z) := \xi(x, h(x,z)) + \mathcal{A}(x) - \mathcal{A}(h(x,z))$. Then for all $x \in A \cap E_{x_0}$, $f(x, F_x) \subset a + H$. Then, by continuity arguments, one can easily show that $f(x, \mathbb{R}) \subset a + H$. In the same way, $f(A \cap E_{x_0}, \mathbb{R}) \subset a + H$, and finally $f(\mathbb{R}^d, \mathbb{R}) \subset a + H$. Since $f(\mathbb{R}^d, \mathbb{R})$ is connex and $a + H$ is discrete, f is constant on $\mathbb{R}^d \times \mathbb{R}$. □

Fact 2. *The Fourier operators of families* (a) *and* (b) *satisfy condition* (i).

Indeed assumptions 1 and 2 of Lemma 1 are fulfilled (see the comments after Lemma 1 concerning assumption 1 and property (17) concerning assumption 2). Then, using Lemma 1, Fact 2 follows from Fact 1.

Fact 3. *The Fourier operators of families* (a) *and* (b) *satisfy* (ii) *and* (iii).

Indeed notice that the family $\{F^{(j)}(\alpha, x, y); \alpha \in \mathcal{A}, j = 1,\dots,4\}$ satisfies (D_0) of Proposition 3, and consequently, the assumptions of Proposition 2 are fulfilled (see Lemma 10 below and apply it to the case where $(\xi_p)_{p\in\mathcal{P}}$ is any of the above families (a) and (b)). Then, using Proposition 2, Fact 3 follows from Fact 2.

Fact 4. *Finally the Fourier operators of families* (a) *and* (b) *satisfy Hypothesis* (N-A)′, *and so* (N-A) (*see Lemma* 2 *and see also the end of Section* 2).

The proof of Proposition 4 is now complete, provided that we give the proof of the next lemma. □

Lemma 10. *Assume that $(\xi_p)_{p\in\mathcal{P}}$ and its derivative with respect to p fulfill (D_{m_0}) with some $m_0 \in \mathbb{N}$. Let $Q_{\theta,p}(t)$ be the Fourier operator defined by* (8) *where the transition kernel Q_θ is the particular one given in* (36). *Then, conditions* (C1) *and* (C2) *of Proposition* 2 *hold true.*

Proof. Let us prove that the family $(\xi_p)_{p\in\mathcal{P}}$ verifies the assumptions of Proposition 2 with $\mathcal{B} := \mathcal{B}_\beta \hookrightarrow \widetilde{\mathcal{B}} := \mathcal{B}_1$ where β is defined in assumption (F) (see page 792 for the definition of the spaces).

First, condition (C1) of Proposition 2 is exactly (D–F) (see page 792 for the definition of condition (D–F)). Then, concerning condition (C2) of Proposition 2, let us prove that the following properties are valid:

1. The map $t \mapsto Q_{\theta,p}(t)$ is continuous from $\mathbb{R}$ into $\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_1)$ uniformly in $(\theta, p) \in \Theta \times \mathcal{P}$;
2. For all $t \in \mathbb{R}$, the map $(\theta, p) \mapsto Q_{\theta,p}(t)$ is continuous from $\Theta \times \mathcal{P}$ into $\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_1)$.

Then one obviously has $\|Q_{\theta,p}(t) - Q_{\theta_0,p_0}(t_0)\|_{\mathcal{B}_\beta,\mathcal{B}_1} \to 0$ when $(t, \theta, p) \to (t_0, \theta_0, p_0)$, which completes the proof of Lemma 10.

Let us first introduce the real number $E_\beta := c_\beta \kappa_\beta + b_1$, where $b_1 := \sup_{\theta\in\Theta} \pi_\theta(V) < +\infty$ from (VG1) and κ_β and c_β are defined in (D–F). Then, using (D–F), $V \geq 1$ and $\beta > 0$, we obtain

$$\forall \theta \in \Theta, \quad Q_\theta V^\beta \leq E_\beta V^\beta.$$

On the other hand, let us state the following obvious inequality:

$$\forall a \in \mathbb{R}, \quad \left|\mathrm{e}^{\mathrm{i}a} - 1\right| \leq \min\bigl(2, |a|\bigr) \leq 2|a|^\alpha.$$

Now recall that $0 < \beta < 1$ and let us define $0 < \alpha \leq 1$ such that $\beta + \alpha/(m_0 + \varepsilon) \leq 1$ where $\varepsilon > 0$ is defined in (D_{m_0}).

(1) Let us define $\Delta := Q_{\theta,p}(t) - Q_{\theta,p}(t_0)$. One has for all $f \in \mathcal{B}_\beta$ and $x \in E$:

$$\begin{aligned}
\bigl|\Delta f(x)\bigr| &\leq \int_E \bigl|\mathrm{e}^{\mathrm{i}t\xi_p(x,y)} - \mathrm{e}^{\mathrm{i}t_0\xi_p(x,y)}\bigr|\bigl|f(y)\bigr| Q_\theta(x, \mathrm{d}y) \\
&\leq 2|t - t_0|^\alpha \|f\|_{\mathcal{B}_\beta} \int_E \bigl|\xi_p(x, y)\bigr|^\alpha V(y)^\beta Q_\theta(x, \mathrm{d}y) \\
&\leq 2C_\xi^{\alpha/(m_0+\varepsilon)} |t - t_0|^\alpha \|f\|_{\mathcal{B}_\beta} \int_E \bigl(V(x) + V(y)\bigr)^{\alpha/(m_0+\varepsilon)} V(y)^\beta Q_\theta(x, \mathrm{d}y) \\
&\leq 2^{1+\alpha/(m_0+\varepsilon)} C_\xi^{\alpha/(m_0+\varepsilon)} |t - t_0|^\alpha \|f\|_{\mathcal{B}_\beta} \bigl(V(x)^{\alpha/(m_0+\varepsilon)} Q_\theta V^\beta(x) + Q_\theta V^{\beta+\alpha/(m_0+\varepsilon)}(x)\bigr) \\
&\leq 2^{1+\alpha/(m_0+\varepsilon)} C_\xi^{\alpha/(m_0+\varepsilon)} |t - t_0|^\alpha \|f\|_{\mathcal{B}_\beta} (E_\beta + E_{\beta+\alpha/(m_0+\varepsilon)}) V(x)^{\beta+\alpha/(m_0+\varepsilon)}
\end{aligned}$$

from which we deduce $\|\Delta f\|_{\mathcal{B}_1} \leq D_\xi |t - t_0|^\alpha \|f\|_{\mathcal{B}_\beta}$ where D_ξ does not depend on (θ, p).

(2) In the same way, let us define $\Delta := Q_{\theta,p}(t) - Q_{\theta_0,p_0}(t)$. Let x denote some d-dimensional column vector $(x_d, \ldots, x_1)'$ and $\underline{x}_{d-1}$ denote the associated $(d-1)$-dimensional column vector $(x_d, \ldots, x_2)'$. We have for all $f \in \mathcal{B}_\beta$ and $x \in E$:

$$\begin{aligned}
\Delta f(x) &= \int_{\mathbb{R}} \exp\left(\mathrm{i}t\xi_p\left(x, \begin{pmatrix} \langle x, g(\theta)\rangle + z \\ \underline{x}_{d-1} \end{pmatrix}\right)\right) f\begin{pmatrix} \langle x, g(\theta)\rangle + z \\ \underline{x}_{d-1} \end{pmatrix} f_Z(z)\,\mathrm{d}z \\
&\quad - \int_{\mathbb{R}} \exp\left(\mathrm{i}t\xi_{p_0}\left(x, \begin{pmatrix} \langle x, g(\theta_0)\rangle + z \\ \underline{x}_{d-1} \end{pmatrix}\right)\right) f\begin{pmatrix} \langle x, g(\theta_0)\rangle + z \\ \underline{x}_{d-1} \end{pmatrix} f_Z(z)\,\mathrm{d}z \\
&= \int_{\mathbb{R}} \exp\left(\mathrm{i}t\xi_p\left(x, \begin{pmatrix} y \\ \underline{x}_{d-1} \end{pmatrix}\right)\right) f\begin{pmatrix} y \\ \underline{x}_{d-1} \end{pmatrix} f_Z\bigl(y - \langle x, g(\theta)\rangle\bigr)\,\mathrm{d}y \\
&\quad - \int_{\mathbb{R}} \exp\left(\mathrm{i}t\xi_{p_0}\left(x, \begin{pmatrix} y \\ \underline{x}_{d-1} \end{pmatrix}\right)\right) f\begin{pmatrix} y \\ \underline{x}_{d-1} \end{pmatrix} f_Z\bigl(y - \langle x, g(\theta_0)\rangle\bigr)\,\mathrm{d}y, \\
|\Delta f(x)| &\leq \left|\int_{\mathbb{R}} f\begin{pmatrix} y \\ \underline{x}_{d-1} \end{pmatrix} \bigl(f_Z\bigl(y - \langle x, g(\theta)\rangle\bigr) - f_Z\bigl(y - \langle x, g(\theta_0)\rangle\bigr)\bigr)\,\mathrm{d}y\right| \\
&\quad + 2|t|^\alpha \int_{\mathbb{R}} \left|(\xi_p - \xi_{p_0})\left(x, \begin{pmatrix} y \\ \underline{x}_{d-1} \end{pmatrix}\right)\right|^\alpha \left|f\begin{pmatrix} y \\ \underline{x}_{d-1} \end{pmatrix}\right| f_Z\bigl(y - \langle x, g(\theta_0)\rangle\bigr)\,\mathrm{d}y.
\end{aligned}$$

Thus one has $|\Delta f(x)| \le \|f\|_{\mathcal{B}_\beta}(|\theta-\theta_0|I_1 + |t|^\alpha|p-p_0|^\alpha I_2)$ where, thanks to differentiation under the integral sign and assumptions (E) and (F), I_1 satisfies for some $\widetilde{\theta} \in \mathbb{R}$ such that $|\widetilde{\theta}-\theta| \le |\widetilde{\theta}-\theta_0|$

$$\begin{aligned} I_1 &\le \sup_{\widetilde{\theta}\in\Theta}\bigl|\bigl\langle x, g^{(1)}(\widetilde{\theta})\bigr\rangle\bigr| \sup_{z\in\mathbb{R}}\left|\frac{f_Z^{(1)}(z)}{f_Z(z)}\right| \int_{\mathbb{R}} V\begin{pmatrix} y \\ \underline{x}_{d-1}\end{pmatrix}^\beta f_Z\bigl(y-\bigl\langle x, g(\widetilde{\theta})\bigr\rangle\bigr)\,\mathrm{d}y \\ &= \sup_{\widetilde{\theta}\in\Theta}\bigl|\bigl\langle x, g^{(1)}(\widetilde{\theta})\bigr\rangle\bigr| \sup_{z\in\mathbb{R}}\left|\frac{f_Z^{(1)}(z)}{f_Z(z)}\right| Q_{\widetilde{\theta}}V^\beta(x) \\ &\le E_\beta V(x)^{(1/\gamma)+\beta} \sup_{\widetilde{\theta}\in\Theta}\bigl\|g^{(1)}(\widetilde{\theta})\bigr\| \sup_{z\in\mathbb{R}}\left|\frac{f_Z^{(1)}(z)}{f_Z(z)}\right| \end{aligned}$$

and where I_2 satisfies on the other hand

$$\begin{aligned} I_2 &\le 2\int_{\mathbb{R}}\left(V(x) + V\begin{pmatrix} y \\ \underline{x}_{d-1}\end{pmatrix}\right)^{\alpha/(m_0+\varepsilon)} V\begin{pmatrix} y \\ \underline{x}_{d-1}\end{pmatrix}^\beta f_Z\bigl(y-\bigl\langle x, g(\theta_0)\bigr\rangle\bigr)\,\mathrm{d}y \\ &= 2^{1+\alpha/(m_0+\varepsilon)}\bigl(V(x)^{\alpha/(m_0+\varepsilon)}Q_{\theta_0}V^\beta(x) + Q_{\theta_0}V^{\beta+\alpha/(m_0+\varepsilon)}(x)\bigr) \\ &\le 2^{1+\alpha/(m_0+\varepsilon)}(E_\beta + E_{\beta+\alpha/(m_0+\varepsilon)})V^{\beta+\alpha/(m_0+\varepsilon)}(x). \end{aligned}$$

Since $0<\beta\le 1-1/\gamma$, one has $\|\Delta f\|_{\mathcal{B}_1} \le D'_\xi(|\theta-\theta_0| + |t|^\alpha|p-p_0|^\alpha)\|f\|_{\mathcal{B}_\beta}$. □

Appendix: Proof of Theorem 3

The investigation of the case $|u| > 2\sqrt{\ln n}$ is similar[6] to the one of [17], Section 2.2, Proposition 1, so that the details are omitted. By contrast, the case $|u|\le 2\sqrt{\ln n}$ is quite different. First let us introduce for all $\theta\in\Theta$ and $u\in\mathbb{R}$, $|u|\le 2\sqrt{\ln n}$

$$\tau = \tau_n(u,\theta) := \alpha_0 + \frac{\sigma(\theta)}{\sqrt{n}}u \quad \text{and} \quad \varsigma_n(u,\theta) := \frac{\sigma(\theta)}{\sqrt{n}}u = \tau - \alpha_0.$$

For the sake of simplicity, let us define for all $\theta\in\Theta$ and $u\in\mathbb{R}$, $|u|\le 2\sqrt{\ln n}$

$$P_{n,\theta}(u) := \mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma(\theta)}(\hat{\alpha}_n - \alpha_0) \le u\right\} = \mathbb{P}_\theta\{\hat{\alpha}_n \le \tau\}, \qquad Q_{n,\theta}(u) := \mathbb{P}_\theta\bigl\{M_n^{(1)}(\tau)\ge 0\bigr\}.$$

At a first stage we prove that

$$\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\bigl|P_{n,\theta}(u) - Q_{n,\theta}(u)\bigr| = \mathrm{o}\bigl(n^{-1/2}\bigr) \tag{43}$$

and then we determine A_θ such that

$$\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\bigl|Q_{n,\theta}(u) - \bigl(\mathcal{N}(u) + \eta(u)n^{-1/2}A_\theta(u)\bigr)\bigr| = \mathrm{o}\bigl(n^{-1/2}\bigr) \tag{44}$$

to complete the proof of Theorem 3.

Let us prove that (43) holds true. It follows from (26a) that there exists some real r.v. $\widetilde{\alpha}'_n$ taken between $\hat{\alpha}_n$ and τ such that $0 = M_n^{(1)}(\hat{\alpha}_n) = M_n^{(1)}(\tau) + (\hat{\alpha}_n - \tau)M_n^{(2)}(\widetilde{\alpha}'_n)$. Next, introducing the event $\{M_n^{(2)}(\widetilde{\alpha}'_n) > 0\}$ and its complement, one has

$$P_{n,\theta}(u) = \mathbb{P}_\theta\bigl\{M_n^{(1)}(\tau)\ge 0, M_n^{(2)}\bigl(\widetilde{\alpha}'_n\bigr) > 0\bigr\} + \mathbb{P}_\theta\bigl\{\hat{\alpha}_n \le \tau, M_n^{(2)}\bigl(\widetilde{\alpha}'_n\bigr)\le 0\bigr\},$$

[6] One difference is that our uniform consistency property (26b) is stronger than the one of [17].

so that

$$\left|P_{n,\theta}(u)-Q_{n,\theta}(u)\right| \le 2\mathbb{P}_\theta\left\{M_n^{(2)}\left(\widetilde{\alpha}_n'\right)\le 0\right\}.$$

Introducing the events $\{M_n^{(2)}(\widetilde{\alpha}_n') < M_n^{(2)}(\alpha_0)-|\widetilde{\alpha}_n'-\alpha_0|l(\theta)\}$, $\{M_n^{(2)}(\alpha_0)\le m_2(\theta)/2\}$ and their complements, where the function $l(\cdot)$ is defined in (A.1), one has

$$\mathbb{P}_\theta\left\{M_n^{(2)}\left(\widetilde{\alpha}_n'\right)\le 0\right\}\le P_1+P_2+P_3,$$

where $(P_i)_{i=1,2,3}$ denote

$$P_1 := \sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\mathbb{P}_\theta\left\{M_n^{(2)}\left(\widetilde{\alpha}_n'\right)<M_n^{(2)}(\alpha_0)-\left|\widetilde{\alpha}_n'-\alpha_0\right|l(\theta)\right\},$$

$$P_2 := \sup_{\theta\in\Theta}\mathbb{P}_\theta\left\{M_n^{(2)}(\alpha_0)\le\frac{m_2(\theta)}{2}\right\},$$

$$P_3 := \sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\left\{\frac{m_2(\theta)}{2}-\left|\widetilde{\alpha}_n'-\alpha_0\right|l(\theta)<M_n^{(2)}\left(\widetilde{\alpha}_n'\right)\le 0\right\}.$$

- Introducing the event $\{W_n\ge l(\theta)\}$ and its complement, it is easy to check from (A.1) that $P_1=\mathrm{o}(n^{-1/2})$.
- One has $P_2\le\mathbb{P}_\theta\{(\sqrt{n}/\sigma_2(\theta))(M_n^{(2)}(\alpha_0)-m_2(\theta))\le -b\sqrt{n}\}$ where $b:=\inf_{\theta\in\Theta}m_2(\theta)/2\sigma_2(\theta)$, $b>0$ from (A.2), which implies $P_2=\mathrm{o}(n^{-1/2})$.
- Introducing the event $\{|\widetilde{\alpha}_n'-\alpha_0|\ge m_2(\theta)/2l(\theta)\}$ and its complement, it is easily checked that

$$P_3\le\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\mathbb{P}_\theta\left\{\left|\widetilde{\alpha}_n'-\alpha_0\right|\ge m_2(\theta)/2l(\theta)\right\}.$$

  Furthermore $\widetilde{\alpha}_n'$ satisfies $|\widetilde{\alpha}_n'-\alpha_0|\le|\hat{\alpha}_n-\alpha_0|+|\tau-\alpha_0|$, where $\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}|\tau-\alpha_0|\to 0$ when $n\to+\infty$ (recall that $\sup_{\theta\in\Theta}\sigma(\theta)<+\infty$). Thus $P_3\le\sup_{\theta\in\Theta}\mathbb{P}_\theta\{|\hat{\alpha}_n-\alpha_0|\ge d\}$ for n sufficiently large, and where the real number d is defined by $d:=\inf_{\theta\in\Theta}m_2(\theta)/(4l(\theta))>0$, so that $P_3=\mathrm{o}(n^{-1/2})$ according to (26b).

Therefore the estimate (43) holds true.

In a second and last step, let us determine A_θ such that (44) holds true. There exists some real r.v. $\widetilde{\alpha}_n''$ taken between τ and α_0 such that

$$M_n^{(1)}(\tau)=M_n^{(1)}(\alpha_0)+\varsigma_n(u,\theta)M_n^{(2)}(\alpha_0)+\frac{\varsigma_n(u,\theta)^2}{2}M_n^{(3)}\left(\widetilde{\alpha}_n''\right).$$

Let us introduce the r.v.

$$Z_n(u,\theta):=M_n^{(1)}(\alpha_0)+\varsigma_n(u,\theta)M_n^{(2)}(\alpha_0)+\frac{\varsigma_n(u,\theta)^2}{2}M_n^{(3)}(\alpha_0),$$

the event $C_{n,\theta}:=\{W_n<l(\theta)\}$ and the positive number $c=c_n(u,\theta):=|\varsigma_n(u,\theta)|^3 l(\theta)/2$, where the r.v. W_n and the function $l(\cdot)$ are defined in (A.1).

Consider the following events

$$\begin{aligned}
\underbrace{B_{n,u,\theta}^{1-}}&:=\left\{Z_n(u,\theta)-c\ge 0\right\}, & \underbrace{B_{n,u,\theta}^{2-}}&:=\underbrace{B_{n,u,\theta}^{1-}}\cap C_{n,\theta},\\
B_{n,u,\theta}^{1}&:=\left\{M_n^{(1)}(\tau)\ge 0\right\}, & B_{n,u,\theta}^{2}&:=B_{n,u,\theta}^{1}\cap C_{n,\theta},\\
B_{n,u,\theta}^{1+}&:=\left\{Z_n(u,\theta)+c\ge 0\right\}, & B_{n,u,\theta}^{2+}&:=B_{n,u,\theta}^{1+}\cap C_{n,\theta},
\end{aligned}$$

and notice that $Q_{n,\theta}(u)=\mathbb{P}_\theta\{\widetilde{B_{n,u,\theta}^{1}}\}$ and the following facts

- since $\sup_{\theta\in\Theta}\mathbb{P}_\theta\{C^c_{n,\theta}\} = \mathrm{o}(n^{-1/2})$ from (A.1), one has $\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}|Q_{n,\theta}(u) - \mathbb{P}_\theta\{\widetilde{B^2_{n,u,\theta}}\}| = \mathrm{o}(n^{-1/2})$;
- one obviously has $M_n^{(1)}(\tau) = Z_n(u,\theta) + (\varsigma_n(u,\theta)^2/2)(M_n^{(3)}(\widetilde{\alpha}''_n) - M_n^{(3)}(\alpha_0))$, and hence from (A.1), one has

$$\forall\theta\in\Theta, \forall u\in\mathbb{R},\quad B^{2-}_{n,u,\theta}\subset \widetilde{B^2_{n,u,\theta}}\subset B^{2+}_{n,u,\theta};$$

- again since $\sup_{\theta\in\Theta}\mathbb{P}_\theta\{C^c_{n,\theta}\} = \mathrm{o}(n^{-1/2})$, one obtains $\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}|\mathbb{P}_\theta\{B^{2+}_{n,u,\theta}\} - \mathbb{P}_\theta\{B^{1+}_{n,u,\theta}\}| = \mathrm{o}(n^{-1/2})$ and $\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}|\mathbb{P}_\theta\{B^{2-}_{n,u,\theta}\} - \mathbb{P}_\theta\{B^{1-}_{n,u,\theta}\}| = \mathrm{o}(n^{-1/2})$.

Then it only remains to determine A_θ such that

$$\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\bigl|\mathbb{P}_\theta\bigl\{B^{1\pm}_{n,u,\theta}\bigr\} - \bigl(\mathcal{N}(u) + \eta(u)n^{-1/2}A_\theta(u)\bigr)\bigr| = \mathrm{o}\bigl(n^{-1/2}\bigr).$$

Let us introduce

$$\begin{aligned}\Delta_n^{\pm}(u,\theta) &:= \frac{\sqrt{n}}{\sigma_{n,u}(\theta)}\left[m_2(\theta)\varsigma_n(u,\theta) + \frac{\varsigma_n(u,\theta)^2}{2}m_3(\theta)\pm c\right] - u\\ &= u\left(\sigma_{n,u}(\theta)^{-1}\left[\sigma_1(\theta) + \sigma(\theta)\frac{m_3(\theta)}{2}\varsigma_n(u,\theta)\pm\sigma(\theta)l(\theta)\frac{\varsigma_n(u,\theta)^2}{2}\right] - 1\right)\end{aligned}$$

so that

$$\mathbb{P}_\theta\bigl\{B^{1\pm}_{n,u,\theta}\bigr\} = 1 - \mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_{n,u}(\theta)}\widetilde{M}_n(u,\theta) < -u - \Delta_n^{\pm}(u,\theta)\right\}.$$

From the last property of (A.3) applied to $v = -u - \Delta_n^{\pm}(u,\theta)$, we obtain

$$\begin{aligned}&\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\bigl|\mathbb{P}_\theta\bigl\{B^{1\pm}_{n,u,\theta}\bigr\}\\ &\quad - \mathcal{N}\bigl(u + \Delta_n^{\pm}(u,\theta)\bigr) + \eta\bigl(u + \Delta_n^{\pm}(u,\theta)\bigr)n^{-1/2}E_\theta\bigl(-u - \Delta_n^{\pm}(u,\theta)\bigr)\bigr| = \mathrm{o}\bigl(n^{-1/2}\bigr).\end{aligned}$$

From the first property of (A.3), both $\Delta_n^+(u,\theta)$ and $\Delta_n^-(u,\theta)$ admit the following expansion:

$$\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\left|\Delta_n^{\pm}(u,\theta) - \left(\sigma_1(\theta)D_\theta(u) + \frac{\sigma(\theta)^2}{2\sigma_1(\theta)}m_3(\theta)u\right)un^{-1/2}\right| = \mathrm{o}\bigl(n^{-1/2}\bigr),$$

and hence

$$\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\bigl|\mathbb{P}_\theta\bigl\{B^{1+}_{n,u,\theta}\bigr\} - \mathbb{P}_\theta\bigl\{B^{1-}_{n,u,\theta}\bigr\}\bigr| = \mathrm{o}\bigl(n^{-1/2}\bigr).$$

Finally we define the polynomial function A_θ by (31); (30) follows from:

$$\begin{aligned}&\sup_{\theta\in\Theta}\sup_{|u|\le 2\sqrt{\ln n}}\bigl|\bigl(\mathcal{N}(u) + \eta(u)n^{-1/2}A_\theta(u)\bigr)\\ &\quad - \bigl(\mathcal{N}\bigl(u + \Delta_n^{+}(u,\theta)\bigr) - \eta\bigl(u + \Delta_n^{+}(u,\theta)\bigr)n^{-1/2}E_\theta\bigl(-u - \Delta_n^{+}(u,\theta)\bigr)\bigr)\bigr| = \mathrm{o}\bigl(n^{-1/2}\bigr).\end{aligned}$$ □

References

[1] V. Baladi. *Positive Transfer Operators and Decay of Correlations*. *Adv. Ser. Nonlinear Dynam.* **16**. World Scientific Publishing Co., River Edge, NJ, 2000. MR1793194

[2] P. Bertail and S. Clémençon. A renewal approach to markovian U-statistics. *Math. Methods Statist.* **20** (2011) 79–105. MR2882153
[3] P. Billingsley. *Statistical Inference for Markov Processes*. Univ. Chicago Press, Chicago, IL, 1961. MR0123419
[4] E. Bolthausen. The Berry–Esseen theorem for strongly mixing Harris recurrent Markov chains. *Probab. Theory Related Fields* **60** (1982) 283–289. MR0664418
[5] R. Dahlhaus. Efficient parameter estimation for self similar processes. *Ann. Statist.* **17** (1989) 1749–1766. MR1026311
[6] D. Dehay and J. F. Yao. On likelihood estimation for discretely observed Markov jump processes. *Aust. N. Z. J. Stat.* **49** (2007) 93–107. MR2345413
[7] J. Durbin. Approximations for densities of sufficient estimators. *Biometrika* **67** (1980) 311–333. MR0581728
[8] W. Feller. *An Introduction to Probability Theory and Its Applications. Vol. II*. Wiley, New York, 1971. MR0270403
[9] D. Ferré. Développements d'Edgeworth en statistique des modèles markoviens. Ph.D. thesis, Institut National des Sciences Appliquées de Rennes.
[10] D. Ferré. Développement d'Edgeworth d'ordre 1 pour des M-estimateurs dans le cas de chaînes V-géométriquement ergodiques. *C. R. Math. Acad. Sci. Paris* **348** (2010) 331–334. MR2600134
[11] D. Ferré and D. Guibourg. Regularity of the characteristic function of additive functionals for iterated function systems. Statistical applications. *Markov Process. Related Fields* **19** (2013) 299–342. MR3113946
[12] D. Ferré, L. Hervé and J. Ledoux. Limit theorems for stationary Markov processes with L^2-spectral gap. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** (2012) 396–423. MR2954261
[13] C.-D. Fuh. Efficient likelihood estimation in state space models. *Ann. Statist.* **34** (2006) 2026–2068. MR2283726
[14] P. Gänssler. Note on minimum contrast estimators for Markov processes. *Metrika* **19** (1972) 115–130. MR0418374
[15] F. Götze and C. Hipp. Asymptotic distribution of statistics in time series. *Ann. Statist.* **22** (1994) 2062–2088. MR1329183
[16] H. Hennion and L. Hervé. *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness*. *Lecture Notes Math.* **1766**. Springer, Berlin, 2001. MR1862393
[17] L. Hervé, J. Ledoux and V. Patilea. A Berry-Esseen theorem on M-estimators for V-geometrically ergodic Markov chains. *Bernoulli* **18** (2012) 703–734. MR2922467
[18] L. Hervé and F. Pène. The Nagaev–Guivarc'h method via the Keller–Liverani theorem. *Bull. Soc. Math. France* **138** (3) (2010) 415–489. MR2729019
[19] J. L. Jensen. Asymptotic expansions for strongly mixing Harris recurrent Markov chains. *Scand. J. Statist.* **16** (1989) 47–63. MR1003968
[20] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (4)* **XXVIII** (1999) 141–152. MR1679080
[21] O. Lieberman, J. Rousseau and D. M. Zucker. Valid asymptotic expansions for the maximum likelihood estimator of the parameter of a stationary, Gaussian, strongly dependent process. *Ann. Statist.* **31** (2003) 586–612. MR1983543
[22] C. Liverani. Invariant measures and their properties. A functional analytic point of view. In *Dynamical Systems*. *Part II* 185–237. Cent. Ric. Mat. Ennio Giorgi. Scuola Norm. Sup., Pisa, 2003. MR2071241
[23] V. K. Malinovskii. Limit theorems for Harris Markov chains. I. *Theory Probab. Appl.* **31** (1987) 269–285. MR0850991
[24] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer, Berlin, 1993. MR1287609
[25] J. Pfanzagl. Asymptotic expansions related to minimum contrast estimators. *Ann. Statist.* **1** (6) (1973) 993–1026. MR0359151
[26] B. L. S. Prakasa Rao. Maximum likelihood estimation for Markov processes. *Ann. Inst. Statist. Math.* **24** (1972) 333–345. MR0336936
[27] F. Räbiger and M. P. H. Wolff. On the approximation of positive operators and the behaviour of the spectra of the approximants. *Integral Equations Operator Theory* **28** (1) (1997) 72–86. MR1446832
[28] G. G. Roussas. Asymptotic inferencer in Markov process. *Ann. Math. Statist.* **36** (1965) 978–982. MR0179882
[29] W. L. Smith. Regenerative stochastic processes. *Proc. Roy. Soc. London. Ser. A.* **232** (1955) 6–31. MR0073877