Bernoulli **18**(2), 2012, 703–734
DOI: 10.3150/10-BEJ347

A uniform Berry–Esseen theorem on M-estimators for geometrically ergodic Markov chains

LOÏC HERVÉ[1,*], JAMES LEDOUX[1,**] and VALENTIN PATILEA[1,2]

[1] *Université Européenne de Bretagne, INSA-IRMAR, UMR-CNRS 6625. Institut National des Sciences Appliquées de Rennes, 20, Avenue des Buttes de Coësmes CS 14315, 35043 Rennes cedex, France.*
E-mail: [*] *loic.herve@insa-rennes.fr;* [**] *james.ledoux@insa-rennes.fr*
[2] *CREST (Ensai), Campus de Ker-Lann, Rue Blaise Pascal – BP 37203, 35172 Bruz cedex, France.*
E-mail: valentin.patilea@insa-rennes.fr

Let $\{X_n\}_{n\geq 0}$ be a V-geometrically ergodic Markov chain. Given some real-valued functional F, define $M_n(\alpha) := n^{-1}\sum_{k=1}^n F(\alpha, X_{k-1}, X_k)$, $\alpha \in \mathcal{A} \subset \mathbb{R}$. Consider an M estimator $\widehat{\alpha}_n$, that is, a measurable function of the observations satisfying $M_n(\widehat{\alpha}_n) \leq \min_{\alpha\in\mathcal{A}} M_n(\alpha) + c_n$ with $\{c_n\}_{n\geq 1}$ some sequence of real numbers going to zero. Under some standard regularity and moment assumptions, close to those of the i.i.d. case, the estimator $\widehat{\alpha}_n$ satisfies a Berry–Esseen theorem uniformly with respect to the underlying probability distribution of the Markov chain.

Keywords: asymptotic properties of estimators; Markov chains; weak spectral method

1. Introduction

Let $(E, \mathcal{E})$ be a measurable space with $\mathcal{E}$ a countably generated σ-field, and let $\{X_n\}_{n\geq 0}$ be a Markov chain with state space E and transition kernels $\{Q_\theta(x, \cdot)\colon x \in E\}$, where θ is a parameter in some general set Θ. The initial distribution of the chain, that is, the probability distribution of X_0, is denoted by μ and may or may not depend on θ. Although $\{X_n\}_{n\geq 0}$ does not need to be the canonical version, we use the standard notation $\mathbb{P}_{\theta,\mu}$ to refer to the probability distribution of $\{X_n\}_{n\geq 0}$ (and $\mathbb{E}_{\theta,\mu}$ for the expectation w.r.t. $\mathbb{P}_{\theta,\mu}$). We consider that $\{X_n\}_{n\geq 0}$ is a V-geometrically ergodic Markov chain, where $V : E \to [1, +\infty)$ is some fixed *unbounded* function. This class of Markov chains is large enough to cover interesting applications (see [16], Sections 16.4 and 16.5).

The parameter of interest is $\alpha_0 = \alpha_0(\theta) \subset \mathcal{A}$, where $\alpha_0(\cdot)$ is a function of the parameter θ and $\mathcal{A}$ is an open interval of $\mathbb{R}$. To estimate α_0, let us introduce the statistic

$$M_n(\alpha) := \frac{1}{n}\sum_{k=1}^n F(\alpha, X_{k-1}, X_k), \tag{1}$$

where F is a real-valued measurable functional on $\mathcal{A} \times E^2$. We define an M-*estimator* (this is slightly more general than the usual definition of M-estimators or minimum contrast estimators, where $c_n = 0$, see [1]) to be a random variable $\widehat{\alpha}_n$ depending on the observations $(X_0, \ldots, X_n)$ such that

$$M_n(\widehat{\alpha}_n) \le \min_{\alpha \in \mathcal{A}} M_n(\alpha) + c_n,$$

where $\{c_n\}_{n\ge 1}$ is a sequence of non-negative real numbers going to zero to be specified later. Assume that for all $\theta \in \Theta$

$$M_\theta(\alpha) := \lim_{n\to\infty} \mathbb{E}_{\theta,\mu}[M_n(\alpha)]$$

is well defined everywhere on $\mathcal{A}$ and does not depend on μ. In addition, assume that there exists a unique "true" value α_0 of the parameter of interest, that is, $M_\theta(\alpha_0) < M_\theta(\alpha)$, $\forall \alpha \ne \alpha_0$. We want to prove the following *uniform* Berry–Esseen bound for $\widehat{\alpha}_n$

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\mu}\left\{\frac{\sqrt{n}}{\tau(\theta)}(\widehat{\alpha}_n - \alpha_0) \le u\right\} - \Gamma(u)\right| = \mathrm{O}\left(\frac{1}{\sqrt{n}}\right), \tag{BE}$$

where Γ denotes the standard normal distribution function, and $\tau(\theta)$ is some positive real number defined in Theorem 3.

To derive (BE), we use Pfanzagl's approach [20]. Besides technical assumptions, this approach relies on several ingredients. First, we need the uniform consistency condition:

(UC) $\forall d > 0$, $\sup_{\theta\in\Theta} \mathbb{P}_{\theta,\mu}\{|\widehat{\alpha}_n - \alpha_0| \ge d\} = \mathrm{O}(1/\sqrt{n})$.

Second, consider the following two convergence properties: If $S_n(\alpha_0) := \sum_{k=1}^n \xi(\alpha_0, X_{k-1}, X_k)$ with $\xi(\alpha_0, X_{k-1}, X_k)$ centered,

(a) *the sequence* $\{\mathbb{E}_{\theta,\mu}[S_n^2(\alpha_0)]/n\}_{n\ge1}$ *converges to a real number* $\sigma^2(\theta)$;
(b) *there exists a positive constant* $B(\xi)$ *such that for any* $n \ge 1$

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\mu}\left\{\frac{S_n(\alpha_0)}{\sigma(\theta)\sqrt{n}} \le u\right\} - \Gamma(u)\right| \le \frac{B(\xi)}{\sqrt{n}}.$$

The properties (a) and (b) will be required for certain $\xi(\alpha_0, x, y)$ defined as linear combinations of some functionals related to F. To obtain (a) and (b) for such $\xi(\alpha_0, x, y)$ with V-geometrically ergodic Markov chains, a natural moment (or V-domination) condition is used: There exist positive constants C_ξ and m such that

$$\forall (x, y) \in E^2, \forall \alpha \in \mathcal{A} \qquad |\xi(\alpha, x, y)|^m \le C_\xi\bigl(V(x) + V(y)\bigr). \tag{2}$$

The paper is organized as follows. In Section 2, an extended version of Pfanzagl's theorem [20], is stated for any sequence of observations, not necessarily Markovian. Section 3 is devoted to a Berry–Esseen bound for the additive functional $\sum_{k=1}^n \xi(\alpha_0, X_{k-1}, X_k)$ of a V-geometrically ergodic Markov chain $\{X_n\}_{n\ge0}$ with ξ satisfying inequality (2). In Section 3.2, we prove that the properties (a) and (b) are fulfilled when inequality (2) holds with the (almost expected) order m, namely: $m > 2 \Rightarrow$ (a), and $m > 3 \Rightarrow$ (b). These results follow from the weak spectral method based on the theorem of Keller and Liverani [14]. This approach, introduced in [10], is fully described in [12] in the Markov context (see also [8,9] and other references given in [12]). It is important to notice that Pfanzagl's method requires the precise control of the constant $B(\xi)$ in property (b) as a function of the size of ξ. The present operator-type approach shows that $B(\xi)$ depends only on the constant C_ξ in inequality (2). Thanks to these preliminary results, in Section 4 we prove our main statement, that is:

(R) *Under some technical assumptions and the uniform consistency condition* (UC)*, if two functionals F' and F'' related to F (in the basic case F' and F'' are the first- and second-order derivatives of F with respect to α) satisfy inequality* (2) *for some $m > 3$ and constants $C_{F'}$, $C_{F''}$ that do not depend on α, then $\widehat{\alpha}_n$ satisfies property* (BE).

To the best of our knowledge, the result (R) is new. It completes the central limit theorem for $\{\widehat{\alpha}_n\}_{n\geq 1}$ proved in [5] when inequality (2) holds with $m = 2$. The domination condition (2) required by (R) is almost optimal in the sense that we impose $m > 3$ in place of the best possible value $m = 3$ obtained in the i.i.d. case. In Section 5, our results are applied to the AR(1) process with ARCH (autoregressive conditional heteroscedastic) of order-1 errors. The paper ends with a conclusion section.

Let us close the Introduction with a brief review of previous related works in the literature. In [20], $\{X_n\}_{n\in\mathbb{N}}$ is a sequence of i.i.d. random variables and Pfanzagl proved a Berry–Esseen theorem for minimum contrast estimators (which are special instances of M-estimators) associated with functionals of the form $F(\alpha, X_k)$. In [20], the moment conditions on $F' := \partial F/\partial\alpha$, $F'' := \partial^2 F/\partial\alpha^2$ are the expected ones since the property (b) is fulfilled under the expected third moment condition [6], Chapter XVI. Using convexity arguments, Bentkus *et al.* [2] proposed an alternative method for deriving Berry–Esseen bounds for M-estimators with i.i.d. data. In the Markov context, the method proposed by Pfanzagl is extended, first by Rao to cover the case of uniformly ergodic Markov chains [21], second in [19] to the case of the linear autoregressive model. However, their assumptions to get the property (BE) include much stronger moment conditions involving both the functional F and the Markov chain. Here, as already mentioned, the weak spectral method of [12] enables us to have an (almost) optimal treatment of (a) and (b), and hence an improved Berry–Esseen result (BE).

2. The Pfanzagl method revisited

We state and prove a general result that allows us to derive uniform Berry–Esseen bounds for M-estimators. This result is an extended version of Theorem 1 in [20] and is applied to our Markov context in Section 4.

2.1. The result

Consider a statistical model $(\Omega, \mathcal{F}, \{\mathbb{P}_\theta, \theta \in \Theta\})$, where Θ denotes some parameter space, and let $\{X_n\}_{n\geq 0}$ be any sequence of observations (not necessarily Markovian). Let us denote the expectation with respect to $\mathbb{P}_\theta$ by $\mathbb{E}_\theta$.

For each n, let $M_n(\alpha)$ be a measurable functional of the observations $X_0, \dots, X_n$ and the parameter of interest $\alpha \in \mathcal{A}$, where $\mathcal{A}$ is some open interval of $\mathbb{R}$. Let $\{c_n\}_{n\geq 1}$ be a sequence of non-negative real numbers going to zero at some rate to be specified later. An *M-estimator* is a measurable function $\widehat{\alpha}_n$ of the observations $(X_0, \dots, X_n)$ such that

$$M_n(\widehat{\alpha}_n) \le \min_{\alpha \in \mathcal{A}} M_n(\alpha) + c_n. \tag{3}$$

This is the usual definition of minimum contrast estimators as soon as $c_n \equiv 0$.

Assumptions. *Suppose that for all $n \ge 1$ and $\alpha \in \mathcal{A}$, there exist $M'_n(\alpha)$, $M''_n(\alpha)$ some measurable functions depending on $X_0, X_1, \dots, X_n$ and on the parameter of interest, such that the following properties hold true*:

(A1) $\forall \theta \in \Theta$, *there exists a unique $\alpha_0 = \alpha_0(\theta) \in \mathcal{A}$ such that $M'_\theta(\alpha_0) = 0$, where $M'_\theta(\alpha) := \lim_{n\to\infty} \mathbb{E}_\theta[M'_n(\alpha)]$ (the limit is assumed to be well defined for all $(\theta, \alpha) \in \Theta \times \mathcal{A}$)*;

(A2) $0 < \inf_{\theta\in\Theta} m(\theta) \le \sup_{\theta\in\Theta} m(\theta) < \infty$, *where $m(\theta) := \lim_{n\to\infty} \mathbb{E}_\theta[M''_n(\alpha_0)]$ (the limit is assumed to be well defined for all θ)*;

(A3) *for every $n \ge 1$, there exists $r_n > 0$ independent of θ such that $r_n = \mathrm{o}(n^{-1/2})$ and*

$$\sup_{\theta\in\Theta} \mathbb{P}_\theta\{|M'_n(\widehat{\alpha}_n)| \ge r_n\} = \mathrm{O}(n^{-1/2});$$

(A4) *for $j = 1, 2$, there exists a function $\sigma_j(\cdot)$ such that $0 < \inf_{\theta\in\Theta} \sigma_j(\theta) \le \sup_{\theta\in\Theta} \sigma_j(\theta) < \infty$ and there exists a positive constant B such that for all $n \ge 1$*

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_1(\theta)} M'_n(\alpha_0) \le u\right\} - \Gamma(u)\right| \le \frac{B}{\sqrt{n}},$$

$$\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_2(\theta)}\big(M''_n(\alpha_0) - m(\theta)\big) \le u\right\} - \Gamma(u)\right| \le \frac{B}{\sqrt{n}};$$

(A4′) *for $n \ge 1$, $|u| \le 2\sqrt{\ln n}$ and $\theta \in \Theta$, there is a positive number $\sigma_{n,u}(\theta)$ such that*

$$|\sigma_{n,u}(\theta) - \sigma_1(\theta)| \le A' \frac{|u|}{\sqrt{n}},$$

$$\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_{n,u}(\theta)}\left(M'_n(\alpha_0) + \frac{u\sigma_1(\theta)}{\sqrt{n}m(\theta)}\big(M''_n(\alpha_0) - m(\theta)\big)\right) \le u\right\} - \Gamma(u)\right| \le \frac{B'}{\sqrt{n}}$$

with some positive constants A', B' independent of n, u, θ;

(A5) *for any $(\alpha, \tilde{\alpha}) \in \mathcal{A}^2$, let $R_n(\alpha, \tilde{\alpha})$ be defined by the equation*

$$M'_n(\tilde{\alpha}) = M'_n(\alpha) + [M''_n(\alpha) + R_n(\alpha, \tilde{\alpha})](\tilde{\alpha} - \alpha).$$

For each n, there exist $\omega_n \ge 0$ and a real-valued measurable function W_n depending on $X_0, \dots, X_n$, both independent of θ, such that $\omega_n = \mathrm{o}(1)$ and

$$\forall (\alpha, \tilde{\alpha}) \in \mathcal{A}^2 \qquad |R_n(\alpha, \tilde{\alpha})| \le \{|\alpha - \tilde{\alpha}| + \omega_n\} W_n,$$

and there is a constant $c_W > 0$ *such that*

$$\sup_{\theta\in\Theta} \mathbb{P}_\theta\{c_W \le W_n\} = \mathrm{O}(n^{-1/2});$$

(A6) $\widehat{\alpha}_n$ *is assumed to be uniformly consistent*, *that is*, *there exists* $\gamma_n = \mathrm{o}(1)$ *such that*

$$\sup_{\theta\in\Theta} \mathbb{P}_\theta\{|\widehat{\alpha}_n - \alpha_0| \ge d\} \le \gamma_n,$$

where $d := \inf_{\theta\in\Theta} m(\theta)/8c_W$ *with* c_W *and* $m(\theta)$ *defined in* (A5) *and* (A2), *respectively.*

Let us comment on these assumptions. Condition (A1) identifies the true value of the parameter. In conditions (A1) and (A2), the expectations $\mathbb{E}_\theta[M_n'(\alpha)]$ and $\mathbb{E}_\theta[M_n''(\alpha_0)]$ may depend on n, as in the Markovian framework considered in the sequel when the initial distribution is not the stationary distribution. Condition (A3) ensures that the estimator (approximately) satisfies a kind of first-order condition. Such a condition allows us to take into account the numerical errors with which we are faced when computing $\widehat{\alpha}_n$. It may also be useful when the estimator of the parameter α_0 depends on some "nuisance" parameters (see the example in the second part of Section 5). Conditions (A4) and (A4′) are the uniform Berry–Esseen bounds for $M_n'(\alpha_0)$, $M_n''(\alpha_0)$ and for some of their linear combinations. The identity defining $R_n(\alpha,\tilde{\alpha})$ in condition (A5) is guaranteed by a Taylor expansion when the criterion $M_n(\alpha)$ is twice differentiable with respect to α. In this case M_n' and M_n'' are nothing else but the first- and second-order derivatives of M_n with respect to α. The reminder $R_n(\alpha,\tilde{\alpha})$ must satisfy a Lipschitz condition. For instance, when $\omega_n = 0$, this holds true if $\alpha \mapsto M_n(\alpha)$ is three times continuously differentiable with a bounded third-order derivative. Condition (A6) is a standard consistency condition (see [2]). General sufficient conditions for (A6) with $\gamma_n = \mathrm{O}(n^{-1})$ have been proposed in the case of i.i.d. observations or uniformly ergodic Markov chains (see [18], Lemma 4, and [21], Lemma 4.1, resp.). Such general arguments can easily be adapted to the geometrically ergodic Markov chain framework. In specific examples, like the one investigated in Section 5, condition (A6) can be checked by direct arguments.

The proof of Theorem 1, which adapts the arguments of [20], is given in Section 2.2.

Theorem 1. *Under conditions* (A1)–(A6), *there exists a positive constant* C *such that*

$$\forall n \ge 1 \qquad \sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}} \left| \mathbb{P}_\theta\left\{ \frac{\sqrt{n}}{\tau(\theta)}(\widehat{\alpha}_n - \alpha_0) \le u \right\} - \Gamma(u) \right| \le C\left(\frac{1}{\sqrt{n}} + \sqrt{n} r_n + \omega_n + \gamma_n \right) \tag{4}$$

with $\tau(\theta) := \sigma_1(\theta)/m(\theta)$.

To obtain the classical order $\mathrm{O}(n^{-1/2})$ of the Berry–Esseen bound, one needs $\gamma_n = \mathrm{O}(n^{-1/2})$, $r_n = \mathrm{O}(n^{-1})$ and $\omega_n = \mathrm{O}(n^{-1/2})$. Note that this usually requires that the sequence $\{c_n\}_{n\ge1}$ in (3) decreases at the rate $n^{-3/2}$. This is to be compared to the rate n^{-1} that is usually required to obtain the asymptotic normality of M-estimators (see [1]).

Remark 1. A close inspection of the proof of Theorem 1 below shows that the constant C in inequality (4) can be tracked provided that the $\mathrm{O}(\cdot)$ and $\mathrm{o}(\cdot)$ rates in assumptions (A3)–(A6) are more explicit. For the sake of brevity, we only consider the case where $c_n = r_n = \omega_n = 0$, $\alpha(\theta) = \theta$ and (A3) is: for any $n \geq 1$, $|M_n'(\widehat{\theta}_n)| = 0$. The constants C in the various inequalities of assumptions (A4)–(A6) are denoted by C_1, C_2 in (A4), C_3, C_4 in (A4′) and C_5 in (A5) and we choose $\gamma_n \leq C_6 n^{-1/2}$ in (A6). Then we can obtain from Propositions 1 and 2 that

$$\forall n \geq 1 \qquad \sup_{\theta\in\Theta}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\tau(\theta)}(\widehat{\alpha}_n - \alpha_0) \leq u\right\} - \Gamma(u)\right| \leq \frac{C}{\sqrt{n}},$$

where $C := \frac{1}{2} + \frac{1}{\sqrt{2\pi}} + 2C_1 + 2C_2 + \frac{\exp(-a^2/2)}{a} + C_5 + C_6$ when $|u| \geq 2\sqrt{\ln n}$; or $C := 2[\frac{1}{\sqrt{2\pi}} + 2C_1 + 4C_2 + 2\frac{\exp(-a^2/2)}{a} + 2C_5 + C_6] + C_4 + \frac{16\mathrm{e}^{-1}(C_3 + \overline{\sigma}^2 c_W)}{\underline{\sigma}_1\sqrt{2\pi}}$ when $|u| < 2\sqrt{\ln n}$ provided that $\sqrt{n/\ln n} \geq \max(8c_W\overline{\sigma}^2, 4)/\underline{\sigma}_1$; with $a := \inf_{\theta\in\Theta}(m(\theta)/4\sigma_2(\theta))$, $\overline{\sigma} := \sup_{\theta\in\Theta}\sigma_1(\theta)/m(\theta)$, $\underline{\sigma}_1 := \inf_{\theta\in\Theta}\sigma_1(\theta)$.

2.2. Proof of Theorem 1

The hypotheses of Theorem 1 are assumed to hold. For the sake of brevity, the sequence $\{r_n\}_{n\geq 1}$ in (A3) is supposed to be such that $r_n = \mathrm{o}(n^{-1/2})$ and $|M_n'(\widehat{\alpha}_n)| \leq r_n$ for every $n \geq 1$. In the general case, it suffices to work on the event $\{|M_n'(\widehat{\alpha}_n)| \leq r_n\}$ and to bound the probability of the event $\{|M_n'(\widehat{\alpha}_n)| > r_n\}$ using (A3). From conditions (A2) and (A4),

$$\tau(\theta) := \frac{\sigma_1(\theta)}{m(\theta)}, \qquad \underline{m} := \inf_{\theta\in\Theta} m(\theta), \qquad \overline{m} := \sup_{\theta\in\Theta} m(\theta),$$
$$\underline{\sigma}_j := \inf_{\theta\in\Theta}\sigma_j(\theta), \qquad \overline{\sigma}_j := \sup_{\theta\in\Theta}\sigma_j(\theta),$$

$j = 1, 2$, are well defined. Recall that $0 < \underline{m} \leq \overline{m} < \infty$ and $0 < \underline{\sigma}_j \leq \overline{\sigma}_j < \infty$. Note that the function $\tau(\cdot)$ is positive and bounded. In the following, C denotes a positive constant whose value may be different from line to line.

Inequality (4) is proved, first for $|u| \geq 2\sqrt{\ln n}$, second for $|u| < 2\sqrt{\ln n}$. In fact, for $|u| \geq 2\sqrt{\ln n}$, the bound in inequality (4) does not involve r_n and ω_n.

Proposition 1. *There exists a positive constant C such that for each $n \geq 1$ and all $u \in \mathbb{R}$ such that $|u| \geq 2\sqrt{\ln n}$*

$$\sup_{\theta\in\Theta}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\tau(\theta)}(\widehat{\alpha}_n - \alpha_0) \leq u\right\} - \Gamma(u)\right| \leq \frac{C}{\sqrt{n}} + \gamma_n. \tag{5}$$

Proof. For $|u| \geq 2\sqrt{\ln n}$, it is easily checked that

$$\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\tau(\theta)}(\widehat{\alpha}_n - \alpha_0) \leq u\right\} - \Gamma(u)\right| \leq \mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\tau(\theta)}|\widehat{\alpha}_n - \alpha_0| \geq 2\sqrt{\ln n}\right\} + \Gamma\bigl(-2\sqrt{\ln n}\bigr).$$

Now,

$$\Gamma\big(-2\sqrt{\ln n}\big) \le \frac{1}{2\sqrt{\ln n}} \frac{1}{\sqrt{2\pi}} \int_{2\sqrt{\ln n}}^{+\infty} v e^{-v^2/2}\,dv = \frac{1}{2\sqrt{\ln n}} \frac{1}{\sqrt{2\pi}} \frac{1}{n^2}.$$

Finally, the proof is complete if there exists $C > 0$ such that (see [18], Lemma 6)

$$\forall n \ge 1 \qquad \sup_{\theta\in\Theta} \mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\tau(\theta)}|\widehat{\alpha}_n - \alpha_0| > 2\sqrt{\ln n}\right\} \le \frac{C}{\sqrt{n}} + \gamma_n. \tag{6}$$

It follows from (A5) and (A3) that $|M_n'(\alpha_0)| + r_n \ge |\widehat{\alpha}_n - \alpha_0||M_n''(\alpha_0) + R_n(\widehat{\alpha}_n, \alpha_0)|$. Then,

$$\frac{\sqrt{n}}{\sigma_1(\theta)}|\widehat{\alpha}_n - \alpha_0| > 2\frac{\sqrt{\ln n}}{m(\theta)} \Longrightarrow \frac{\sqrt{n}}{\sigma_1(\theta)}\big(|M_n'(\alpha_0)| + r_n\big) > 2\frac{\sqrt{\ln n}}{m(\theta)}|M_n''(\alpha_0) + R_n(\widehat{\alpha}_n, \alpha_0)|,$$

provided that $M_n'(\widehat{\alpha}_n) \ne M_n'(\alpha_0)$. Next, introducing the event $\{2|M_n''(\alpha_0) + R_n(\widehat{\alpha}_n, \alpha_0)| > m(\theta)\}$ and its complement (which includes the event $\{M_n'(\widehat{\alpha}_n) = M_n'(\alpha_0)\}$), we obtain

$$\begin{aligned}
&\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\tau(\theta)}|\widehat{\alpha}_n - \alpha_0| > 2\sqrt{\ln n}\right\} \\
&\quad \le \mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_1(\theta)}\{|M_n'(\alpha_0)| + r_n\} > \sqrt{\ln n}\right\} + \mathbb{P}_\theta\{2|M_n''(\alpha_0) + R_n(\widehat{\alpha}_n, \alpha_0)| \le m(\theta)\}.
\end{aligned}$$

It is easily checked from (A4) and $r_n = o(n^{-1/2})$ that

$$\sup_{\theta\in\Theta} \mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\sigma_1(\theta)}\{|M_n'(\alpha_0)| + r_n\} > \sqrt{\ln n}\right\} = O\left(\frac{1}{\sqrt{n}}\right) + 2\Gamma\left(-\sqrt{\ln n} + \frac{\sqrt{n} r_n}{\sigma_1(\theta)}\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

Finally, to obtain the bound (6), it remains to justify the use of the following bound:

$$\sup_{\theta\in\Theta} \mathbb{P}_\theta\{2|M_n''(\alpha_0) + R_n(\widehat{\alpha}_n, \alpha_0)| \le m(\theta)\} = O(n^{-1/2}) + \gamma_n. \tag{7}$$

Using elementary inequalities and assumption (A5),

$$\begin{aligned}
&\mathbb{P}_\theta\{2|M_n''(\alpha_0) + R_n(\widehat{\alpha}_n, \alpha_0)| \le m(\theta)\} \\
&\quad \le \mathbb{P}_\theta\{|M_n''(\alpha_0) - m(\theta)| \ge m(\theta)/4\} + \mathbb{P}_\theta\{|R_n(\widehat{\alpha}_n, \alpha_0)| \ge m(\theta)/4\} \\
&\quad \le \mathbb{P}_\theta\{|M_n''(\alpha_0) - m(\theta)| \ge m(\theta)/4\} + \mathbb{P}_\theta\{[|\widehat{\alpha}_n - \alpha_0)| + \omega_n]W_n \ge m(\theta)/4\} \\
&\quad =: P_{1,n,\theta} + P_{2,n,\theta}.
\end{aligned}$$

It follows from (A4) that $a := \inf_{\theta\in\Theta}(m(\theta)/4\sigma_2(\theta))$ is well defined and positive, and

$$\sup_{\theta\in\Theta} P_{1,n,\theta} \le O(n^{-1/2}) + 2\Gamma\big(-a\sqrt{n}\big) = O(n^{-/1/2}). \tag{8}$$

Now, let $d(\theta) := m(\theta)/4c_W$ with c_W defined in (A5) and notice that $d = \inf_{\theta' \in \Theta} d(\theta')/2$ in (A6). Use the event $\{|\widehat{\alpha}_n - \alpha_0| \le d(\theta) - \omega_n\}$ and its complement to write

$$\begin{aligned} P_{2,n,\theta} &\le \mathbb{P}_\theta\left\{\frac{m(\theta)}{4} \le [|\widehat{\alpha}_n - \alpha_0| + \omega_n]W_n \le W_n d(\theta)\right\} + \mathbb{P}_\theta\{|\widehat{\alpha}_n - \alpha_0| > d(\theta) - \omega_n\} \\ &\le \sup_{\theta\in\Theta} \mathbb{P}_\theta\{c_W \le W_n\} + \sup_{\theta\in\Theta} \mathbb{P}_\theta\{|\widehat{\alpha}_n - \alpha_0| > d\} = \mathrm{O}(n^{-1/2}) + \gamma_n, \end{aligned}$$

from (A5)–(A6) and provided that $\omega_n \le d$. Therefore, inequality (7) holds true. □

Now, it remains to investigate the case $|u| < 2\sqrt{\ln n}$.

Proposition 2. *There exists a positive constant C such that, for any $|u| < 2\sqrt{\ln n}$,*

$$\sup_{\theta\in\Theta}\left|\mathbb{P}_\theta\left\{\frac{\sqrt{n}}{\tau(\theta)}(\widehat{\alpha}_n - \alpha_0) \le u\right\} - \Gamma(u)\right| \le C\left(\frac{1}{\sqrt{n}} + \sqrt{n}r_n + \omega_n + \gamma_n\right). \tag{9}$$

Proof. We just have to prove that (9) holds true for all $n \ge n_0$, for some $n_0 \in \mathbb{N}$. Let us introduce some sets and derive their probability bounds:

- $E_{n,\theta} := \{\sqrt{n}|\widehat{\alpha}_n - \alpha_0|/\tau(\theta) \le 2\sqrt{\ln n}\}$. From (6), $\sup_{\theta\in\Theta} \mathbb{P}_\theta(E^c_{n,\theta}) = \mathrm{O}(n^{-1/2} + \gamma_n)$.
- $A_n := \{0 \le W_n \le c_W\}$ where the r.v. W_n and the constant c_W are defined in (A5). Then $\sup_{\theta\in\Theta} \mathbb{P}_\theta(A^c_n) = \mathrm{O}(n^{-1/2})$.
- $D_{n,\theta} := \{2M''_n(\alpha_0) > m(\theta)\}$. We have $\mathbb{P}_\theta\{D^c_{n,\theta}\} \le \mathbb{P}_\theta\{|M''_n(\alpha_0) - m(\theta)| \ge m(\theta)/2\} \le \mathbb{P}_\theta\{|M''_n(\alpha_0) - m(\theta)| \ge m(\theta)/4\}$. We know from (8) that $\sup_{\theta\in\Theta} \mathbb{P}_\theta(D^c_{n,\theta}) = \mathrm{O}(n^{-1/2})$.

Then, we obtain from the previous estimates that the following set

$$B_{n,\theta} := E_{n,\theta} \cap A_n \cap D_{n,\theta}$$

is such that

$$\sup_{\theta\in\Theta} \mathbb{P}_\theta(B^c_{n,\theta}) \le \mathrm{O}(n^{-1/2} + \gamma_n). \tag{10}$$

Now, if $D_{n,\theta,u} := \{\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta) \le u\}$, then we can write from (10)

$$|\mathbb{P}_\theta(D_{n,\theta,u}) - \Gamma(u)| \le |\mathbb{P}_\theta(D_{n,\theta,u} \cap B_{n,\theta}) - \Gamma(u)| + \mathrm{O}(n^{-1/2} + \gamma_n). \tag{11}$$

From (A2) and (A4), $0 < \overline{\sigma} := \sup_{\theta\in\Theta} \tau(\theta) < \infty$. Define the piecewise quadratic functions

$$g^-(v) := c^- + b^- v + a^- v^2, \qquad g^+(v) := c^+ + b^+ v + a^+ v^2, \tag{12}$$

where $c^\pm := n[M'_n(\alpha_0) \pm r_n]$, $b^\pm := \tau(\theta)\sqrt{n}[M''_n(\alpha_0) \pm \operatorname{sign}(v)c_W\omega_n]$, $a^\pm := \pm\overline{\sigma}^2 c_W$, and $\operatorname{sign}(v)$ denotes the sign of v when $v \ne 0$ and $\operatorname{sign}(0) = 0$. Notice that g^- and g^+ are continuous on the whole real line. To bound the term $|\mathbb{P}_\theta(D_{n,\theta,u} \cap B_{n,\theta}) - \Gamma(u)|$ in (11), let us introduce the events

$$E^\pm_{n,\theta,u} := \{g^\pm(u) \ge 0\}. \tag{13}$$

It follows from Lemma A.2 in Appendix A that, for n large enough and $|u| < 2\sqrt{\ln n}$,

$$\mathbb{P}_\theta(E^-_{n,\theta,u} \cap B_{n,\theta}) \le \mathbb{P}_\theta(D_{n,\theta,u} \cap B_{n,\theta}) \le \mathbb{P}_\theta(E^+_{n,\theta,u} \cap B_{n,\theta})$$

so that

$$\begin{aligned}
&|\mathbb{P}_\theta(D_{n,\theta,u} \cap B_{n,\theta}) - \Gamma(u)| \\
&\quad \le \max\{|\mathbb{P}_\theta(E^-_{n,\theta,u} \cap B_{n,\theta}) - \Gamma(u)|, |\mathbb{P}_\theta(E^+_{n,\theta,u} \cap B_{n,\theta}) - \Gamma(u)|\} \qquad (14) \\
&\quad \le \max\{|\mathbb{P}_\theta(E^-_{n,\theta,u}) - \Gamma(u)|, |\mathbb{P}_\theta(E^+_{n,\theta,u}) - \Gamma(u)|\} + \mathbb{P}_\theta(B^c_{n,\theta}).
\end{aligned}$$

Then the proof of Proposition 2 is easily completed using (10) and the following estimate: There exists a constant C such that for n large enough and $|u| < 2\sqrt{\ln n}$

$$\sup_{\theta\in\Theta} |\mathbb{P}_\theta(E^\pm_{n,\theta,u}) - \Gamma(u)| \le C\left(\frac{1}{\sqrt{n}} + \sqrt{n}r_n + \omega_n\right). \qquad (15)$$

Indeed, $E^\pm_{n,\theta,u} = \{g^\pm(u) \ge 0\}$ with $g^\pm$ defined in (12). We can write

$$\begin{aligned}
E^\pm_{n,\theta,u} &= \{n[M'_n(\alpha_0) \pm r_n] + u\tau(\theta)\sqrt{n}[M''_n(\alpha_0) \pm \operatorname{sign}(u)c_W\omega_n] \pm u^2\overline{\sigma}^2 c_W \ge 0\} \\
&= \left\{\frac{\sqrt{n}}{\sigma_{n,u}(\theta)}\left(M'_n(\alpha_0) + \frac{u\sigma_1(\theta)}{\sqrt{n}m(\theta)}\big(M''_n(\alpha_0) - m(\theta)\big)\right) \ge -\frac{a_n(u,\theta) + b_n(u,\theta)}{\sigma_{n,u}(\theta)}\right\},
\end{aligned}$$

where the positive real number $\sigma_{n,u}(\theta)$ is that of condition (A4′) and

$$a_n(u,\theta) = u\left[\sigma_1(\theta)(1 \pm \frac{\operatorname{sign}(u)c_W\omega_n}{m(\theta)}) \pm \frac{u\overline{\sigma}^2 c_W}{\sqrt{n}}\right], \qquad b_n(u,\theta) = \pm\sqrt{n}r_n.$$

From the second statement of (A4′), it follows that there exists a constant B' such that we have, for n large enough and $|u| < 2\sqrt{\ln n}$,

$$\sup_{\theta\in\Theta}\left|\mathbb{P}_\theta(E^\pm_{n,\theta,u}) - \Gamma\left(\frac{a_n(u,\theta) + b_n(u,\theta)}{\sigma_{n,u}(\theta)}\right)\right| \le \frac{B'}{\sqrt{n}}.$$

Now, from $\underline{\sigma}_1 := \inf_{\theta\in\Theta}\sigma_1(\theta) > 0$ and from the first property of $\sigma_{n,u}(\theta)$ in (A4′), it follows that, for n large enough and $|u| < 2\sqrt{\ln n}$, and for all $\theta \in \Theta$, we have $\sigma_{n,u}(\theta) \ge \underline{\sigma}_1/2$ and

$$\begin{aligned}
\left|\frac{a_n(u,\theta)}{\sigma_{n,u}(\theta)} - u\right| &\le \frac{|u|}{\sigma_{n,u}(\theta)}\left(|\sigma_{n,u}(\theta) - \sigma_1(\theta)| + \frac{c_W\omega_n}{m(\theta)} + \frac{|u|\overline{\sigma}^2 c_W}{\sqrt{n}}\right) \\
&\le \frac{2|u|}{\underline{\sigma}_1}\left[(A' + \overline{\sigma}^2 c_W)\frac{|u|}{\sqrt{n}} + \frac{c_W}{\underline{m}}\omega_n\right] \le C'\left(\frac{u^2}{\sqrt{n}} + |u|\omega_n\right),
\end{aligned}$$

where C' is independent of n, u, θ. We obtain from estimates on the characteristic function of the standard Gaussian distribution reported in [20], page 89, that, for n large enough, $|u| < 2\sqrt{\ln n}$, and $\theta \in \Theta$,

$$\left|\Gamma\left(\frac{a_n(u,\theta)}{\sigma_{n,u}(\theta)}\right) - \Gamma(u)\right| \le C_1\left(\frac{1}{\sqrt{n}} + \omega_n\right)$$

for some $C_1 > 0$. We deduce from similar arguments that, for some constant C_2,

$$\left|\Gamma\left(\frac{a_n(u,\theta)}{\sigma_{n,u}(\theta)}\right) - \Gamma\left(\frac{a_n(u,\theta)+b_n(u,\theta)}{\sigma_{n,u}(\theta)}\right)\right| \le C_2\sqrt{n}r_n.$$

Since C_1, C_2 only depend on A', $\underline{\sigma}_1$, $\underline{m}$, $\overline{\sigma}$ and c_W, the proof of (15) is complete. □

3. A Berry–Esseen bound for an additive functional of geometrically ergodic Markov chains

The main focus of the paper is to apply the general Berry–Esseen result of Theorem 1 to the case of M-estimators as defined in the Introduction when the observations come from a geometrically ergodic Markov chain. To check conditions (A4) and (A4′) in Theorem 1, we need the next probabilistic results based on a recent version of the Berry–Esseen theorem derived by [12] in the geometrically ergodic Markov chain setting.

3.1. The statistical model

Let $(E,\mathcal{E})$ be a measurable space with a countably generated σ-field $\mathcal{E}$ and Θ be some general parameter space. Let $\{X_n\}_{n\ge 0}$ be a Markov chain with state space E, transition kernels $\{Q_\theta(x,\cdot), x\in E\}$, $\theta\in\Theta$ and an initial distribution μ that may or may not depend on θ.

Assumption $(\mathcal{M})$. *Let* $V : E \to [1,+\infty)$ *be an unbounded function* (*independent of* θ). *For each* $\theta \in \Theta$, *there exists a* Q_θ*-invariant probability distribution*, *denoted by* π_θ, *such that*

(VG1) $b_1 := \sup_{\theta\in\Theta} \pi_\theta(V) < +\infty$.
(VG2) *For all* $\gamma \in (0,1]$, *there exist real numbers* $\kappa_\gamma < 1$ *and* $C_\gamma \ge 0$ *such that we have*, *for any* $\theta\in\Theta$, $n \ge 1$ *and* $x \in E$,

$$\sup\{|Q_\theta^n f(x) - \pi_\theta(f)|, f : E \to \mathbb{C} \textit{ measurable}, |f| \le V^\gamma\} \le C_\gamma \kappa_\gamma^n V(x)^\gamma.$$

Throughout Section 3, we assume that $\overline{\mu}(V) := \sup_{\theta\in\Theta}\mu(V) < \infty$. Notice that (VG2) with $\gamma = 1$ implies the following property: For any measurable real-valued function f defined on E such that $|f| \le DV$, for some constant $D > 0$,

$$\forall n \ge 1 \qquad \sup_{\theta\in\Theta} |\mathbb{E}_{\theta,\mu}[f(X_n)] - \pi_\theta(f)| \le DC_1\kappa_1^n \overline{\mu}(V). \tag{16}$$

Moreover, conditions (VG1) and (VG2) imply that, for any $\gamma\in(0,1]$ and $\theta\in\Theta$, Q_θ is V^γ-geometrically ergodic, but it is worth noticing that the constants C_γ and κ_γ do not depend on θ. In the following remark, the properties (VG1) and (VG2) are related to the so-called drift condition w.r.t. the function V for each Q_θ.

Remark 2. Assume that for each $\theta \in \Theta$, Q_θ is aperiodic and ψ-irreducible w.r.t. a certain positive σ-finite measure ψ on E (which may depend on θ).

1. For $\gamma = 1$ and any fixed θ, the properties (VG1)–(VG2) follow from the drift condition: $Q_\theta V \leq \varrho V + \varsigma 1_S$, with $\varrho < 1$, $\varsigma > 0$ and some set S (S is the so-called small set) satisfying the minorization condition $Q_\theta(x, \cdot) \geq c\nu(\cdot)1_S(x)$, where $c > 0$ and ν is a probability measure concentrated on S (see [16], Theorem 16.0.1). In addition, the constants C_1 and κ_1 can be bounded by a quantity involving ϱ, ς, c, the measure ν and the set S (see [17]). To obtain the uniformity in θ, it suffices to check that all these elements do not depend on θ.
2. For any $\gamma \in (0, 1]$, we have $\pi_\theta(V^\gamma) \leq \pi_\theta(V)$ and thus condition (VG1) implies that $\sup_{\theta\in\Theta} \pi_\theta(V^\gamma) < \infty$. Furthermore, under the drift condition, it follows from Jensen's inequality that $Q_\theta V^\gamma \leq \varrho^\gamma V + \varsigma^\gamma 1_S$. Using [17], one obtains (VG2).

3.2. A preliminary uniform Berry–Esseen statement

Let $\alpha_0 = \alpha_0(\theta) \in \mathcal{A}$ be the parameter of interest for the statistical applications we have in mind (see condition (A1), page 706), where θ is the parameter of the Markov chain model and $\mathcal{A}$ is an open interval of the real line.

Let $\xi(\alpha, x, y)$ be a real-valued measurable function defined on $\mathcal{A} \times E^2$ such that the random variable $\xi(\alpha, X_{k-1}, X_k)$ is (integrable and) centered with respect to the stationary distribution π_θ, that is,

$$\mathbb{E}_{\theta,\pi_\theta}[\xi(\alpha_0, X_0, X_1)] = 0,$$

and let

$$S_n(\alpha) := \sum_{k=1}^{n} \xi(\alpha, X_{k-1}, X_k).$$

We investigate the following uniform Berry–Esseen property:

$$\sup_{\theta\in\Theta} \sup_{u\in\mathbb{R}} \left| \mathbb{P}_{\theta,\mu}\left\{ \frac{S_n(\alpha_0)}{\sigma(\theta)\sqrt{n}} \leq u \right\} - \Gamma(u) \right| = \mathrm{O}\left(\frac{1}{\sqrt{n}}\right),$$

where $\sigma^2(\theta)$ will be defined below as the asymptotic variance associated with the random variables $\xi(\alpha, X_{k-1}, X_k)$. When $\{X_n\}_{n\geq 0}$ are i.i.d. and $\xi(\alpha, X_{k-1}, X_k) \equiv \xi(\alpha, X_k)$, this property follows from the Berry–Esseen theorem [6], provided that $\xi(\alpha, X_0)$ has finite third-order moment, uniformly bounded in α, and a variance greater than some positive constant that does not depend on α.

In our Markov framework, the following moment (or V-domination) condition is natural for the functional ξ. In the sequel, this condition will be required for $m_0 = 1, 2$ or 3.

Condition (D_{m_0}). *There exist real constants $m > m_0 \geq 1$ and $C_\xi > 0$ such that*

$$\forall \alpha \in \mathcal{A}, \forall (x, y) \in E^2 \qquad |\xi(\alpha, x, y)|^m \leq C_\xi\big(V(x) + V(y)\big). \tag{D_{m_0}}$$

This domination condition implies that

$$\begin{aligned}\mathbb{E}_{\theta,\pi_\theta}[|\xi(\alpha, X_0, X_1)|^m] &= \int |\xi(\alpha, x, y)|^m Q_\theta(x, \mathrm{d}y)\,\mathrm{d}\pi_\theta(x) \\ &\leq C_\xi\big(\pi_\theta(V) + \pi_\theta(Q_\theta V)\big) < \infty,\end{aligned} \tag{17}$$

and since $m \geq 1$, observe that $\mathbb{E}_{\theta,\pi_\theta}[|\xi(\theta, X_0, X_1)|] < \infty$.

Proposition 3. *Suppose that Assumption $(\mathcal{M})$ holds true and that ξ is centered and satisfies condition (D_1). Then, we have $\sup_{\theta\in\Theta}\sup_{n\geq 1}|\mathbb{E}_{\theta,\mu}[S_n(\alpha_0)]| < \infty$. In particular, for each $\theta \in \Theta$, $\lim_n \mathbb{E}_{\theta,\mu}[S_n(\alpha_0)/n] = 0$. If, in addition, ξ satisfies condition (D_2), then for each $\theta \in \Theta$, the non-negative real number*

$$\sigma^2(\theta) := \lim_n \frac{\mathbb{E}_{\theta,\mu}[S_n(\alpha_0)^2]}{n}$$

is well defined and does not depend on μ. Furthermore, the function $\sigma^2(\cdot)$ is bounded on Θ, and there exists a positive constant C, only depending on C_ξ and $\overline{\mu}(V)$, such that

$$\forall \theta \in \Theta \qquad \forall n \geq 1, \left|\sigma^2(\theta) - \frac{\mathbb{E}_{\theta,\mu}[S_n(\alpha_0)^2]}{n}\right| \leq \frac{C}{n}.$$

Now, we are ready to state our uniform Berry–Esseen statement for $S_n(\alpha_0)$.

Theorem 2. *Let us assume that*:

1. *Assumption $(\mathcal{M})$ holds true*;
2. *the functional ξ is centered and satisfies condition (D_3)*;
3. $\sigma_0^2 := \inf_{\theta\in\Theta}\sigma^2(\theta) > 0$.

Then, there exists a constant $B(\xi)$ such that

$$\forall n \geq 1 \qquad \sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\mu}\left\{\frac{S_n(\alpha_0)}{\sigma(\theta)\sqrt{n}} \leq u\right\} - \Gamma(u)\right| \leq \frac{B(\xi)}{\sqrt{n}}.$$

Furthermore, the constant $B(\xi)$ depends on the functional ξ, but only through σ_0 and the constant C_ξ of condition (D_3).

The fact that we look for a Berry–Esseen bound with a constant $B(\xi)$ independent of θ is natural given our main purpose, that is, to prove a uniform Berry–Esseen theorem for M-estimators.

There are several methods for deriving Berry–Esseen bound for the functionals of Markov chains (see [3,13]). But to prove Proposition 3 and Theorem 2, we use the weak spectral method developed in [12]. (A Berry–Esseen theorem is established in [11] for sequences of the form $\{\xi(X_k)\}_{k\geq 0}$ under the conditions $\mu(V) < \infty$ and $|\xi|^3 \leq CV$; however, the case of sequences of the form $\{\xi(X_{k-1}, X_k)\}_{k\geq 0}$ is not a direct corollary of this work since the Markov chain $\{(X_{k-1}, X_k)\}_{k\geq 0}$ may not be geometrically ergodic.) This method allows us to control the constant $B(\xi)$ as a function of C_ξ for checking assumption (A4′) of Theorem 1 (see the arguments following equation (32) in Section 4). This follows from the next key technical result. Although the proof of the Berry–Esseen theorem only requires Taylor expansions up to the order m_0 and Condition (D_{m_0}) with $m_0 = 3$, for the purpose of possible further applications, Lemma 1 below is stated for any $m_0 \in \mathbb{N}^*$.

Lemma 1. *If ξ is centered and satisfies Condition (D_{m_0}) with $m_0 \in \mathbb{N}^*$, then there exists $\beta > 0$ such that*

$$\forall\theta \in \Theta, \forall n \geq 1, \forall t \in [-\beta, \beta] \qquad \mathbb{E}_{\theta,\mu}\big[\mathrm{e}^{\mathrm{i}tS_n(\alpha_0)}\big] = \lambda_\theta(t)^n\big(1 + L_\theta(t)\big) + r_{\theta,n}(t), \tag{18}$$

where $\lambda_\theta(\cdot)$, $L_\theta(\cdot)$ and $r_{\theta,n}(\cdot)$ are some m_0 times continuously differentiable functions from $[-\beta, \beta]$ into $\mathbb{C}$ satisfying $\lambda_\theta(0) = 1$, $\lambda'_\theta(0) = 0$, $L_\theta(0) = 0$ and $r_{\theta,n}(0) = 0$. Furthermore, there exists $\rho \in (0, 1)$ such that we have for $\ell = 0, \dots, m_0$:

$$G_\ell := \sup\big\{\rho^{-n}\big|r_{\theta,n}^{(\ell)}(t)\big|, |t| \leq \beta, \theta \in \Theta, n \geq 1\big\} < \infty.$$

Finally, the constants β, ρ, G_ℓ and the following ones (for $\ell = 0, \dots, m_0$),

$$E_\ell := \sup\big\{\big|\lambda_\theta^{(\ell)}(t)\big|, |t| \leq \beta, \theta \in \Theta\big\} < \infty,$$

$$F_\ell := \sup\big\{\big|L_\theta^{(\ell)}(t)\big|, |t| \leq \beta, \theta \in \Theta\big\} < \infty,$$

depend on ξ, but only through the constant C_ξ of Condition (D_{m_0}).

Lemma 1 is proved in Section 3.3. The definition of $L_\theta(t)$ and $r_{\theta,n}(t)$ (see (25) and (26)) shows that the constants F_ℓ and G_ℓ also depend on $\overline{\mu}(V)$ (see Remark 3). Now Lemma 1 allows us to derive Proposition 3 and Theorem 2.

Proof of Proposition 3. Assume that ξ is centered and satisfies (D_{m_0}) with $m_0 \in \mathbb{N}^*$. Proceeding as in (17) and using (16), (VG1) and $\overline{\mu}(V) < \infty$, we obtain that

$$\sup_{\theta\in\Theta}\sup_{k\geq 1}\mathbb{E}_{\theta,\mu}[|\xi(\alpha_0, X_{k-1}, X_k)|^m] < \infty \qquad \text{for some } m > m_0. \tag{19}$$

Now assume $m_0 = 1$, and let $\phi(t) := \mathbb{E}_{\theta,\mu}[\mathrm{e}^{\mathrm{i}tS_n(\alpha_0)}]$, $t \in \mathbb{R}$. Then $\phi'(0) = \mathrm{i}\mathbb{E}_{\theta,\mu}[S_n(\alpha_0)]$, but Lemma 1 also gives $\phi'(0) = L'_\theta(0) + r'_{\theta,n}(0)$. Hence $\sup_{\theta\in\Theta}\sup_{n\geq 1}|\mathbb{E}_{\theta,\mu}[S_n(\alpha_0)]| \leq F_1 + G_1$. Next, assume $m_0 = 2$. From (19) we have $\mathbb{E}_{\theta,\mu}[S_n(\alpha_0)^2] < \infty$, and thus we can write $\phi''(0) = -\mathbb{E}_{\theta,\mu}[S_n(\alpha_0)^2]$, and $\phi''(0) = n\lambda''_\theta(0) + L''_\theta(0) + r''_{\theta,n}(0)$ by Lemma 1. Thus we obtain $|\lambda''_\theta(0) + \mathbb{E}_{\theta,\mu}[S_n(\alpha_0)^2]/n| \leq (|L''_\theta(0)| + |r''_{\theta,n}(0)|)/n \leq (F_2 + G_2)/n$. Set $\sigma^2(\theta) := -\lambda''_\theta(0)$. Then $\sup_{\theta\in\Theta}\sigma^2(\theta) \leq E_2$ (by Lemma 1), and the proof is complete with $C := F_2 + G_2$. □

Proof of Theorem 2. Recall that ξ is centered and satisfies condition (D_3). To prove the result, we use Lemma 1 with $m_0 = 3$ and we adapt the arguments of the i.i.d. case. Recall that $\sigma^2(\theta) = -\lambda''_\theta(0)$. According to the classical Berry–Esseen inequality (see [6]), we must prove that for some suitable positive constant c, $\sup_{\theta\in\Theta} A_n(\theta) = \mathrm{O}(n^{-1/2})$, where

$$A_n(\theta) := \int_{-c\sqrt{n}}^{c\sqrt{n}} \left| \frac{\mathbb{E}[\mathrm{e}^{\mathrm{i}tS_n(\alpha_0)/(\sigma(\theta)\sqrt{n})}] - \mathrm{e}^{-t^2/2}}{t} \right| \mathrm{d}t.$$

For the moment, we just assume that $0 < c \le \beta\sigma_0$, where β is the real number in Lemma 1. Notice that $|t| \le c$ implies $|t/\sigma(\theta)| \le \beta$ for all $\theta \in \Theta$. Using Lemma 1, we have

$$\begin{aligned} A_n(\theta) &\le \int_{-c\sqrt{n}}^{c\sqrt{n}} \left| \frac{\lambda_\theta(t/(\sigma(\theta)\sqrt{n}))^n - \mathrm{e}^{-t^2/2}}{t} \right| \mathrm{d}t \\ &\quad + \int_{-c\sqrt{n}}^{c\sqrt{n}} \left| \lambda_\theta\left(\frac{t}{\sigma(\theta)\sqrt{n}}\right)\right|^n \left| \frac{L_\theta(t/(\sigma(\theta)\sqrt{n}))}{t} \right| \mathrm{d}t \\ &\quad + \int_{-c\sqrt{n}}^{c\sqrt{n}} \left| \frac{r_{\theta,n}(t/(\sigma(\theta)\sqrt{n}))}{t} \right| \mathrm{d}t \\ &:= I_n(\theta) + J_n(\theta) + K_n(\theta). \end{aligned}$$

By a Taylor expansion, for all $\theta \in \Theta$ and $|v| \le c$,

$$\left| \lambda_\theta\left(\frac{v}{\sigma(\theta)}\right) - 1 + \frac{v^2}{2} \right| \le \frac{E_3}{6\sigma_0^3}|v|^3,$$

where E_3 is defined in Lemma 1. Hereafter, set $c := \min\{\beta\sigma_0, 3\sigma_0^3/2E_3, \sqrt{2}\}$. From the last inequality, deduce that for any $|v| \le c$

$$\left| \lambda_\theta\left(\frac{v}{\sigma(\theta)}\right) \right| \le 1 - \frac{v^2}{2} + \frac{v^2}{4} \le \mathrm{e}^{-v^2/4}.$$

Therefore, for any $t \in \mathbb{R}$ such that $|t| \le c\sqrt{n}$,

$$\left| \lambda_\theta\left(\frac{t}{\sigma(\theta)\sqrt{n}}\right) \right|^n \le \mathrm{e}^{-t^2/4}. \tag{20}$$

Let us write

$$\lambda_\theta\left(\frac{t}{\sigma(\theta)\sqrt{n}}\right)^n - \mathrm{e}^{-t^2/2} = \left(\lambda\left(\frac{t}{\sigma(\theta)\sqrt{n}}\right) - \mathrm{e}^{-t^2/(2n)}\right) \sum_{k=0}^{n-1} \lambda_\theta\left(\frac{t}{\sigma(\theta)\sqrt{n}}\right)^{n-k-1} \mathrm{e}^{-kt^2/(2n)}.$$

Notice that $|\lambda_\theta(t/\sigma(\theta)\sqrt{n}) - \exp(-t^2/2n)| \le (a + E_3/6\sigma_0^3)|t/\sqrt{n}|^3$ if $a := \sup_{|v|\le c} |\psi^{(3)}(v)|$ with $\psi(v) := 6\exp(-v^2/2)$. Moreover,

$$\sum_{k=0}^{n-1}\left|\lambda_\theta\left(\frac{t}{\sigma(\theta)\sqrt{n}}\right)\right|^{n-k-1} \mathrm{e}^{-kt^2/2n} \le \sum_{k=0}^{n-1} \mathrm{e}^{-t^2(n-k-1)/(4n)}\mathrm{e}^{-kt^2/(4n)} \le bn\mathrm{e}^{-t^2/4},$$

where $b := \sup_{|v|\le c}\exp(v^2/4)$. Hence

$$\left|\lambda_\theta\left(\frac{t}{\sigma(\theta)\sqrt{n}}\right)^n - \mathrm{e}^{-t^2/2}\right| \le \left(a + \frac{E_3}{6\sigma_0^3}\right) bn^{-1/2}|t|^3\mathrm{e}^{-t^2/4},$$

which yields $\sup_{\theta\in\Theta} I_n(\theta) \le bn^{-1/2}(a + E_3/6\sigma_0^3)\int_{\mathbb{R}} t^2\exp(-t^2/4)\,\mathrm{d}t$. Next, using (20) and $L_\theta(0) = 0$,

$$\sup_{\theta\in\Theta} J_n(\theta) \le \frac{F_1}{\sigma_0\sqrt{n}}\int_{\mathbb{R}} \mathrm{e}^{-t^2/4}\,\mathrm{d}t.$$

Finally, using $r_{\theta,n}(0) = 0$, we have $\sup_{\theta\in\Theta}|r_{\theta,n}(t/\sigma(\theta)\sqrt{n})| \le (|t|/\sigma_0\sqrt{n})G_1\rho^n$, so that $\sup_{\theta\in\Theta} K_n(\theta) \le (2cG_1/\sigma_0)\rho^n$. Gathering the results, we deduce that

$$\sup_{\theta\in\Theta} A_n(\theta) \le \frac{A}{\sqrt{n}} + \frac{2cG_1}{\sigma_0}\rho^n,$$

where the constants A, ρ, G_1 and c depend on C_ξ of condition (D_3). The Berry–Esseen inequality [6] then yields

$$\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\mu}\left\{\frac{S_n(\theta)}{\sigma(\theta)\sqrt{n}} \le u\right\} - \Gamma(u)\right| \le \frac{1}{\pi}\left(\frac{A}{\sqrt{n}} + \frac{2cG_1}{\sigma_0}\rho^n + \frac{24\eta}{c\sqrt{n}}\right),$$

where $\eta = \sup_{u\in\mathbb{R}}|\Gamma'(u)|$. The proof of Theorem 2 is complete. □

3.3. Proof of Lemma 1

For $\theta \in \Theta$ fixed, Lemma 1 follows from [12], Section 10. Here, we must prove that all the constants in Lemma 1 are uniform in θ and depend on ξ as claimed. For this purpose, the weak spectral method is outlined below (in the V-geometrical ergodicity context) and we give the main statements by paying special attention to the constants. For convenience, the technical proofs are postponed in Appendix B.

• *Geometrical ergodicity of Q_θ*. Let $0 < \gamma \le 1$. We denote by $\mathcal{B}_\gamma$ the weighted supremum-normed space of measurable complex-valued functions f on E such that

$$\|f\|_\gamma := \sup_{x\in E}\frac{|f(x)|}{V(x)^\gamma} < \infty.$$

$(\mathcal{B}_\gamma, \|\cdot\|_\gamma)$ is a Banach space. The space of bounded operators on $\mathcal{B}_\gamma$ is denoted by $\mathcal{L}(\mathcal{B}_\gamma)$, and the associated operator norm is still denoted by $\|\cdot\|_\gamma$. We have from (VG1)

$$\sup_{\theta\in\Theta} \pi_\theta(V^\gamma) \le b_1 = \sup_{\theta\in\Theta} \pi_\theta(V) < \infty, \tag{21}$$

so that π_θ is a continuous linear form on $\mathcal{B}_\gamma$. Define the following rank-one projection on $\mathcal{B}_\gamma$:

$$\forall f \in \mathcal{B}_\gamma \qquad \Pi_\theta f := \pi_\theta(f) 1_E.$$

Then condition (VG2) in Assumption ($\mathcal{M}$) can be rewritten as follows: $Q_\theta \in \mathcal{L}(\mathcal{B}_\gamma)$ and there exist $\kappa_\gamma < 1$ and $C_\gamma > 0$ such that

$$\forall \theta \in \Theta, \forall f \in \mathcal{B}_\gamma, \forall n \ge 1 \qquad \|Q_\theta^n f - \Pi_\theta f\|_\gamma \le C_\gamma \kappa_\gamma^n \|f\|_\gamma. \tag{22}$$

From (21) and (22), $\|Q_\theta^n\|_\gamma = \sup_{x\in E}(Q_\theta^n V^\gamma)(x)/V(x)^\gamma$ is uniformly bounded in $n \in \mathbb{N}^*$ and $\theta \in \Theta$.

• *The Fourier kernels associated with Q_θ and ξ.* Assume that, for all $\alpha \in \mathcal{A}$, $\xi(\alpha, \cdot, \cdot)$ is measurable. The Fourier kernels associated with Q_θ and ξ are denoted by $\{Q_\theta(t)(x, \mathrm{d}y), t \in \mathbb{R}\}$ and defined by

$$\forall x \in E \qquad Q_\theta(t)(x, \mathrm{d}y) := \mathrm{e}^{\mathrm{i}t\xi(\alpha_0,x,y)} Q_\theta(x, \mathrm{d}y).$$

Let us recall that $S_n(\alpha_0) := \sum_{k=1}^n \xi(\alpha_0, X_{k-1}, X_k)$. The following link between $Q_\theta(t)$ and the characteristic function of $S_n(\alpha_0)$ is well-known in the spectral method:

$$\forall n \ge 1, \forall t \in \mathbb{R} \qquad \mathbb{E}_{\theta,\mu}\big[\mathrm{e}^{\mathrm{i}tS_n(\alpha_0)}\big] = \mu(Q_\theta(t)^n 1_E). \tag{23}$$

In fact, we have $\mathbb{E}_{\theta,\mu}[e^{itS_n(\alpha_0)} f(X_n)] = \mu(Q_\theta(t)^n f)$ for any real-valued measurable bounded function f on E. This can be easily checked by induction using the Markov property and the following equality:

$$\forall n \ge 2 \qquad \mathbb{E}_{\theta,\mu}\big[\mathrm{e}^{\mathrm{i}tS_n(\alpha_0)} f(X_n)\big] = \mathbb{E}_{\theta,\mu}\big[\mathrm{e}^{\mathrm{i}tS_{n-1}(\alpha_0)} (Q_\theta(t) f)(X_{n-1})\big].$$

• *Spectral study of $Q_\theta(t)$ on $\mathcal{B}_\gamma$ (for t near 0).* It can be easily seen that, for all $t \in \mathbb{R}$, we have $Q_\theta(t) \in \mathcal{L}(\mathcal{B}_\gamma)$. For $\kappa \in (0, 1)$, we set

$$\mathcal{D}_\kappa := \{z \in \mathbb{C}\colon |z| \ge \kappa, |z - 1| \ge (1 - \kappa)/2\}.$$

Lemma 2. *Let $\gamma \in (0, 1)$. For all $\kappa \in (\kappa_\gamma, 1)$, there exists $\beta_{\gamma,\kappa} > 0$ such that, for $\theta \in \Theta$, $|t| \le \beta_{\gamma,\kappa}$ and $z \in \mathcal{D}_\kappa$, we have $(z - Q_\theta(t))^{-1} \in \mathcal{L}(\mathcal{B}_\gamma)$ and*

$$\mathcal{R}_{\gamma,\kappa} := \sup\big\{\big\|\big(z - Q_\theta(t)\big)^{-1}\big\|_\gamma\colon \theta \in \Theta, |t| \le \beta_{\gamma,\kappa}, z \in \mathcal{D}_\kappa\big\} < \infty.$$

Moreover, the constants $\beta_{\gamma,\kappa}$ and $\mathcal{R}_{\gamma,\kappa}$ depend on ξ, but only via the constant C_ξ of Condition (D_{m_0}).

For θ fixed, Lemma 2 is established in [12], Proposition 10.1, thanks to the theorem of Keller and Liverani [14,15]. Here, we only have to prove that the constants $\beta_{\gamma,\kappa}$ and $\mathcal{R}_{\gamma,\kappa}$ are uniform in θ and depend on ξ as stated above. According to [14], Remark, page 145, it is enough to check that the constants are so involved in the hypotheses of the Keller–Liverani theorem. This is due to Lemmas B.1–B.2 in Appendix B.

• *Proof of formula* (18). Now assume that ξ satisfies Condition (D_{m_0}) for some $m_0 \in \mathbb{N}^*$. Let $\gamma_0 \in (0, 1)$ be fixed such that $\gamma_0 + m_0/m < 1$. For any $\kappa \in (\kappa_{\gamma_0}, 1)$, denote by $\Gamma_{0,\kappa}$ the oriented circle centered at $z = 0$, with radius κ, and by $\Gamma_{1,\kappa}$ the oriented circle centered at $z = 1$, with radius $(1 - \kappa)/2$. Note that both $\Gamma_{0,\kappa}$ and $\Gamma_{1,\kappa}$ are contained in $\mathcal{D}_\kappa$. From (22) and Lemma 2, one can deduce that we have, for all $n \ge 1$, $\theta \in \Theta$, and $t \in [-\beta_{\gamma_0,\kappa}; \beta_{\gamma_0,\kappa}]$, the following equality in $\mathcal{L}(\mathcal{B}_{\gamma_0})$:

$$Q_\theta(t)^n = \lambda_\theta(t)^n \Pi_\theta(t) + N_\theta(t)^n, \tag{24}$$

where $\lambda_\theta(t)$ is the dominating simple eigenvalue of $Q_\theta(t)$ and $\Pi_\theta(t)$ and $N_\theta(t)^n$ are the elements of $\mathcal{L}(\mathcal{B}_{\gamma_0})$ defined by the following line integrals:

$$\Pi_\theta(t) := \frac{1}{2i\pi} \oint_{\Gamma_{1,\kappa}} \bigl(z - Q_\theta(t)\bigr)^{-1} \, dz \quad \text{and} \quad N_\theta(t)^n := \oint_{\Gamma_{0,\kappa}} z^n \bigl(z - Q_\theta(t)\bigr)^{-1} \, dz.$$

Note that we have $\lambda_\theta(0) = 1$ and $\Pi_\theta(0) = \Pi_\theta$ from (22). Also observe that, from Lemma 2 and the definition of $\Gamma_{0,\kappa}$, we have $\|N_\theta(t)^n\|_\gamma = O(\kappa^n)$. Since $1_E \in \mathcal{B}_{\gamma_0}$ and $\mu(V) < \infty$ (μ is a continuous linear form on $\mathcal{B}_{\gamma_0}$), the equalities (23) and (24) give:

$$\mathbb{E}_{\theta,\mu}\bigl[e^{itS_n(\alpha_0)}\bigr] = \lambda_\theta(t)^n \mu(\Pi_\theta(t)1_E) + \mu(N_\theta(t)^n 1_E).$$

Therefore, formula (18) holds true with

$$L_\theta(t) := \mu(\Pi_\theta(t)1_E) - 1, \qquad r_{\theta,n}(t) := \mu(N_\theta(t)^n 1_E) \qquad (n \in \mathbb{N}^*).$$

We have $L_\theta(0) = \mu(\Pi_\theta 1_E) - 1 = 0$ and $r_{\theta,n}(0) = \mu(N_\theta(0)^n 1_E) = \mu(Q_\theta^n 1_E - \Pi_\theta 1_E) = 0$. Finally, to make the link with Lemma 3 below easier, let us observe that

$$1 + L_\theta(t) = \frac{1}{2i\pi} \oint_{\Gamma_{1,\kappa}} \mu\bigl(\bigl(z - Q_\theta(t)\bigr)^{-1} 1_E\bigr) \, dz, \tag{25}$$

$$r_{\theta,n}(t) = \frac{1}{2i\pi} \oint_{\Gamma_{0,\kappa}} z^n \mu\bigl(\bigl(z - Q_\theta(t)\bigr)^{-1} 1_E\bigr) \, dz. \tag{26}$$

• *Regularity properties of* $\lambda(\cdot)$, $L_\theta(\cdot)$, $r_{\theta,n}(\cdot)$. Let γ_0' be such that $\gamma_0 + m_0/m < \gamma_0' < 1$. We denote by $\mathcal{L}(\mathcal{B}_{\gamma_0}, \mathcal{B}_{\gamma_0'})$ the space of the bounded linear operators from $\mathcal{B}_{\gamma_0}$ to $\mathcal{B}_{\gamma_0'}$, and by $\|\cdot\|_{\gamma_0,\gamma_0'}$ the associated operator norm.

Lemma 3. *We have the following regularity properties*:

(a) *The map* $Q_\theta(\cdot)$ *is* m_0*-times continuously differentiable from* $\mathbb{R}$ *to* $\mathcal{L}(\mathcal{B}_{\gamma_0}, \mathcal{B}_{\gamma_0'})$, *and we have* $\mathcal{Q}_\ell := \sup_{t\in\mathbb{R},\theta\in\Theta} \|Q_\theta^{(\ell)}(t)\|_{\gamma_0,\gamma_0'} < \infty$ *for* $\ell = 0, \dots, m_0$.

(b) *There exist some real numbers* $\kappa \in (\kappa_{\gamma_0}, 1)$ *and* $0 < \beta < \beta_{\gamma_0,\kappa}$ *such that, for all* $\theta \in \Theta$ *and* $z \in \mathcal{D}_\kappa$, *the function* $R_{\theta,z} : t \mapsto (z - Q_\theta(t))^{-1}$ *is* m_0*-times continuously differentiable from* $[-\beta, \beta]$ *into* $\mathcal{L}(\mathcal{B}_{\gamma_0}, \mathcal{B}_{\gamma_0'})$, *and we have for* $\ell = 0, \dots, m_0$:

$$\sup\{\|R^{(\ell)}_{\theta,z}(t)\|_{\gamma_0,\gamma_0'}: |t| \le \beta, z \in \mathcal{D}_\kappa, \theta \in \Theta\} < \infty.$$

The scalars β, κ *and all the bounds in* (a) *and* (b) *depend on* ξ *only via the constant* C_ξ *of Condition* (D_{m_0}).

For θ fixed, Lemma 3 is established in [12], Proposition 10.3. It can be also derived from [8], which relaxes the assumptions used in [9,10] to obtain Taylor expansions of the resolvent maps. (As observed in [8], the passage to the differentiability properties can be derived from [4].) However, a fine control of the constants is still required. Using either [8] or [12], Section 10, this control is derived from Lemma 2 and from Lemma B.3 in Appendix B.

Since $1_E \in \mathcal{B}_{\gamma_0}$ and μ is a continuous linear form on $\mathcal{B}_{\gamma_0'}$ (use $\overline{\mu}(V) < \infty$), Lemma 3(b) gives that, for any $z \in \Gamma_{0,\kappa} \cup \Gamma_{1,\kappa}$, the $\mathbb{C}$-valued function $t \mapsto \mu((z - Q_\theta(t))^{-1} 1_E)$ is m_0-times continuously differentiable on $[-\beta, \beta]$ and that its m_0 first derivatives are uniformly bounded in θ and $z \in \Gamma_{0,\kappa} \cup \Gamma_{0,\kappa}$. The regularity properties (and the related bounds) for $L_\theta(\cdot)$ and $r_{\theta,n}(\cdot)$ then follow from (25) and (26), while those concerning the function $\lambda_\theta(\cdot)$ follow from both Lemma 3(a) and Lemma 3(b), according to a formula given in [12], Section 7.2. Finally the property $\lambda_\theta'(0) = 0$ can be proved as follows. By deriving (18) (applied with $\mu = \pi_\theta$) at $t = 0$ and by using the fact that ξ is centered, we have $0 = \mathrm{i}\mathbb{E}_{\theta,\pi_\theta}[S_n(\alpha_0)] = n\lambda_\theta'(0) + L_\theta'(0) + r_{\theta,n}'(0)$. Hence $\lambda_\theta'(0) = 0$.

Remark 3. Notice that, according to (25)–(26), the constants F_ℓ and G_ℓ in Lemma 1 also depend on the supremum in θ of the norm of μ in $\mathcal{B}'_{\gamma_0'}$, namely $\sup_{\theta\in\Theta} \mu(V^{\gamma_0'})$.

4. A Berry–Esseen theorem for M-estimators

Consider a Markov chain satisfying Assumption ($\mathcal{M}$) of Section 3.1. Let us introduce the statistic

$$M_n(\alpha) := \frac{1}{n} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k), \tag{27}$$

where α is the parameter of interest, F is a real-valued measurable function on $\mathcal{A} \times E^2$ and $\mathcal{A}$ is an open interval of the real line.

Assume that F satisfies condition (D_1) and let

$$M_\theta(\alpha) := \lim_{n\to\infty} \mathbb{E}_{\theta,\mu}[M_n(\alpha)] = \mathbb{E}_{\theta,\pi_\theta}[F(\alpha, X_0, X_1)],$$

which is well defined by Proposition 3. Assume also that, for each $\theta \in \Theta$, there exists a unique $\alpha_0 = \alpha_0(\theta) \in \mathcal{A}$, the so-called true value of the parameter of interest, such that $M_\theta(\alpha) > M_\theta(\alpha_0)$,

$\forall\alpha \neq \alpha_0$. To estimate $\alpha_0 = \alpha_0(\theta)$, we consider an M-estimator $\widehat{\alpha}_n$ as defined in Section 2, that is, $M_n(\widehat{\alpha}_n) \leq \min_{\alpha\in\mathcal{A}} M_n(\alpha) + c_n$, where $\{c_n\}_{n\geq 1}$ is a sequence of non-negative real numbers going to zero.

Let F' and F'' be real-valued measurable functions defined on $\mathcal{A} \times E^2$ and let

$$M'_n(\alpha) := \frac{1}{n}\sum_{k=1}^{n} F'(\alpha, X_{k-1}, X_k), \qquad M''_n(\alpha) := \frac{1}{n}\sum_{k=1}^{n} F''(\alpha, X_{k-1}, X_k). \tag{28}$$

The functionals F' and F'' could be the first- and second-order partial derivatives of F with respect to α, but this is not necessary to deduce our next result. Consider the following assumptions on F' and F'' (and, implicitly, on c_n; see (V3)).

Assumptions.

(V0) *F' and F'' satisfy condition (D_3).*
(V1) *$\forall\theta \in \Theta, \mathbb{E}_{\theta,\pi_\theta}[F'(\alpha_0, X_0, X_1)] = 0$ and $\alpha_0 = \alpha_0(\theta)$ is unique with this property.*
(V2) *$m(\theta) := \mathbb{E}_{\theta,\pi_\theta}[F''(\alpha_0, X_0, X_1)]$ satisfies $\inf_{\theta\in\Theta} m(\theta) > 0$.*
(V3) *$M'_n(\widehat{\alpha}_n)$ satisfies condition* (A3), *that is, $\forall n \geq 1$ and there exists $r_n > 0$ independent of θ such that $r_n = \mathrm{o}(1/\sqrt{n})$ and $\sup_{\theta\in\Theta} \mathbb{P}_{\theta,\mu}\{|M'_n(\widehat{\alpha}_n)| \geq r_n\} = \mathrm{O}(n^{-1/2})$.*

Notice that (V0) ensures $\sup_{\theta\in\Theta} m(\theta) < \infty$ (see (17)). Now, as a consequence of Proposition 3 applied to F' and F'', the conditions (V0)–(V2) enable us to define the asymptotic variances:

$$\sigma_1^2(\theta) := \lim_n \frac{1}{n}\mathbb{E}_{\theta,\mu}\left[\left(\sum_{k=1}^{n} F'(\alpha_0, X_{k-1}, X_k)\right)^2\right],$$
$$\sigma_2^2(\theta) := \lim_n \frac{1}{n}\mathbb{E}_{\theta,\mu}\left[\left(\sum_{k=1}^{n} F''(\alpha_0, X_{k-1}, X_k) - nm(\theta)\right)^2\right].$$

Moreover, condition (V0) and Proposition 3 ensure that $\sup_{\theta\in\Theta} \sigma_j(\theta) < \infty$ for $j = 1, 2$. The following conditions are also assumed to hold.

(V4) $\inf_{\theta\in\Theta} \sigma_j(\theta) > 0$ *for* $j = 1, 2$.
(V5) *There exist $\eta \in (0, 1/2)$ and $C > 0$ such that*

$$\forall(\alpha, \tilde{\alpha}) \in \mathcal{A}^2, \forall(x, y) \in E^2 \qquad |F''(\alpha, x, y) - F''(\tilde{\alpha}, x, y)| \leq C|\alpha - \tilde{\alpha}|\big(V(x) + V(y)\big)^\eta.$$

(V6) *Set $d := \inf_{\theta\in\Theta} m(\theta)/8\pi_\theta(V^\eta)$ with η defined in* (V5). *There exists $\gamma_n = \mathrm{o}(1)$ such that*

$$\sup_{\theta\in\Theta} \mathbb{P}_{\theta,\mu}\{|\widehat{\alpha}_n - \alpha_0| \geq d\} \leq \gamma_n.$$

Theorem 3. *Assume that Assumption $(\mathcal{M})$ holds true, F satisfies condition (D_1) and conditions* (V0)–(V6) *are fulfilled. Let $\tau(\theta) := \sigma_1(\theta)/m(\theta)$. Then there exists a positive constant C such* *that*

$$\forall n \geq 1 \qquad \sup_{\theta \in \Theta} \sup_{u \in \mathbb{R}} \left| \mathbb{P}_{\theta,\mu}\left\{ \frac{\sqrt{n}}{\tau(\theta)}(\widehat{\alpha}_n - \alpha_0) \leq u \right\} - \Gamma(u) \right| \leq C\left(\frac{1}{\sqrt{n}} + \sqrt{n} r_n + \gamma_n \right).$$

The statement in the above theorem corresponds to that of the i.i.d. case in [20] up to few changes: First, the variances of the i.i.d. context (namely, $\mathbb{E}_\theta[F'(\theta, X_0)^2]$ and $\mathbb{E}_\theta[(F''(\theta, X_0) - m(\theta))^2]$ for an i.i.d. sequence $\{X_n\}_{n \geq 0}$ and a functional $F(\theta, x)$) are replaced by the above asymptotic variances $\sigma_1^2(\theta)$ and $\sigma_2^2(\theta)$ (this is natural in a general Markovian context); second, the uniform (in θ) third-order moment conditions (namely, $\sup_{\theta \in \Theta} \mathbb{E}_\theta[|F'(\theta, X_0)|^3 + |F''(\theta, X_0)|^3] < \infty$) on both F', F'' are replaced by the domination condition (D_3) for F', F''; third, even when $F' = \partial F / \partial \alpha$, here we allow for a positive sequence r_n, $n \geq 1$, provided it decreases to zero sufficiently fast. The second point is specific to the geometrically ergodic Markov chain case. Indeed, in the same statistical model, Dehay and Yao [5] proved a CLT for maximum likelihood estimates under a second-order domination assumption on the two first derivatives of the functional, which corresponds to inequality (D_{m_0}) with $m_0 = 2$. Here the previous second-order assumption is replaced by the (almost) optimal condition (D_3) for deriving the Berry–Esseen theorem for M-estimators.

Proof of Theorem 3. It suffices to check the conditions (A1)–(A6) of Theorem 1. The limit $M'_\theta(\alpha) := \lim_n \mathbb{E}_{\theta,\mu}[M'_n(\alpha)]$ is well defined by Proposition 3 and condition (V0), the uniqueness of α_0 is guaranteed by (V1) and hence (A1) holds true. One more application of Proposition 3 ensures that $\mathbb{E}_{\theta,\pi_\theta}[F''(\alpha_0, X_0, X_1)] = \lim_n \mathbb{E}_{\theta,\mu}[M''_n(\alpha_0)]$, hence (A2) is satisfied. Condition (V3) is nothing else but (A3). The Berry–Esseen properties in (A4) are associated with the functionals $F'(\alpha_0, x, y)$ and $F''(\alpha_0, x, y)$ respectively, so that they directly follow from Theorem 2.

Now, let us check that (A5) holds true with $\omega_n \equiv 0$. Define $W := V^\eta$, where $\eta \in (0, 1/2)$ is the scalar in (V5) and notice that $\mathbb{E}_{\theta,\pi_\theta}[W(X_0)^{1/\eta}] = \pi_\theta(V)$. Next, since $V \geq 1$ and $\eta \in (0, 1/2)$, we have $1 \leq W \leq W^2 \leq V$ so that $1 \leq \pi_\theta(W) \leq \pi_\theta(W^2) \leq \pi_\theta(V) \leq b_1$ by property (VG1). Deduce that $\sup_{\theta \in \Theta} \pi_\theta(W) < \infty$, and by Proposition 3 applied to $\xi(\theta, x, y) = W(y)$

$$\sup_{n \geq 1} \sup_{\theta \in \Theta} \frac{1}{n} \mathbb{E}_{\theta,\mu}\left[\left(\sum_{k=1}^n W(X_k) - n\pi_\theta(W) \right)^2 \right] < \infty.$$

Now, condition (A5) is guaranteed by the properties $(\mathcal{M})$ and (V5) with $\omega_n \equiv 0$, $c_W := \sup_{\theta \in \Theta} \pi_\theta(W)$ and $W_n := (1/n)\sum_{k=1}^n (W(X_{k-1}) + W(X_k))$ provided that

$$\sup_{\theta \in \Theta} \mathbb{P}_{\theta,\mu}\{8\pi_\theta(W) \leq W_n\} = \mathrm{O}(n^{-1}). \tag{29}$$

To prove (29), set $S_n := \sum_{k=1}^n W(X_k)$. Since $W_n \leq 2S_n/n + (W(X_0) + W(X_n))/n$ and $\pi_\theta(W) \geq 1$,

$$\begin{aligned} \mathbb{P}_{\theta,\mu}\{8\pi_\theta(W) \leq W_n\} &\leq \mathbb{P}_{\theta,\mu}\{S_n \geq 2n\pi_\theta(W)\} + \mathbb{P}_{\theta,\mu}\{W(X_0) + W(X_n) \geq 4n\pi_\theta(W)\} \\ &\leq \mathbb{P}_{\theta,\mu}\{S_n - n\pi_\theta(W) \geq n\} + \mathbb{P}_{\theta,\mu}\{W(X_0) + W(X_n) \geq 4n\}. \end{aligned}$$

Equality (29) is then obtained by Markov's inequality,

$$\mathbb{P}_{\theta,\mu}\{8\pi_\theta(W) \le W_n\} \le \frac{1}{n^2}\mathbb{E}_{\theta,\mu}\big[\big(S_n - n\pi_\theta(W)\big)^2\big] + \left(\frac{1}{4n}\right)^{1/\eta}\mathbb{E}_{\theta,\mu}\big[\big(W(X_0)+W(X_n)\big)^{1/\eta}\big]$$
$$= \mathrm{O}(n^{-1}),$$

since

$$\sup_{\theta\in\Theta}\sup_{n\ge 1}\mathbb{E}_{\theta,\mu}\big[\big(W(X_0)+W(X_n)\big)^{1/\eta}\big] \le 2^{1/\eta-1}[\overline{\mu}(V) + C_1\overline{\mu}(V) + b_1],$$

using $(a+b)^{1/\eta} \le 2^{1/\eta-1}(a^{1/\eta}+b^{1/\eta})$ for any $a, b \ge 0$ and (VG1)–(VG2). Notice also that now condition (V6) is identical to condition (A6).

The difficult part is to check the Berry–Esseen-type property (A4′). For this purpose, let $\Xi := \{\xi_i(\cdot,\cdot,\cdot), i \in I\}$ denote an arbitrary family of real-valued functionals defined on $\mathcal{A}\times E^2$. Suppose that each ξ_i is centered, that is, $\mathbb{E}_{\theta,\pi_\theta}[\xi_i(\alpha_0, X_0, X_1)] = 0$ for all $i \in I$ and $\theta \in \Theta$, and that condition (D_3) is fulfilled uniformly in $i \in I$, that is,

$$\exists m > 3, \exists C \ge 0, \forall i \in I, \forall \alpha \in \mathcal{A}, \forall (x,y) \in E^2 \qquad |\xi_i(\alpha,x,y)|^m \le C\big(V(x)+V(y)\big). \tag{30}$$

For each $i \in I$, set $S_n(\alpha_0, i) := \sum_{k=1}^n \xi_i(\alpha_0, X_{k-1}, X_k)$, and using Proposition 3, associate the corresponding asymptotic variance denoted by $\sigma_i^2(\theta)$. Moreover, assume that

$$0 < \inf\{\sigma_i(\theta), \theta\in\Theta, i \in I\} \le \sup\{\sigma_i(\theta), \theta\in\Theta, i\in I\} < \infty. \tag{31}$$

Then, we deduce from Theorem 2 that, under conditions $(\mathcal{M})$, (30), (31) and $\overline{\mu}(V) < \infty$, there exists a constant B such that

$$\forall n \ge 1 \qquad \sup_{i\in I}\sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\mu}\left\{\frac{S_n(\alpha_0,i)}{\sigma_i(\theta)\sqrt{n}} \le u\right\} - \Gamma(u)\right| \le \frac{B}{\sqrt{n}}. \tag{32}$$

This allows us to establish the two conditions in (A4′). Indeed, for $(p, v) \in \mathbb{N}^*\times\mathbb{R}$ with v such that $|v| \le 2\sqrt{\ln p}$, let us introduce the functional $\xi_{p,v}$ defined by

$$\xi_{p,v}(\alpha_0,x,y) := F'(\alpha_0,x,y) + \frac{v}{\sqrt{p}}\frac{\sigma_1(\theta)}{m(\theta)}\big(F''(\alpha_0,x,y) - m(\theta)\big).$$

Set $S_n(\alpha_0, p, v) := \sum_{k=1}^n \xi_{p,v}(\alpha_0, X_{k-1}, X_k)$, and

$$\alpha_\theta(p,v) := \frac{v}{\sqrt{p}}\frac{\sigma_1(\theta)}{m(\theta)},$$

$$S'_n(\theta) := \sum_{k=1}^n F'(\alpha_0, X_{k-1}, X_k),$$

$$S''_n(\theta) := \sum_{k=1}^n F''(\alpha_0, X_{k-1}, X_k) - nm(\theta),$$

so that $S_n(\alpha_0, p, v) = S'_n(\alpha_0) + \alpha_\theta(p, v)S''_n(\alpha_0)$. Notice that $\mathbb{E}_{\theta,\pi_\theta}[\xi_{p,v}(\alpha_0, X_0, X_1)] = 0$ by (V1)–(V2). We have

$$\begin{aligned}&\mathbb{E}_{\theta,\pi_\theta}[S_n(\alpha_0, p, v)^2] - \mathbb{E}_{\theta,\pi_\theta}[S'_n(\alpha_0)^2]\\ &\quad= \alpha_\theta(p, v)^2\mathbb{E}_{\theta,\pi_\theta}[S''_n(\alpha_0)^2] + 2\alpha_\theta(p, v)\mathbb{E}_{\theta,\pi_\theta}[S'_n(\alpha_0)S''_n(\alpha_0)].\end{aligned}$$

From (V2) and the fact that $\sigma_1(\cdot)$ is bounded, we have $|\alpha_\theta(p, v)| \leq A|v|/\sqrt{p}$ for some $A > 0$ that does not depend on θ. Besides, as already mentioned in this section, one can define the asymptotic variances $\sigma_1^2(\theta)$ and $\sigma_2^2(\theta)$ associated with the functionals F' and F'' by

$$\sigma_1^2(\theta) := \lim_n \frac{1}{n}\mathbb{E}_{\theta,\pi_\theta}[S'_n(\alpha_0)^2], \qquad \sigma_2^2(\theta) := \lim_n \frac{1}{n}\mathbb{E}_{\theta,\pi_\theta}[S''_n(\alpha_0)^2].$$

Similarly, the asymptotic variance $\sigma_{p,v}^2(\theta)$ associated with $\xi_{p,v}$ can be defined by:

$$\sigma_{p,v}^2(\theta) := \lim_n \frac{1}{n}\mathbb{E}_{\theta,\pi_\theta}[S_n(\alpha_0, p, v)^2].$$

Then it follows from $|\mathbb{E}_{\theta,\pi_\theta}[S'_n(\alpha_0)S''_n(\alpha_0)]| \leq \mathbb{E}_{\theta,\pi_\theta}[S'_n(\alpha_0)^2]^{1/2}\mathbb{E}_{\theta,\pi_\theta}[S''_n(\alpha_0)^2]^{1/2}$ that

$$|\sigma_{p,v}^2(\theta) - \sigma_1^2(\theta)| \leq A^2\frac{v^2}{p}\sigma_2^2(\theta) + 2A\frac{|v|}{\sqrt{p}}\sigma_1(\theta)\sigma_2(\theta).$$

Since $\sigma_j(\cdot)$ is bounded ($j = 1, 2$) and $|v| \leq 2\sqrt{\ln p} \leq 2\sqrt{p}$, the previous inequality shows that there exists $C' > 0$, independent of θ, such that

$$|\sigma_{p,v}^2(\theta) - \sigma_1^2(\theta)| \leq C'\frac{|v|}{\sqrt{p}}.$$

Set $\overline{\sigma}_1 := \sup_{\theta\in\Theta}\sigma_1(\theta)$ and $\underline{\sigma}_1 := \inf_{\theta\in\Theta}\sigma_1(\theta)$ (we have $\underline{\sigma}_1 > 0$ from (V4)). Using $|v|/\sqrt{p} \leq 2\sqrt{\ln p/p}$ and $\sqrt{\ln p/p} = \mathrm{o}(1)$, the above inequality implies that there exists $P_0 \in \mathbb{N}$ such that we have, for all $p \geq P_0$ and v such that $|v| \leq 2\sqrt{\ln p}$,

$$\forall\theta \in \Theta \qquad \tfrac{1}{2}\underline{\sigma}_1 \leq \sigma_{p,v}(\theta) \leq \tfrac{3}{2}\overline{\sigma}_1.$$

In particular, under the same condition on (p, v), this gives $\sigma_{p,v}(\theta) + \sigma_1(\theta) \geq 3\underline{\sigma}_1/2$, hence $|\sigma_{p,v}(\theta) - \sigma_1(\theta)| \leq 2C'|v|/3\underline{\sigma}_1\sqrt{p}$. This proves the first assertion in (A4′).

Now, let us define

$$I = \{(p, v) \in \mathbb{N}^* \times \mathbb{R}:\ p \geq P_0, |v| \leq 2\sqrt{\ln p}\}.$$

It follows from (V0), (V2) and $\overline{\sigma}_1 < +\infty$ that the family $\Xi := \{\xi_{p,v}, (p, v) \in I\}$ satisfies (30). Besides, the above bounds of $\sigma_{p,v}(\theta)$ give the property (31). Then equation (32) shows that there exists $B' > 0$ such that we have for all $n \geq 1$, $(p, v) \in I$, $\theta \in \Theta$ and $u \in \mathbb{R}$:

$$\left|\mathbb{P}_{\theta,\mu}\left\{\frac{S_n(\alpha_0, p, v)}{\sigma_{p,v}(\theta)\sqrt{n}} \leq u\right\} - \Gamma(u)\right| \leq \frac{B'}{\sqrt{n}}.$$

Finally, let us fix any integer $n \geq P_0$ and any real number u such that $|u| \leq 2\sqrt{\ln n}$. Then, the previous Berry–Esseen bound with $p := n$ and $v := u$ provides the second property of (A4′). Indeed, we obtain from $S_n'(\alpha_0) = nM_n'(\alpha_0)$ and $S_n''(\alpha_0) = n(M_n''(\alpha_0) - m(\theta))$ that

$$\begin{aligned}\frac{S_n(\alpha_0, n, u)}{\sigma_{n,u}(\theta)\sqrt{n}} &= \frac{1}{\sigma_{n,u}(\theta)\sqrt{n}}\left(S_n'(\alpha_0) + \frac{u}{\sqrt{n}}\frac{\sigma_1(\theta)}{m(\theta)}S_n''(\alpha_0)\right) \\ &= \frac{\sqrt{n}}{\sigma_{n,u}(\theta)}\left(M_n'(\alpha_0) + \frac{u\sigma_1(\theta)}{\sqrt{n}m(\theta)}\big(M_n''(\alpha_0) - m(\theta)\big)\right).\end{aligned}$$

Now the proof of Theorem 3 is complete. □

5. An example: AR(1) process with ARCH(1) errors

Let us apply our theoretical results to an AR(1) process with ARCH(1) errors that belongs to the class of ARMA–GARCH models (see [7] and the references therein). The observations are generated by the process

$$X_n = \rho_0 X_{n-1} + \sigma(X_{n-1}; a_0, b_0)\varepsilon_n, \qquad n = 1, 2, \dots, \tag{33}$$

where X_0 has some probability distribution μ, $\sigma^2(x; a, b) := a + bx^2$ and $|\rho_0| < 1$, $a_0, b_0 > 0$ are the true values of the parameters. $\{\varepsilon_n\}_{n\geq 1}$ is a sequence of i.i.d. random variables with zero mean and variance equal to 1, with finite pth order moment for some p to be specified below and (unknown) density f_ε that is continuous and positive on $\mathbb{R}$. $\{\varepsilon_n\}_{n\geq 1}$ is independent of X_0. For simplicity, hereafter μ is assumed to be the Dirac distribution δ_0. The "true" parameter θ in the associated statistical model is the vector $(\rho_0, a_0, b_0) \in \Theta \subset [-\overline{\rho}, \overline{\rho}] \times [m_a, M_a] \times [m_b, M_b] \subset \mathbb{R}^3$, where $\overline{\rho} \in (0, 1)$, $0 < m_a < M_a < \infty$ and $0 < m_b < M_b < 1$ are given such that $(\overline{\rho} + \sqrt{M_b})^p \int_{\mathbb{R}} (1 + |y|)^p f_\varepsilon(y)\,\mathrm{d}y < 1$. For illustration, we apply our results to estimate ρ_0 and b_0.

First, let us check that the Markov chain defined by (33) satisfies Assumption $(\mathcal{M})$ of Section 3.1 with $V(x) = (1 + |x|)^p$. To check (VG1)–(VG2) and the existence of the Q_θ-invariant probability measure π_θ, by [17], Theorem 2.3, it suffices to prove that there exist constants $\varrho \in (0, 1)$, $c, \varsigma > 0$, a Borel subset S of the real line and a probability measure ν concentrated on S such that the following two conditions hold true (see Remark 2): For all $\theta \in \Theta$,

$$\forall x \in \mathbb{R} \qquad Q_\theta V(x) \leq \varrho V(x) + \varsigma 1_S(x) \quad \text{and} \quad Q_\theta(x, \cdot) \geq c\nu(\cdot)1_S(x). \tag{34}$$

In our setting, the transition probability of $\{X_n\}_{n\geq 0}$ is given by

$$Q_\theta(x, B) = \int 1_B\big(\rho_0 x + \sigma(x, a_0, b_0)y\big) f_\varepsilon(y)\,\mathrm{d}y$$

for any Borel set $B \subset \mathbb{R}$. As a consequence, for all $\theta \in \Theta$ and $x \in \mathbb{R}$,

$$\begin{aligned}\frac{Q_\theta V(x)}{V(x)} &= \int_{\mathbb{R}} \frac{V(\rho_0 x + \sigma(x, a_0, b_0)y)}{V(x)} f_\varepsilon(y)\,\mathrm{d}y \\ &\leq \int_{\mathbb{R}} \left(\frac{1 + \overline{\rho}|x| + (\sqrt{M_a} + \sqrt{M_b}|x|)|y|}{1 + |x|}\right)^p f_\varepsilon(y)\,\mathrm{d}y.\end{aligned}$$

By Fatou's lemma,

$$\limsup_{|x|\to\infty}\left(\sup_{\theta\in\Theta}\frac{Q_\theta V(x)}{V(x)}\right)\le\left(\overline{\rho}+\sqrt{M_b}\right)^p\int_{\mathbb{R}}(1+|y|)^p f_\varepsilon(y)\,\mathrm{d}y=:\iota<1.$$

Next, fix $\varrho\in(\iota,1)$. There exists $s>0$ such that for each $|x|>s$, $Q_\theta V(x)\le\varrho V(x)$ for all $\theta\in\Theta$. Set $S:=[-s;s]$. For all $x\in S$ and $\theta\in\Theta$,

$$Q_\theta V(x)\le\int_{\mathbb{R}}\left(1+\overline{\rho}s+\left(\sqrt{M_a}+\sqrt{M_b}s\right)|y|\right)^p f_\varepsilon(y)\,\mathrm{d}y<\infty,$$

so that the first condition in (34) is guaranteed. To check the second condition in (34), define

$$0<\delta(u):=\inf_{x\in S,\theta\in\Theta} f_\varepsilon\left(\sigma^{-1}(x,a_0,b_0)(u-\rho_0x)\right),\qquad u\in\mathbb{R}.$$

Then, for any $x\in S$, Borel set $B\subset\mathbb{R}$ and $\theta\in\Theta$,

$$\begin{aligned}Q_\theta(x,B)&=\int_{\mathbb{R}}1_B\left(\rho_0x+\sigma(x,a_0,b_0)y\right)f_\varepsilon(y)\,\mathrm{d}y\\&=\int_B\frac{f_\varepsilon(\sigma^{-1}(x,a_0,b_0)(u-\rho_0x))}{\sigma(x,a_0,b_0)}\,\mathrm{d}u\ge\int_B\frac{\delta(u)}{m_a}\,\mathrm{d}u.\end{aligned}$$

Define the measure $m(\mathrm{d}u):=m_a^{-1}\delta(u)\,\mathrm{d}u$ and notice that $m(S)>0$. We deduce from above that all $\theta\in\Theta$, $x\in S$ and Borel set $B\subset\mathbb{R}$,

$$Q_\theta(x,B)\ge m(B)\ge m(B\cap S)=m(S)\nu(B),$$

where ν is the probability measure $\nu(B):=m(B\cap S)/m(S)$. Hence the second condition in (34) is fulfilled and Assumption $(\mathcal{M})$ is satisfied for $\{X_n\}_{n\ge0}$ defined in (33).

Second, to estimate ρ_0, one can use the least-squares estimator,

$$\widehat{\rho}_n:=\frac{\sum_{k=1}^n X_kX_{k-1}}{\sum_{k=1}^n X_{k-1}^2}=\arg\min_\rho\frac{1}{n}\sum_{k=1}^n F(\rho,X_{k-1},X_k),$$

where $F(\rho,X_{k-1},X_k):=(X_k-\rho X_{k-1})^2$. We show that the assumptions of Theorem 3 are satisfied so that we have a uniform Berry–Esseen bound for $\widehat{\rho}_n$. Fix some $p>6$ and recall that $\int_{\mathbb{R}}|y|^p f_\varepsilon(y)\,\mathrm{d}y<\infty$. Take $F'(\rho,X_{k-1},X_k):=-2X_{k-1}(X_k-\rho X_{k-1})$ and $F''(\rho,X_{k-1},X_k):=2X_{k-1}^2$. The conditions (V0) and (V1) are obviously fulfilled. Next, define $m(\theta):=\mathbb{E}_{\theta,\pi_\theta}[F''(\rho_0,X_{k-1},X_k)]$ and notice that $m(\theta)/2=a_0+(b_0+\rho_0^2)m(\theta)/2$. It follows that $m(\theta)=2a_0/(1-\rho_0^2-b_0)>2m_a$ and thus (V2) holds. Condition (V3) is satisfied with $r_n\equiv0$. From Proposition 3, we can use the Q_θ-invariant probability measure π_θ to check condition (V4). Notice that $\lim_n\mathbb{E}_{\theta,\pi_\theta}[X_n^2]=m(\theta)/2>m_a$ and recall that $\{\varepsilon_n\}_{n\ge1}$ is i.i.d. We deduce that

$$\sigma_1^2(\theta)=\lim_n\frac{4}{n}\sum_{k=1}^n\mathbb{E}_{\theta,\pi_\theta}[X_{k-1}^2\sigma^2(X_{k-1},a_0,b_0)\varepsilon_k^2]\ge4a_0\lim_n\mathbb{E}_{\theta,\pi_\theta}[X_n^2]\ge4m_a^2.$$

To derive a lower bound for $\sigma_2^2(\theta)$, let us decompose

$$\mathbb{E}_{\theta,\pi_\theta}\left[\sum_{k=1}^n \bigl(F''(\rho_0, X_{k-1}, X_k) - m(\theta)\bigr)\right]^2 = \sum_{k=1}^n v_{k,k} + 2 \sum_{1\le k<l\le n} v_{k,l},$$

where $v_{k,l} := \mathbb{E}_{\theta,\pi_\theta}[(F''(\rho_0, X_{k-1}, X_k) - m(\theta))(F''(\rho_0, X_{l-1}, X_l) - m(\theta))]$, $k \le l$. It is easily checked that $v_{k,l} = (\rho_0^2 + b_0)v_{k,l-1}$ for $k < l$. In particular, this implies $v_{k,l} > 0$, $k \le l$. Next, by elementary inequalities, we can obtain $\inf_\theta \mathbb{E}_{\theta,\pi_\theta}[(F''(\rho_0, X_0, X_1) - m(\theta))^2] \ge K$ for some positive constant K depending on the variance of ε_1^2. Deduce that $\sigma_2^2(\theta) \ge K$, hence (V4) holds true. Condition (V5) is trivially satisfied. To check the consistency of condition (V6), we take advantage of the explicit form of $\widehat{\rho}_n$. Indeed, we have

$$\begin{aligned}\widehat{\rho}_n - \rho_0 &= \frac{n^{-1}\sum_{k=1}^n (X_k X_{k-1} - \rho_0 \mathbb{E}_{\theta,\pi_\theta}[X_1^2]) - \rho_0 n^{-1}\sum_{k=1}^n (X_{k-1}^2 - \mathbb{E}_{\theta,\pi_\theta}[X_1^2])}{n^{-1}\sum_{k=1}^n (X_{k-1}^2 - \mathbb{E}_{\theta,\pi_\theta}[X_1^2]) + \mathbb{E}_{\theta,\pi_\theta}[X_1^2]} \\ &=: \frac{\Delta_{1n} - \rho_0 \Delta_{2n}}{\Delta_{2n} + \mathbb{E}_{\theta,\pi_\theta}[X_1^2]}.\end{aligned}$$

By Chebyshev's inequality, for any $d > 0$, $\mathbb{P}_{\theta,\delta_0}\{|\Delta_{1n}| > d\} \le d^{-2}n^{-1}\mathbb{E}_{\theta,\delta_0}[n\Delta_{1n}^2]$. Proposition 3 guarantees that $\mathbb{E}_{\theta,\delta_0}[n\Delta_{1n}^2]$ is uniformly bounded (with respect to θ). Similar arguments apply to Δ_{2n}. Since $\mathbb{E}_{\theta,\pi_\theta}[X_1^2] > m_a$ for all θ, we deduce that (V6) holds with $\gamma_n = \mathrm{O}(n^{-1})$. Finally, by Theorem 3, there exists $C > 0$ such that

$$\forall n \ge 1 \qquad \sup_{\theta\in\Theta} \sup_{u\in\mathbb{R}} \left| \mathbb{P}_{\theta,\delta_0}\left\{ \frac{\sqrt{n}}{\sigma_1(\theta)m(\theta)^{-1}}(\widehat{\rho}_n - \rho_0) \le u \right\} - \Gamma(u) \right| \le \frac{C}{\sqrt{n}}. \tag{35}$$

Third, let us now turn to the estimation of b_0. For this purpose, assume that the ε_n's have a moment of order p for some $p > 12$. Recall that $a_0 = m(\theta)(1 - \rho_0^2 - b_0)/2$ and notice that $\tau_0^2 := m(\theta)/2$ is easily estimated by $\widehat{\tau}_n^2 := n^{-1}\sum_{k=1}^n X_k^2$. Next, define

$$T_n(b; r, v) := \frac{1}{n}\sum_{k=1}^n \eta_k(b, r, v)^2$$

$$\text{with } \eta_k(b, r, v) := (X_k - rX_{k-1})^2 - v(1 - r^2 - b) - bX_{k-1}^2,$$

$$\text{with } \frac{\partial T_n}{\partial b}(b; r, v) = \frac{2}{n}\sum_{k=1}^n (v - X_{k-1}^2)\eta_k(b, r, v), \ \frac{\partial^2 T_n}{\partial b^2}(b; r, v) = \frac{2}{n}\sum_{k=1}^n (v - X_{k-1}^2)^2.$$

If ρ_0 and a_0 were known, one could easily estimate b_0 by least squares, more precisely by minimizing $T_n(b; \rho_0, \tau_0^2)$ with respect to b. With this idea in mind, our feasible estimator of b_0 is defined as follows:

$$\widehat{b}_n := \arg\min_{b\in[m_b, M_b]} M_n(b) \qquad \text{with } M_n(b) := T_n(b; \widehat{\rho}_n, \widehat{\tau}_n^2).$$

Define $F'(b, X_{k-1}, X_k) := 2(\tau_0^2 - X_{k-1}^2)\eta_k(b, \rho_0, \tau_0^2)$, $F''(b, X_{k-1}, X_k) := 2(\tau_0^2 - X_{k-1}^2)^2$ and $M_n'(b) := \partial T_n/\partial b(b; \rho_0, \tau_0^2)$, $M_n''(b) := \partial^2 T_n/\partial b^2(b; \rho_0, \tau_0^2)$. Let us point out that, in this case, $M_n'(\cdot)$ and $M_n''(\cdot)$ are only approximations of the derivatives of $M_n(\cdot)$. Checking assumptions (V0)–(V2) is obvious and therefore we skip the details. To check condition (V3) for $M_n'(\widehat{b}_n)$, we use the decomposition $M_n'(\widehat{b}_n) = A_n + \Delta_n = A_n + \Delta_{1n} + \Delta_{2n} + \Delta_{3n}$ with

$$A_n := \frac{2}{n}\sum_{k=1}^{n}(\tau_0^2 - X_{k-1}^2)\eta_k(\widehat{b}_n, \widehat{\rho}_n, \widehat{\tau}_n^2),$$

$$\Delta_n := \frac{2}{n}\sum_{k=1}^{n}(\tau_0^2 - X_{k-1}^2)\bigl(\eta_k(\widehat{b}_n, \rho_0, \tau_0^2) - \eta_k(\widehat{b}_n, \widehat{\rho}_n, \widehat{\tau}_n^2)\bigr),$$

$$\Delta_{1n} := \frac{4(\widehat{\rho}_n - \rho_0)}{n}\sum_{k=1}^{n}(\tau_0^2 - X_{k-1}^2)(X_k - \rho_0 X_{k-1})X_{k-1},$$

$$\Delta_{2n} := -\frac{2(\widehat{\rho}_n - \rho_0)^2}{n}\sum_{k=1}^{n}(\tau_0^2 - X_{k-1}^2)X_{k-1}^2,$$

$$\Delta_{3n} := 2\{\widehat{\tau}_n^2(1 - \widehat{\rho}_n^2 - \widehat{b}_n) - \tau_0^2(1 - \rho_0^2 - \widehat{b}_n)\}(\tau_0^2 - \widehat{\tau}_n^2 + X_n^2/n).$$

We check that each term satisfies condition (V3) with a suitable r_n. First, we can write

$$0 = \frac{\partial M_n}{\partial b}(\widehat{b}_n) = A_n + B_n \qquad \text{with } B_n := \frac{2(\widehat{\tau}_n^2 - \tau_0^2)}{n}\sum_{k=1}^{n}\eta_k(\widehat{b}_n, \widehat{\rho}_n, \widehat{\tau}_n^2).$$

By elementary algebra $B_n = 2(\widehat{\tau}_n^2 - \tau_0^2)(\widehat{b}_n + \widehat{\rho}_n^2)X_n^2/n$. Using the Berry–Esseen bound for $\widehat{\tau}_n^2$ (see Theorem 2) and Markov's inequality for X_n^{2+a} for some small $a > 0$, we can prove that $\mathbb{P}_{\theta,\delta_0}\{|B_n| \geq n^{-1}\} = \mathrm{O}(n^{-1/2})$ so that $\mathbb{P}_{\theta,\delta_0}\{|A_n| \geq n^{-1}\} = \mathrm{O}(n^{-1/2})$. By the bound in equation (35), we have $\sup_\theta \mathbb{P}_{\theta,\delta_0}\{|\widehat{\rho}_n - \rho_0|^j \geq n^{-j/2}\log^{j/2} n\} = \mathrm{O}(n^{-1/2})$, $j = 1, 2$. Use this with $j = 1$ and our Theorem 2 for the centered functional $\xi(X_k, X_{k-1}) = (\tau_0^2 - X_{k-1}^2)(X_k - \rho_0 X_{k-1})X_{k-1}$ to deduce that $\mathbb{P}_{\theta,\delta_0}\{|\Delta_{1n}| \geq n^{-1}\log n\} = \mathrm{O}(n^{-1/2})$. Next, the bound on $|\widehat{\rho}_n - \rho_0|^2$ and Theorem 2 applied to the centered functional $\xi(X_k, X_{k-1}) = (\tau_0^2 - X_{k-1}^2)X_{k-1}^2 - \tau_0^4 + \mathbb{E}_{\theta,\pi_\theta}[X_{k-1}^4]$ allow us to deduce that $\mathbb{P}_{\theta,\delta_0}\{|\Delta_{2n}| \geq n^{-1}\log n\} = \mathrm{O}(n^{-1/2})$. Finally, use the Berry–Esseen bounds for $\widehat{\rho}_n$ and $\widehat{\tau}_n^2$ and Markov's inequality for X_n^{2+a} with some $a > 0$ to deduce that $\mathbb{P}_{\theta,\delta_0}\{|\Delta_{3n}| \geq n^{-1}\log n\} = \mathrm{O}(n^{-1/2})$. Combining these facts gives that $M_n'(\widehat{b}_n)$ satisfies condition (V3) with $r_n = n^{-1}\log n$. Condition (V4) can be checked using similar arguments to those used for $\widehat{\rho}_n$ and, therefore, the details are omitted. Condition (V5) is trivially satisfied. Finally, let us note that

$$\widehat{b}_n - b_0 = \frac{\sum_{k=1}^{n}(\widehat{\tau}_n^2 - X_{k-1}^2)\eta_k(b_0, \widehat{\rho}_n, \widehat{\tau}_n^2)}{\sum_{k=1}^{n}(\widehat{\tau}_n^2 - X_{k-1}^2)^2},$$

and thus condition (V6) can be checked by arguments that we already used in this example. We deduce from Theorem 3 that, for some suitable $\tau(\theta)$,

$$\forall n \geq 1 \qquad \sup_{\theta\in\Theta}\sup_{u\in\mathbb{R}}\left|\mathbb{P}_{\theta,\delta_0}\left\{\frac{\sqrt{n}}{\tau(\theta)}(\widehat{b}_n - b_0) \leq u\right\} - \Gamma(u)\right| = \mathrm{O}\left(\frac{\log n}{\sqrt{n}}\right).$$

The log factor in this Berry–Esseen bound is the price we pay for estimating b_0 by a simple two-step procedure, easy to implement, where we first estimate $\widehat{\rho}_n$ and $\widehat{\tau}_n^2$ and then we use the least-squares criterion $M_n(b) = T_n(b; \widehat{\rho}_n, \widehat{\tau}_n^2)$. We feel that the log factor could be removed by using a direct approach where the three parameters are estimated simultaneously, but the investigation of this idea with Markov chain data is left for future work.

6. Conclusion

In this paper, we study the Berry–Esseen theorem for M-estimators (or minimum contrast estimators) of some parameter α_0 on the real line. The estimators are defined from a criterion based on a functional $F(\alpha, X_{n-1}, X_n)$ of the observation process $\{X_n\}_{n\geq 0}$. Our approach to derive such bounds relies on Pfanzagl's method originally proposed for i.i.d. observations [20]. In a first step, Theorem 1 in [20] is extended to obtain Berry–Esseen bounds for M-estimators based on any sequence of observations satisfying suitable conditions. In a second step, the specific case of V-geometrically ergodic Markov observations is considered. We show that such Markov framework allows us to apply our general result provided that F and related functionals F', F'' satisfy suitable domination conditions. This result covers those reported in [19,21], which are proved under much stronger moment conditions. We argue that the domination conditions used in the present paper give an almost optimal treatment of Berry–Esseen bounds for V-geometrically ergodic Markov chains. This is possible due to the operator-type procedure developed in [12].

There are several possible extensions of our results. A straightforward one is to follow the lines of the proof [20], Theorem 2, and to consider an estimator of the standard deviation in the Berry–Esseen bounds when this standard deviation depends on θ only through α_0. The details are omitted. Next, for more effective bounds, we need to carefully evaluate the constants involved throughout the paper. This is a direction of future work. Finally, there is no doubt that the operator-type procedure in [12] could be further used in statistical applications with Markov models, in particular with strongly ergodic Markov chains. This is under investigation.

Appendix A: Complements for the proof of Theorem 1

The reader is referred to Proposition 2 and its proof for the notation and the definitions used throughout this part. The following lemma gives key properties of the random functions $g^{\pm}$.

Lemma A.1. *The following properties hold true.*

1. *If* $\nu_{n,\theta} := \sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta)$, *then* $A_n \subset \{g^-(\nu_{n,\theta}) \leq 0 \leq g^+(\nu_{n,\theta})\}$.

2. *For* $\omega \in D_{n,\theta}$, $g^{\pm}$ *are increasing on the interval* $(-2\sqrt{\ln n}, 2\sqrt{\ln n})$ *provided that*

$$\sqrt{n} \geq \frac{2c_W}{\underline{m}}\left[\frac{4\overline{\sigma}^2\overline{m}\sqrt{\ln n}}{\underline{\sigma}_1} + \sqrt{n}\omega_n\right]. \tag{A.1}$$

Proof. We can write from assumptions (A5) and (A3)

$$\begin{aligned}|nM_n'(\alpha_0) + (\widehat{\alpha}_n - \alpha_0)nM_n''(\alpha_0)| &= |nM_n'(\widehat{\alpha}_n) - (\widehat{\alpha}_n - \alpha_0)nR_n(\alpha_0, \widehat{\alpha}_n)| \\ &\leq nr_n + n|\widehat{\alpha}_n - \alpha_0||R_n(\alpha_0, \widehat{\alpha}_n)| \\ &\leq nr_n + n|\widehat{\alpha}_n - \alpha_0|[|\widehat{\alpha}_n - \alpha_0| + \omega_n]W_n.\end{aligned}$$

If $\omega \in A_n$, then

$$|nM_n'(\alpha_0) + (\widehat{\alpha}_n - \alpha_0)nM_n''(\alpha_0)| \leq n|\widehat{\alpha}_n - \alpha_0|^2 c_W + n\omega_n|\widehat{\alpha}_n - \alpha_0|c_W + nr_n.$$

This last inequality is rewritten as

$$n[M_n'(\alpha_0) - r_n] + \tau(\theta)\sqrt{n}[M_n''(\alpha_0) - \operatorname{sign}(\nu_{n,\theta})c_W\omega_n]\nu_{n,\theta} - \tau(\theta)^2 c_W \nu_{n,\theta}^2 \leq 0$$

and

$$n[M_n'(\alpha_0) + r_n] + \tau(\theta)\sqrt{n}[M_n''(\alpha_0) + \operatorname{sign}(\nu_{n,\theta})c_W\omega_n]\nu_{n,\theta} + \tau(\theta)^2 c_W \nu_{n,\theta}^2 \geq 0,$$

with $\nu_{n,\theta} := \sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta)$. Since $0 < \tau(\theta) \leq \overline{\sigma}$, we obtain that

$$g^-(\nu_{n,\theta}) \leq 0 \text{ and } g^+(\nu_{n,\theta}) \geq 0.$$

The second statement is proved as follows for g^+. Note that $a^+ > 0$ and g^+ is continuous. If we restrict $v < 0$, the minimum of this quadratic function $g^+(v)$ is achieved at

$$v_{\min} = -\frac{b^+}{2a^+} = -\frac{\tau(\theta)\sqrt{n}[M_n''(\alpha_0) - c_W\omega_n]}{2\overline{\sigma}^2 c_W},$$

or at the origin if $v_{\min} \geq 0$. Now, if $\omega \in D_{n,\theta}$ and n satisfies condition (A.1), it is easy to check that

$$v_{\min} < -2\sqrt{\ln n}$$

and g^+ is strictly increasing on $(0, \infty)$. Hence, g^+ is increasing on $(-2\sqrt{\ln n}, 2\sqrt{\ln n})$. Similar arguments apply for g^-. □

Lemma A.2. *We have for n large enough and* $|u| < 2\sqrt{\ln n}$

$$E_{n,\theta,u}^- \cap B_{n,\theta} \subset D_{n,\theta,u} \cap B_{n,\theta} \subset E_{n,\theta,u}^+ \cap B_{n,\theta}. \tag{A.2}$$

Proof. It is understood below that $\omega \in B_{n,\theta}$. Since $B_{n,\theta} \subset E_{n,\theta} \cap D_{n,\theta}$ and $|u| < 2\sqrt{\ln n}$, the second statement in Lemma A.1 guarantees that for n large enough

$$\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta) \le u \quad \Longrightarrow \quad g^+\big(\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta)\big) \le g^+(u).$$

Since $B_{n,\theta} \subset A_n$, the first assertion in Lemma A.1 yields $g^+(\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta)) \ge 0$ so that $g^+(u) \ge 0$ when $\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta) \le u$. This proves the second inclusion in (A.2).

Next, assume that $g^-(u) \ge 0$. Since g^- is increasing, we have

$$\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta) > u \quad \Longrightarrow \quad g^-\big(\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta)\big) > g^-(u) \ge 0.$$

Since $B_{n,\theta} \subset A_n$, we know from Lemma A.1 that $g^-(\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta)) \le 0$ which is in contradiction with the above inequality. Thus, $g^-(u) \ge 0$ gives $\sqrt{n}(\widehat{\alpha}_n - \alpha_0)/\tau(\theta) \le u$. □

Appendix B: Complements for the proof of Lemma 1

A first step to control the constants in Lemma 2 is to study the resolvent map $(z - Q_\theta)^{-1}$ of the transition kernel Q_θ acting on $\mathcal{B}_\gamma$.

Lemma B.1. *Let δ, r be such that $\kappa_\gamma < r < 1$ and $0 < \delta < 1 - r$. Then, for any $z \in \mathbb{C}$ such that $|z| > r$ and $|z - 1| > \delta$, the operator $z - Q_\theta$ is invertible on $\mathcal{B}_\gamma$, and we have*:

$$H_\gamma(\delta, r) := \sup\{\|(z - Q_\theta)^{-1}\|_\gamma, \theta \in \Theta, |z| > r, |z - 1| > \delta\} < \infty.$$

Proof. Let $g \in \mathcal{B}_\gamma$, and let us write $h_\theta = g - \pi_\theta(g)1_E$. Since $\pi_\theta(h_\theta) = 0$, it follows from (VG2) that $\|Q_\theta^n h_\theta\|_\gamma \le C_\gamma \kappa_\gamma^n \|h_\theta\|_\gamma$. Now assume $|z| > r$. Then

$$\sum_{k\ge 0} |z|^{-(k+1)} \|Q_\theta^k h_\theta\|_\gamma \le \frac{C_\gamma}{\kappa_\gamma} \sum_{k\ge 0} \left(\frac{\kappa_\gamma}{r}\right)^{k+1} \|h_\theta\|_\gamma \le \frac{C_\gamma}{r - \kappa_\gamma} \|h_\theta\|_\gamma.$$

Thus, $\psi_\theta := \sum_{k\ge 0} z^{-(k+1)} Q_\theta^k h_\theta$ is absolutely convergent in $\mathcal{B}_\gamma$, we have $(z - Q_\theta)\psi_\theta = h_\theta$ and $\|\psi_\theta\|_\gamma \le C_\gamma \|h_\theta\|_\gamma/(r - \kappa_\gamma)$. Besides, if $z \ne 1$, then we clearly have

$$(z - Q_\theta)\left(\frac{\pi_\theta(g)}{z - 1} 1_E\right) = \pi_\theta(g)1_E.$$

Now assume $|z| > r$ and $|z - 1| > \delta$. Then the function $f_\theta := (\pi_\theta(g)/(z - 1))1_E + \psi_\theta$ is such that $(z - Q_\theta)f_\theta = g$. Thus $(z - Q_\theta)^{-1}g = f_\theta$. From (21), we obtain $|\pi_\theta(g)| \le \pi_\theta(|g|) \le \pi_\theta(V^\gamma)\|g\|_\gamma \le b_1\|g\|_\gamma$ and $\|h_\theta\|_\gamma = \|g - \pi_\theta(g)1_E\|_\gamma \le (1 + b_1)\|g\|_\gamma$. This gives: $\|f_\theta\|_\gamma \le (b_1/\delta)\|g\|_\gamma + C_\gamma(1 + b_1)\|g\|_\gamma/(r - \kappa_\gamma)$, hence $H_\gamma(\delta, r) \le [b_1/\delta + C_\gamma(1 + b_1)/(r - \kappa_\gamma)] < \infty$. □

Second, the constants involved in the Doeblin–Fortet inequality and the weak continuity condition of the Keller–Liverani theorem are proved to be uniform in θ and to depend on ξ only via the constant C_ξ of (D_{m_0}). We appeal to [14], Remark, page 145, and to the improvements given in [15]. In the context of strongly ergodic Markov chains, the hypotheses resulting from [14,15] are stated in [12], Section 4, and used here with the auxiliary norm $\|f\|_1 := \sup|f|/V$ on $\mathcal{B}_\gamma$. In the sequel, for $0 < \gamma < \gamma' \le 1$, we denote by $\mathcal{L}(\mathcal{B}_\gamma, \mathcal{B}_{\gamma'})$ the space of the bounded linear operators from $\mathcal{B}_\gamma$ to $\mathcal{B}_{\gamma'}$, and by $\|\cdot\|_{\gamma,\gamma'}$ the associated operator norm (with the convention $\|\cdot\|_\gamma = \|\cdot\|_{\gamma,\gamma}$ when $\gamma' = \gamma$).

Lemma B.2. *Let $\gamma \in (0,1)$. We have*:

(a) $\forall\theta \in \Theta, \forall t \in \mathbb{R}, \forall n \ge 1, \forall f \in \mathcal{B}_\gamma, \|Q_\theta(t)^n f\|_\gamma \le C_\gamma \kappa_\gamma^n \|f\|_\gamma + b_1\|f\|_1$;

(b) $\forall\theta \in \Theta, \forall t \in \mathbb{R}, \|Q_\theta(t) - Q_\theta\|_{\gamma,1} \le 2^{2-\gamma} C_\xi{}^{(1-\gamma)/m}(E_\gamma + E_1)|t|^{1-\gamma}\|f\|_\gamma$,

where $E_\gamma := \sup_{\theta\in\Theta}\|Q_\theta\|_\gamma$, $E_1 := \sup_{\theta\in\Theta}\|Q_\theta\|_1$ and $C_\gamma, \kappa_\gamma, b_1$ are defined in (21) *and* (22).

Proof. By using the inequality $\|Q_\theta(t)^n f\|_\gamma \le \|Q_\theta^n|f|\|_\gamma$, assertion (a) easily follows from (22) and (21). To establish (b), let us recall that we have from (D_{m_0}) (use $V \ge 1$)

$$\begin{aligned}|\xi(\theta,x,y)|^{1-\gamma} &\le C_\xi^{(1-\gamma)/m}\big(V(x) + V(y)\big)^{1-\gamma}\\ &\le 2^{1-\gamma} C_\xi^{(1-\gamma)/m}\big(V(x)^{1-\gamma} + V(y)^{1-\gamma}\big).\end{aligned}$$

Let $f \in \mathcal{B}_\gamma$. From the definition of $Q_\theta(t)f$ and the inequalities $|f| \le V^\gamma\|f\|_\gamma$, $|\mathrm{e}^{\mathrm{i}a} - 1| \le 2|a|^{1-\gamma}$, we obtain that

$$\begin{aligned}|(Q_\theta(t)f)(x) - (Q_\theta f)(x)| &\le \|f\|_\gamma \int_E \big|\mathrm{e}^{\mathrm{i}t\xi(\alpha_0,x,y)} - 1\big| V(y)^\gamma Q_\theta(x,\mathrm{d}y)\\ &\le 2^{2-\gamma} C_\xi{}^{(1-\gamma)/m}|t|^{1-\gamma}\|f\|_\gamma\big[V(x)^{1-\gamma}(Q_\theta V^\gamma)(x) + (Q_\theta V)(x)\big],\end{aligned}$$

from which we deduce (b). □

For the next lemma (used to prove Lemma 3), we introduce the following notation. For any $\theta \in \Theta$, $k \in \mathbb{N}$, $t \in \mathbb{R}$, let us denote by $Q_{\theta,k}(t)$ the operator associated with the kernel: $Q_{\theta,k}(t)(x,\mathrm{d}y) = \mathrm{i}^k\xi(\alpha_0,x,y)^k \mathrm{e}^{\mathrm{i}t\xi(\alpha_0,x,y)} Q_\theta(x,\mathrm{d}y)$ $(x \in E)$.

Lemma B.3. *Let $0 < \gamma < \gamma' \le 1$ and $k = 0,\dots,m_0$*:

(a) *If $\gamma + k/m < \gamma' \le 1$, then the map $t \mapsto Q_{\theta,k}(t)$ is continuous from $\mathbb{R}$ to $\mathcal{L}(\mathcal{B}_\gamma, \mathcal{B}_{\gamma'})$.*

(b) *If $k \le m_0 - 1$ and $\gamma + (k+1)/m < \gamma' \le 1$, then the map $t \mapsto Q_{\theta,k}(t)$ is continuously differentiable from $\mathbb{R}$ to $\mathcal{L}(\mathcal{B}_\gamma, \mathcal{B}_{\gamma'})$, and for all $t \in \mathbb{R}$, $(\mathrm{d}Q_{\theta,k}/\mathrm{d}t)(t)$ is the operator in $\mathcal{L}(\mathcal{B}_\gamma, \mathcal{B}_{\gamma'})$ associated to the kernel $Q_{\theta,k+1}(t)$.*

Finally, we have $\mathcal{Q}_{k,\gamma,\gamma'} := \sup\{\|Q_{\theta,k}(t)\|_{\gamma,\gamma'}, \theta \in \Theta, t \in \mathbb{R}\} < \infty$, and $\mathcal{Q}_{k,\gamma,\gamma'}$ depends on ξ but only via the constant C_ξ of (D_{m_0}).

Proof. Set $\Delta_{\theta,k} := Q_{\theta,k}(t) - Q_{\theta,k}(t_0)$, and let $0 < \varepsilon \le 1$ be such that $\gamma + (k+\varepsilon)/m \le \gamma'$. Using $|\mathrm{e}^{\mathrm{i}a} - 1| \le 2|a|^\varepsilon$ and (D_{m_0}), we obtain for $f \in \mathcal{B}_\gamma$:

$$|\Delta_{\theta,k} f(x)| \le 2|t-t_0|^\varepsilon \|f\|_\gamma \int |\xi(\alpha_0, x, y)|^{k+\varepsilon} V(y)^\gamma Q_\theta(x, \mathrm{d}y)$$
$$\le 2^{1+(k+\varepsilon)/m} C_\xi{}^{(k+\varepsilon)/m} |t-t_0|^\varepsilon \|f\|_\gamma \big(V^{(k+\varepsilon)/m}(x) Q_\theta V^\gamma(x) + Q_\theta V^{\gamma'}(x)\big).$$

Since the functions $V^{-\gamma} Q_\theta V^\gamma$ and $V^{-\gamma'} Q_\theta V^{\gamma'}$ are bounded on E uniformly in $\theta \in \Theta$, we deduce that $\|\Delta_{\theta,k} f\|_{\gamma'} \le D_\xi |t-t_0|^\varepsilon \|f\|_\gamma$, where D_ξ is a positive constant depending on C_ξ (but independent of θ). This gives (a). The proof of (b) is similar, using the operators $Q_{\theta,k}(t) - Q_{\theta,k}(t_0) - (t-t_0) Q_{\theta,k+1}(t_0)$ and the inequality $|\mathrm{e}^{\mathrm{i}a} - 1 - \mathrm{i}a| \le 2|a|^{1+\varepsilon}$. □

References

[1] Arcones, M.A. (1998). A remark on approximate M-estimators. *Statist. Probab. Lett.* **38** 311–321. MR1631213

[2] Bentkus, V., Bloznelis, M. and Götze, F. (1997). A Berry–Esséen bound for M-estimators. *Scand. J. Stat.* **24** 485–502. MR1615335

[3] Bolthausen, E. (1982). The Berry–Esseén theorem for strongly mixing Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **60** 283–289. MR0664418

[4] Campanato, S. (1964). Proprietà di una famiglia di spazi funzionali. *Ann. Scuola Norm. Sup. Pisa* (3) **18** 137–160. MR0167862

[5] Dehay, D. and Yao, J.F. (2007). On likelihood estimation for discretely observed Markov jump processes. *Aust. N. Z. J. Stat.* **49** 93–107. MR2345413

[6] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. New York: Wiley. MR0270403

[7] Francq, C. and Zakoïan, J.M. (2004). Maximum likelihood estimation of pure GARCH and ARMA–GARCH processes. *Bernoulli* **10** 605–637. MR2076065

[8] Gouëzel, S. (2010). Characterization of weak convergence of Birkhoff sums for Gibbs–Markov maps. *Israel J. Math.* **180** 1–41. MR2735054

[9] Gouëzel, S. and Liverani, C. (2006). Banach spaces adapted to Anosov systems. *Ergodic Theory Dynam. Systems* **26** 189–217. MR2201945

[10] Hennion, H. and Hervé, L. (2004). Central limit theorems for iterated random Lipschitz mappings. *Ann. Probab.* **32** 1934–1984. MR2073182

[11] Hervé, L. (2008). Vitesse de convergence dans le théorème limite central pour des chaînes de Markov fortement ergodiques. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 280–292. MR2446324

[12] Hervé, L. and Pène, F. (2010). The Nagaev–Guivarc'h method via the Keller–Liverani theorem. *Bull. Soc. Math. France* **138** 415–489. MR2729019

[13] Jensen, J.L. (1989). Asymptotic expansions for strongly mixing Harris recurrent Markov chains. *Scand. J. Stat.* **16** 47–63. MR1003968

[14] Keller, G. and Liverani, C. (1999). Stability of the spectrum for transfer operators. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **28** 141–152. MR1679080

[15] Liverani, C. (2003). Invariant measures and their properties. A functional analytic point of view. In *Dynamical Systems. Part II. Pubbl. Cent. Ric. Mat. Ennio Giorgi* 185–237. Pisa: Scuola Norm. Sup. MR2071241

[16] Meyn, S.P. and Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability*. London: Springer. MR1287609

[17] Meyn, S.P. and Tweedie, R.L. (1994). Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab.* **4** 981–1011. MR1304770

[18] Michel, R. and Pfanzagl, J. (1971). The accuracy of the normal approximation for minimum contrast estimates. *Z. Wahrsch. Verw. Gebiete* **18** 73–84. MR0288897

[19] Milhaud, X. and Raugi, A. (1989). Étude de l'estimateur du maximum de vraisemblance dans le cas d'un processus autorégressif: Convergence, normalité asymptotique, vitesse de convergence. *Ann. Inst. Henri Poincaré Probab. Stat.* **25** 383–428. MR1045243

[20] Pfanzagl, J. (1971). The Berry–Esseen bound for minimum contrast estimates. *Metrika* **17** 82–91. MR0295467

[21] Prakasa Rao, B.L.S. (1973). On the rate of convergence of estimators for Markov processes. *Z. Wahrsch. Verw. Gebiete* **26** 141–152. MR0339420

Received September 2009 and revised September 2010