Limit theorems for stationary Markov processes with L^2-spectral gap

Déborah Ferré, Loïc Hervé and James Ledoux

Institut National des Sciences Appliquées de Rennes, 20 avenue des Buttes de Coesmes, CS 70839, 35708 Rennes Cedex 7, France.
E-mail: Deborah.Ferre@insa-rennes.fr; Loic.Herve@insa-rennes.fr; JamesLedoux@insa-rennes.fr

Abstract. Let $(X_t, Y_t)_{t \in \mathbb{T}}$ be a discrete or continuous-time Markov process with state space $\mathbb{X} \times \mathbb{R}^d$ where $\mathbb{X}$ is an arbitrary measurable set. Its transition semigroup is assumed to be additive with respect to the second component, i.e. $(X_t, Y_t)_{t \in \mathbb{T}}$ is assumed to be a Markov additive process. In particular, this implies that the first component $(X_t)_{t \in \mathbb{T}}$ is also a Markov process. Markov random walks or additive functionals of a Markov process are special instances of Markov additive processes. In this paper, the process $(Y_t)_{t \in \mathbb{T}}$ is shown to satisfy the following classical limit theorems:

(a) the central limit theorem,
(b) the local limit theorem,
(c) the one-dimensional Berry–Esseen theorem,
(d) the one-dimensional first-order Edgeworth expansion,

provided that we have $\sup_{t \in (0,1] \cap \mathbb{T}} \mathbb{E}_{\pi,0}[|Y_t|^\alpha] < \infty$ with the expected order α with respect to the independent case (up to some $\varepsilon > 0$ for (c) and (d)). For the statements (b) and (d), a Markov nonlattice condition is also assumed as in the independent case. All the results are derived under the assumption that the Markov process $(X_t)_{t \in \mathbb{T}}$ has an invariant probability distribution π, is stationary and has the $\mathbb{L}^2(\pi)$-spectral gap property (that is, $(X_t)_{t \in \mathbb{N}}$ is ρ-mixing in the discrete-time case). The case where $(X_t)_{t \in \mathbb{T}}$ is non-stationary is briefly discussed. As an application, we derive a Berry–Esseen bound for the M-estimators associated with ρ-mixing Markov chains.

Résumé. Soit $(X_t, Y_t)_{t \in \mathbb{T}}$ un processus de Markov en temps discret ou continu et d'espace d'état $\mathbb{X} \times \mathbb{R}^d$ où $\mathbb{X}$ est un ensemble mesurable quelconque. Son semi-groupe de transition est supposé additif suivant la seconde composante, i.e. $(X_t, Y_t)_{t \in \mathbb{T}}$ est un processus additif Markovien. En particulier, ceci implique que la première composante $(X_t)_{t \in \mathbb{T}}$ est également un processus de Markov. Les marches aléatoires Markoviennes ou les fonctionnelles additives d'un processus de Markov sont des exemples de processus additifs Markoviens. Dans cet article, on montre que le processus $(Y_t)_{t \in \mathbb{T}}$ satisfait les théorèmes limites classiques suivants :

(a) le théorème de la limite centrale,
(b) le théorème limite local,
(c) le théorème uniforme de Berry–Esseen en dimension un,
(d) le développement d'Edgeworth d'ordre un en dimension un,

pourvu que la condition de moment $\sup_{t \in (0,1] \cap \mathbb{T}} \mathbb{E}_{\pi,0}[|Y_t|^\alpha] < \infty$ soit satisfaite, avec l'ordre attendu α du cas indépendant (à un $\varepsilon > 0$ près pour (c) et (d)). Pour les énoncés (b) et (d), il faut ajouter une condition nonlattice comme dans le cas indépendant. Tous les résultats sont obtenus sous l'hypothèse d'un processus de Markov $(X_t)_{t \in \mathbb{T}}$ admettant une mesure de probabilité invariante π et possédant la propriété de trou spectral sur $\mathbb{L}^2(\pi)$ (c'est à dire, $(X_t)_{t \in \mathbb{N}}$ est ρ-mélangeante dans le cas du temps discret). Le cas où $(X_t)_{t \in \mathbb{T}}$ est non-stationnaire est brièvement abordé. Nous appliquons nos résultats pour obtenir une borne de Berry–Esseen pour les M-estimateurs associés aux chaînes de Markov ρ-mélangeantes.

1. Introduction

In this paper, we are concerned with the class of Markov Additive Processes (MAP). The discrete and continuous-time cases are considered so that the time parameter set $\mathbb{T}$ will denote $\mathbb{N}$ or $[0, +\infty)$. Let $\mathbb{X}$ be any set equipped by a σ-algebra $\mathcal{X}$ and let $\mathcal{B}(\mathbb{R}^d)$ be the Borel σ-algebra on $\mathbb{R}^d$ ($d \geq 1$). A (time homogeneous) MAP $(X_t, Y_t)_{t \in \mathbb{T}}$ is a (time homogeneous) Markov process with state space $\mathbb{X} \times \mathbb{R}^d$ and transition semigroup $(Q_t)_{t \in \mathbb{T}}$ satisfying: $\forall t \in \mathbb{T}$, $\forall (x, y) \in \mathbb{X} \times \mathbb{R}^d$, $\forall (A, B) \in \mathcal{X} \times \mathcal{B}(\mathbb{R}^d)$,

$$Q_t(x, y; A \times B) = Q_t(x, 0; A \times B - y). \tag{1.1}$$

In other words, the transition semigroup is additive in the second component. It follows from the definition that the first component $(X_t)_{t \in \mathbb{T}}$ of a MAP is a (time homogeneous) Markov process. The second component $(Y_t)_{t \in \mathbb{T}}$ must be thought of as a process with independent increments given $\sigma(X_s, s \geq 0)$. We refer to [15] for the general structure of such processes. Note that a discrete-time MAP is also called a Markov Random Walk (MRW). In stochastic modelling, the first component of a MAP is usually associated with a random environment which drives or modulates the additive component $(Y_t)_{t \in \mathbb{T}}$. The MAPs have been found to be an important tool in various areas as communication networking (e.g. see [2,71,72]), finance (e.g. see [1,3,56]), reliability (e.g. see [17,37,64,70]), ... Some important instances of MAP are:

- In discrete/continuous-time: $(X_t, Y_t)_{t \in \mathbb{T}}$ where $(Y_t)_{t \in \mathbb{T}}$ is a $\mathbb{R}^d$-valued additive functional (AF) of the Markov process $(X_t)_{t \in \mathbb{T}}$. Therefore any result on the second component of a MAP applies to an AF. Basic discrete and continuous-time AFs are respectively

$$Y_0 = 0, \forall t \in \mathbb{N}^* \quad Y_t = \sum_{k=1}^{t} \xi(X_k); \qquad \forall t \in [0, +\infty[\quad Y_t = \int_0^t \xi(X_s) \, ds, \tag{1.2}$$

where ξ is a $\mathbb{R}^d$-valued function satisfying conditions under which Y_t is well-defined for every $t \in \mathbb{T}$. When $(X_t)_{t \in \mathbb{T}}$ is a regular Markov jump process, then any non-decreasing AF has the form (e.g. [16])

$$\int_0^t \xi_1(X_s) \, ds + \sum_{s \leq t} \xi_2(X_{s-}, X_s),$$

where $X_{t-} = \lim_{s \to t, s < t} X_s$, ξ_1 and ξ_2 are non-negative measurable functions such that $\xi_2(x, x) = 0$ for every $x \in \mathbb{X}$. General representations and properties of AFs may be found in [5,77], and references therein. Such AFs are basically introduced when some kind of "rewards" are collected along with the dynamics of the Markov process $(X_t)_{t \in \mathbb{T}}$ through the state space $\mathbb{X}$. Thus, Y_t is the accumulated reward on the finite interval $[0, t]$. Even if the state space $\mathbb{X}$ is a finite set, the numerical computation of the probability distribution of such AFs is not an easy task (e.g. see [9,82]).
- In discrete-time: the Markov renewal processes when the random variables Y_t, $t \in \mathbb{N}$, are non-negative; if we consider a hidden Markov chain $(X_t, Z_t)_{t \in \mathbb{N}}$, where the so-called observed process $(Z_t)_{t \in \mathbb{N}}$ is $\mathbb{R}^d$-valued ($Z_0 = 0$), then $(X_t, \sum_{k=1}^{t} Z_k)_{t \in \mathbb{N}}$ is a MAP.
- In continuous time: the Markovian Arrival Process where $(X_t)_{t \in \mathbb{T}}$ is a regular jump process and $(Y_t)_{t \in \mathbb{T}}$ is a point process (see [2]), which includes the so-called Markov Modulated Poisson Process.

Seminal works on MAPs are [21,22,59,69,75] and are essentially concerned with a finite Markov process $(X_t)_{t \in \mathbb{T}}$ as first component. The second component $(Y_t)_{t \in \mathbb{T}}$ was sometimes called a process defined on a Markov process. When $\mathbb{X}$ is a finite set, the structure of MAPs are well understood and an account of what is known can be found in [2], Chapter XI. In this paper, we are concerned with Gaussian approximations of the distribution of the second component Y_t of a MAP. Central limit theorems for $(Y_t)_{t \in \mathbb{T}}$ may be found in [7,27,30,50,51,59,61,75,83,84] under various assumptions. Here, such results are derived when $(X_t)_{t \in \mathbb{T}}$ has an invariant probability measure π, is stationary

and has the $\mathbb{L}^2(\pi)$-spectral gap property (see conditions (AS1) and (AS2) below). Moreover, standard refinements of the central limit theorem (CLT) related to the convergence rate are provided. Before, notations and assumptions used throughout the paper are introduced.

Let $(X_t, Y_t)_{t \in \mathbb{T}}$ be a MAP with state space $\mathbb{X} \times \mathbb{R}^d$ and transition semigroup $(Q_t)_{t \in \mathbb{T}}$. $(\mathbb{X}, \mathcal{X})$ is assumed to be a measurable space equipped with a σ-algebra $\mathcal{X}$. In the continuous-time case, $(X_t, Y_t)_{t \in \mathbb{T}}$ is assumed to be progressively measurable. $(X_t)_{t \in \mathbb{T}}$ is also a Markov process with transition semigroup $(P_t)_{t \in \mathbb{T}}$ given by

$$P_t(x, A) := Q_t(x, 0; A \times \mathbb{R}^d).$$

Throughout the paper, we assume that $(X_t)_{t \in \mathbb{T}}$ has a unique invariant probability measure denoted by π ($\forall t \in \mathbb{T}, \pi \circ P_t = \pi$). We denote by $\mathbb{L}^2(\pi)$ the usual Lebesgue space of (classes of) functions $f : \mathbb{X} \to \mathbb{C}$ such that $\|f\|_2 := \sqrt{\pi(|f|^2)} = (\int_{\mathbb{X}} |f|^2 \, d\pi)^{1/2} < \infty$. The operator norm of a bounded linear operator T on $\mathbb{L}^2(\pi)$ is defined by $\|T\|_2 := \sup_{\{f \in \mathbb{L}^2(\pi) : \|f\|_2 = 1\}} \|T(f)\|_2$. We appeal to the following conditions.

AS1. *$(X_t)_{t \in \mathbb{T}}$ is stationary (i.e. $X_0 \sim \pi$).*

AS2. *The semigroup $(P_t)_{t \in \mathbb{T}}$ of $(X_t)_{t \in \mathbb{T}}$ has a spectral gap on $\mathbb{L}^2(\pi)$:*

$$\lim_{t \to +\infty} \|P_t - \Pi\|_2 = 0, \tag{1.3}$$

where Π denotes the rank-one projection defined on $\mathbb{L}^2(\pi)$ by: $\Pi f = \pi(f) 1_{\mathbb{X}}$.

AS3. *The process $(Y_t)_{t \in \mathbb{T}}$ satisfies the moment condition*

$$\sup_{t \in (0,1] \cap \mathbb{T}} \mathbb{E}_{\pi,0}\big[|Y_t|^{\alpha}\big] < \infty, \tag{1.4}$$

where $|\cdot|$ denotes the euclidean norm on $\mathbb{R}^d$ and $\mathbb{E}_{\pi,0}$ is the expectation when $(X_0, Y_0) \sim (\pi, \delta_0)$.

In the discrete-time case, notice that the moment condition (1.4) reduces to (AS3d)

$$\mathbb{E}_{\pi,0}\big[|Y_1|^{\alpha}\big] < \infty \tag{AS3d}$$

and that condition (AS2) is equivalent to the ρ-mixing property of $(X_t)_{t \in \mathbb{N}}$, with ρ-mixing coefficients going to 0 exponentially fast [81]. Condition (AS2) is also related to the notion of essential spectral radius (e.g. see [86]).

Under (AS1) and (AS2), we show that the second component $(Y_t)_{t \in \mathbb{T}}$ of the MAP satisfies, in discrete and continuous time, the following standard limit theorems:

(a) *the central limit theorem, under (AS3) with the optimal value $\alpha = 2$;*
(b) *the local limit theorem, under (AS3) with the optimal value $\alpha = 2$ and the additional classical Markov non-lattice condition;*
(c) *the one-dimensional Berry–Esseen theorem, under (AS3) with the (almost) optimal value ($\alpha > 3$);*
(d) *a one-dimensional first-order Edgeworth expansion, under (AS3) with the (almost) optimal value ($\alpha > 3$) and the Markov non-lattice condition.*

These results correspond to the classical statements for the sequences of independent and identically distributed (i.i.d.) random variables, with the same order α (up to $\varepsilon > 0$ in (c) and (d)). Such results are known for special MAPs satisfying (AS2) (comparison with earlier works is made after each statement), but to the best of our knowledge, the results (a)–(d) are new for general MAPs satisfying (AS2), as, for instance, for AF involving unbounded functionals.

Here, the main arguments are:

• for the statement (a): the ρ-mixing property of the increments $(Y_{t+1} - Y_t)_{t \in \mathbb{T}}$ of the process $(Y_t)_{t \in \mathbb{T}}$ (see Proposition 3.1). This result, which has its own interest, is new to the best of our knowledge. The closest work to this part is a result of [38] which, by using ϕ-mixing properties, gives the CLT for MAPs associated with uniformly ergodic driving Markov chains (i.e. $(P_t)_{t \in \mathbb{T}}$ has a spectral gap on the usual Lebesgue space $\mathbb{L}^{\infty}(\pi)$). Condition (AS2) is less restrictive than uniform ergodicity (which is linked to the so-called Doeblin condition).

• For the refinements (b)–(d): the Nagaev–Guivarc'h spectral method. The closest works to this part are, in discrete-time the paper [49] in which these refinements are obtained for the AF: $Y_t = \sum_{k=1}^{t} \xi(X_k)$, and in continuous-time the work of Lezaud [62] which proves, under the uniform ergodicity assumption, a Berry–Esseen bound for the integral additive functional (1.2). Here, in discrete-time, we borrow to a large extent the weak spectral method of [49]: this is outlined in Proposition 4.2, which gives a precise expansion (close to the i.i.d. case) of the characteristic function of Y_t. For continuous-time MAPs, similar expansions can be derived from the semigroup property of the Fourier operators of the MAP. Proposition 4.2, and its continuous-time counterpart Proposition 4.4, are the key results to establish limit theorems (as for instance the statements (b)–(d)) with the help of Fourier techniques.

The classical (discrete and continuous-time) models for which the spectral gap property (AS2) is met, are briefly reviewed in Sections 2.2–2.4. The above limit theorems (a)–(d) are valid in all these examples and open up possibilities for new applications. First, our moment conditions are optimal (or almost optimal). For instance, in continuous time, the Berry–Esseen bound in [62] requires that ξ in the integral (1.2) is bounded, while our statement (c) holds true under the condition $\pi(|\xi|^{3+\varepsilon}) < \infty$. Second, our results are true for general MAPs. For instance, they apply to $Y_t = \sum_{k=1}^{t} \xi(X_{k-1}, X_k)$. This fact enables us to prove a Berry–Esseen bound for M-estimators associated with ρ-mixing Markov chains, under a moment condition which greatly improves the results in [76].

The paper is organised as follows. The $\mathbb{L}^2(\pi)$-spectral gap assumption for a Markov process is briefly discussed in Section 2 and connections to standard ergodic properties are pointed out. In Section 3, the CLT for $(Y_t)_{t\in\mathbb{T}}$ under (AS1)–(AS3) with $\alpha = 2$ is derived. The functional central limit theorem (FCLT) is also discussed. Section 4 is devoted to refinements of the CLT. First, the Fourier operator is introduced in Section 4.1, the characteristic function of Y_t is investigated in Section 4.2, and our limit theorems are proved for discrete-time MAPs in Section 4.3. Their extension to the non-stationary case is discussed in Section 4.4. The continuous-time case is studied in Section 4.5. The statistical application to M-estimators for ρ-mixing Markov chains is developed in Section 5.

Finally, we point out that the natural way to consider the Nagaev–Guivarc'h method in continuous-time is the semigroup property of the Fourier operators of the MAP (see Section 4.1 for details). To the best of our knowledge, this property, which is closely related to the additivity condition (1.1) defining a MAP, has been introduced and only exploited in [50].

2. The $\mathbb{L}^2(\pi)$-spectral gap property (AS2)

2.1. *Basic facts on property* (AS2)

We discuss the condition (AS2) for the semigroup $(P_t)_{t\in\mathbb{T}}$ of $(X_t)_{t\in\mathbb{T}}$. It is well-known that $(P_t)_{t\in\mathbb{T}}$ is a contraction semigroup on each Lebesgue-space $\mathbb{L}^p(\pi)$ for $1 \le p \le +\infty$, that is: we have $\|P_t\|_p \le 1$ for all $t \in \mathbb{T}$, where $\|\cdot\|_p$ denotes the operator norm on $\mathbb{L}^p(\pi)$. Condition (AS2), introduced by Rosenblatt [81] and also called strong ergodicity on $\mathbb{L}^2(\pi)$, implies that $(P_t)_{t\in\mathbb{T}}$ is strongly ergodic on each $\mathbb{L}^p(\pi)$ ($1 < p < +\infty$), that is $\|P_t - \Pi\|_p \to 0$ when $t \to +\infty$. Moreover, (AS2) is fulfilled under the so-called uniform ergodicity property, i.e. the strong ergodicity on $\mathbb{L}^\infty(\pi)$. These properties, established in [81], can be easily derived from the Riesz–Thorin interpolation theorem [6] which insures, thanks to the contraction property of P_t, that

$$\|P_t - \Pi\|_p \le \|P_t - \Pi\|_{p_1}^\alpha \|P_t - \Pi\|_{p_2}^{1-\alpha} \le 2\min\{\|P_t - \Pi\|_{p_1}^\alpha, \|P_t - \Pi\|_{p_2}^{1-\alpha}\}, \tag{2.1}$$

where $p_1, p_2 \in [1, +\infty]$ and $p \in [1, +\infty]$ satisfy $1/p = \alpha/p_1 + (1-\alpha)/p_2$ for some $\alpha \in [0, 1]$. Indeed, assume that condition (AS2) holds. Then inequality (2.1) with ($p_1 = 2, p_2 = +\infty$) and $\alpha \in (0, 1)$ gives the strong ergodicity on $\mathbb{L}^p(\pi)$ for each $p \in (2, +\infty)$. Notice that the value $p = +\infty$ is obtained with $\alpha = 0$, but in this case, the uniform ergodicity cannot be deduced from (AS2) and (2.1). In fact the uniform ergodicity condition is stronger than (AS2) (see [81]). Next inequality (2.1) with ($p_1 = 2, p_2 = 1$) and $\alpha \in (0, 1)$ gives the strong ergodicity on $\mathbb{L}^p(\pi)$ for each $p \in (1, 2)$. The value $p = 1$ is obtained with $\alpha = 0$, but the strong ergodicity on $\mathbb{L}^1(\pi)$ cannot be deduced from (AS2) and (2.1). Finally, if the uniform ergodicity is assumed, then inequality (2.1) with ($p_1 = +\infty, p_2 = 1$) and $\alpha = 1/2$ yields (AS2).

Also notice that the strong ergodicity property on $\mathbb{L}^p(\pi)$ holds if and only if there exists some strictly positive constants C and ε such that we have for all $t \in \mathbb{T}$:

$$\|P_t - \Pi\|_p \le Ce^{-\varepsilon t}. \tag{2.2}$$

Indeed, if $\kappa_0 := \|P_\tau - \Pi\|_p < 1$ for some $\tau \in \mathbb{T}$ (which holds under the strong ergodicity property), then we have for all $n \in \mathbb{N}^*$: $\|P_{n\tau} - \Pi\|_p = \|P_\tau^n - \Pi\|_p = \|(P_\tau - \Pi)^n\|_p \leq \kappa_0^n$. Writing $t = w + n\tau$ with $n \in \mathbb{N}^*$ and $w \in [0, \tau)$, we obtain: $\|P_t - \Pi\|_p = \|P_w(P_\tau^n - \Pi)\|_p \leq \kappa_0^n \leq Ce^{-\varepsilon t}$ with $C := 1/\kappa_0$ and $\varepsilon := (-1/\tau)\ln\kappa_0$. The converse implication is obvious. Thus, the strong ergodicity property on $\mathbb{L}^2(\pi)$, i.e. condition (AS2), is equivalent to require $\mathbb{L}^2(\pi)$-exponential ergodicity (2.2), that is the $\mathbb{L}^2(\pi)$-spectral gap property.

In the next subsection, Markov models with a spectral gap on $\mathbb{L}^2(\pi)$ arising from stochastic modelling and potentially relevant to our framework are introduced. Assumption (AS2) can be also met in more abstract settings, as for instance in [41] where the $\mathbb{L}^2$-spectral gap property for classic Markov operators (with a state space defined as the d-dimensional torus) is proved.

2.2. *Geometric ergodicity and property* (AS2)

Recall that $(X_t)_{t\in\mathbb{N}}$ is V-geometrically ergodic if its transition kernel P has an invariant probability measure π and is such that there are $r \in (0, 1)$, a finite constant K and a π-a.e finite function $V : \mathbb{X} \mapsto [1, +\infty]$ such that

$$\forall n \geq 0, \pi\text{-a.e. } x \in \mathbb{X}, \quad \sup\{|P^n f(x) - \pi(f)|, f : \mathbb{X} \to \mathbb{C}, |f| \leq V\} \leq KV(x)r^n. \tag{VG}$$

In fact, when $(X_t)_{t\in\mathbb{N}}$ is ψ-irreducible (i.e. $\psi(A) > 0 \Longrightarrow P(x, A) > 0, \forall x \in \mathbb{X}$) and aperiodic [67], condition (VG) is equivalent to the standard geometric ergodicity property [78]: there are functions $r : \mathbb{X} \to (0, 1)$ and $C : \mathbb{X} \mapsto [1, +\infty)$ such that: for all $n \in \mathbb{N}$, π-a.e. $x \in \mathbb{X}$,

$$\left\|P^n(x, \cdot) - \pi(\cdot)\right\|_{\text{TV}} := \sup\{|P^n f(x) - \pi(f)|, f : \mathbb{X} \to \mathbb{C}, |f| \leq 1\} \leq C(x)r(x)^n.$$

There is another equivalent operational condition to geometric ergodicity for ψ-irreducible and aperiodic Markov chains $(X_t)_{t\in\mathbb{N}}$, the so-called "drift-criterion": there exist a function $V : \mathbb{X} \to [1, +\infty]$, a small set $C \subset \mathbb{X}$ and constants $\delta > 0, b < \infty$ such that

$$PV \leq (1 - \delta)V + b1_C.$$

We refer to [67] for details and applications, and to [57] for a recent survey on the CLT for the additive functionals of $(X_t)_{t\in\mathbb{N}}$ in (1.2). Now, the transition kernel P is said to be reversible with respect to π if

$$\pi(\mathrm{d}x)P(x, \mathrm{d}y) = \pi(\mathrm{d}y)P(y, \mathrm{d}x)$$

or equivalently if P is self-adjoint on the space $\mathbb{L}^2(\pi)$. It is well known that a V-geometrically ergodic Markov chain with a reversible transition kernel has the $\mathbb{L}^2(\pi)$-spectral gap property [78]. Moreover, for a ψ-irreducible and aperiodic Markov chain $(X_t)_{t\in\mathbb{N}}$ with reversible transition kernel, (V-)geometric ergodicity is shown to be equivalent to the existence of a spectral gap in $\mathbb{L}^2(\pi)$, and, when $X_0 \sim \mu$, we also have [78], Theorem 2.1, [80]

$$\left\|\mu P^n(\cdot) - \pi(\cdot)\right\|_{\text{TV}} \leq \frac{1}{2}|\mu - \pi|_{L^2(\pi)}r^n, \tag{R}$$

where $r := \lim_{n\to+\infty}(\|P^n - \Pi\|_2)^{1/n}$ and $|\mu - \pi|_{L^2(\pi)} := \|\mathrm{d}\mu/\mathrm{d}\pi - 1\|_2$ if well-defined and ∞ otherwise. Note that the reversibility condition is central to the previous discussion on the $\mathbb{L}^2(\pi)$-spectral gap property. Indeed, there exists a ψ-irreducible and aperiodic Markov chain which is geometrically ergodic but does not admit a spectral gap on $\mathbb{L}^2(\pi)$ [43].

Such a context of geometric ergodicity and reversible kernel is relevant to the Markov Chain Monte Carlo methodology for sampling a given probability distribution, i.e. the target distribution. Indeed, the basic idea is to define a Markov chain $(X_t)_{t\in\mathbb{N}}$ with the target distribution as invariant probability measure π. Then a MCMC algorithm is a scheme to draw samples from the stationary Markov chain $(X_t)_{t\in\mathbb{N}}$. But, the initial condition of the algorithm, i.e. the probability distribution of X_0, is not π since the target distribution is inaccessible. Therefore the convergence in distribution of the Markov chain to π in regard of the probability distribution of X_0 must be guaranteed and the knowledge of the convergence rate is crucial to monitor the sampling. Thus, central limit theorem for the Markov

chains and quantitative bounds as in (R) are highly expected. Geometric ergodicity of Hasting–Metropolis type algorithms has been investigated by many researchers. Two standard instances are the full dimensional and random-scan symmetric random walk Metropolis algorithm [25,55], and references therein. Note that the first algorithm is also referred to as a special instance of the Hasting algorithm and the second one to as a Metropolis-within-Gibbs sampler. Let π be a probability distribution on $\mathbb{R}^d$ which is assumed to have a positive and continuous density with respect to the Lebesgue measure. The so-called proposal densities are assumed to be bounded away from 0 in some region around zero (the moves through the state space $\mathbb{X}$ are based on these probability distributions). These conditions assert that the corresponding transition kernel for each algorithm is ψ-irreducible, aperiodic and is reversible with respect to π. Geometric ergodicity for the Markov chain $(X_t)_{t\in\mathbb{N}}$ (and so the existence of a spectral gap in $\mathbb{L}^2(\pi)$) is closely related to the tails of the target distribution π. For instance, in the first algorithm, it can be shown that π must have an exponential moment [55], Corollary 3.4. A sufficient condition for geometric ergodicity in case of super-exponential target densities, is of the form [55], Theorem 4.1,

$$\lim_{|x|\to+\infty}\left\langle \frac{x}{|x|}, \frac{\nabla\pi(x)}{|\nabla\pi(x)|}\right\rangle < 0.$$

For the second algorithm, sufficient conditions for geometric ergodicity are reported in [25] when the target density decreases either subexponentially or exponentially in the tails. A very large set of examples and their respective merit are discussed in these two references. We refer to [79], and references therein, for a recent survey on the theory of Markov chains in connection with MCMC algorithms.

2.3. *Uniform ergodicity and hidden Markov chains*

As quoted in the introduction, a discrete-time MAP is closely related to a hidden Markov chain. Standard issues for hidden Markov chains require to be aware of the convergence rate of the hidden Markov state process $(X_t)_{t\in\mathbb{N}}$. One of them is the state estimation via filtering or smoothing. In such a context, minorization conditions on P are usually involved. The basic one is: there exists a bounded positive measure φ on $\mathbb{X}$ such that for some $m\in\mathbb{N}^*$:

$$\forall x\in\mathbb{X}, \forall A\in\mathcal{X} \quad P^m(x,A)\geq\varphi(A). \tag{UE}$$

It is well-known that this is equivalent to the uniform ergodicity property or to condition (VG) with $V(x)=1$ [67], Theorems 16.2.1 and 16.2.2. Recall that uniform ergodicity gives the $\mathbb{L}^2(\pi)$-spectral gap property (AS2), but the converse is not true. Another minorization condition is the so-called "Doeblin condition": there exists a probability measure φ such that for some m, $\varepsilon < 1$ and $\delta > 0$ [20]

$$\varphi(A) > \varepsilon \quad\Longrightarrow\quad \forall x\in\mathbb{X}, \quad P^m(x,A)\geq\delta. \tag{D_0}$$

It is well known that, for ergodic and aperiodic Markov chains, (D_0) is equivalent to the uniform ergodicity. We refer to [14], and the references therein, for an excellent overview of the interplay between the Markov chain theory and the hidden Markov models.

2.4. *Property* (AS2) *for continuous time Markov processes*

The Markov jump processes are a basic class of continuous-time Markov models which has a wide interest in stochastic modelling. The $\mathbb{L}^2(\pi)$-exponential convergence has received attention a long time ago. We refer to [18] for a good account of what is known on ergodic properties for such processes. In particular, the $\mathbb{L}^2(\pi)$-spectral gap property is shown to be equivalent to the standard exponential ergodicity for the birth–death processes:

$$\exists\beta > 0 \text{ such that } \forall(i,j)\in\mathbb{X}^2, \exists C_i\geq 0, \quad \left|P_t(i,j)-\pi_j\right|\leq C_i\exp(-\beta t), \quad t\to+\infty,$$

where $(P_t(i,j))_{i,j\in\mathbb{X}}$ is the matrix semigroup of $(X_t)_{t\geq 0}$. This is also true for the reversible Markov jump processes. Hence, in these cases, criteria for exponential ergodicity are also valid to check the $\mathbb{L}^2(\pi)$-exponential convergence. Moreover, explicit bounds on the spectral gap are discussed in details in [18]. For the birth–death processes, we also

refer to [58], and references therein, where explicit formulas are obtained for classical Markov queuing processes. The birth–death processes are often used as reference processes for analyzing general stochastic models. This idea was in force in the Liggetts's derivation of the $\mathbb{L}^2$-exponential convergence of supercritical nearest particle systems [63]. The interacting systems of particles are also a source of examples of processes with a $\mathbb{L}^2$-spectral gap. We refer to [63] for such a discussion on various classes of stochastic Ising models. In physics and specially in statistical physics, many evolution models are given by stochastic ordinary/partial equations. When the solutions are finite/infinite dimensional Markov processes, standard issues arise: existence and uniqueness of an invariant probability measure, ergodic properties which include the rate of convergence to the invariant measure with respect to some norm. Such issues may be included in the general topic of the stability of solutions of stochastic differential equations (SDEs). Thus, it is not surprising that ergodic concepts as the V-geometric ergodicity and Lyapunov-type criteria associated with, originally developed by Meyn and Tweedie [67] for studying the stability of discrete-time Markov models, have been found to be of value (e.g. see [32], and references therein). Here, we are only concerned with the $\mathbb{L}^2(\pi)$-exponential convergence so that we only mention some results related with.

An instance of $\mathbb{L}^2(\pi)$-spectral gap can be found in [28] where the following SDE is considered

$$\mathrm{d}X_t = -\frac{1}{2}b(X_t)\,\mathrm{d}t + \mathrm{d}W_t, \quad X_0 = x \in \mathbb{R}^d,$$

where $(W_t)_{t\geq 0}$ is the standard d-dimensional Brownian motion and $b(\cdot)$ is a gradient field from $\mathbb{R}^d$ to $\mathbb{R}^d$ (with suitable properties ensuring essentially the existence of a unique strong solution to the equation, which has a unique invariant probability measure). When $b(\cdot)$ is a radial function satisfying $b(x) \sim C|x|^\alpha$ for $\alpha > 1$ when $x \to +\infty$, then the semigroup is shown to be ultracontractive and to have a $\mathbb{L}^2(\pi)$-spectral gap [28].

Another instance of $\mathbb{L}^2(\pi)$-spectral gap is related to the $\mathbb{R}$-valued Markov process solution to

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + a(X_t)\,\mathrm{d}W_t, \tag{2.3}$$

where $(W_t)_{t\geq 0}$ is the standard 1-dimensional Brownian motion and X_0 is a random variable independent of $(W_t)_{t\geq 0}$. Standard assumptions ensure that the solution of the SDE above is a positive recurrent diffusion on some interval and a (strictly) stationary ergodic time-reversible process. Under additional conditions on the scale and the speed densities of the diffusion $(X_t)_{t\geq 0}$ [29], (A4) and reinforced (A5), Proposition 2.8, the transition semigroup of $(X_t)_{t\geq 0}$ is shown to have the $\mathbb{L}^2(\pi)$-spectral gap property (explicit bounds on the spectral gap are also provided). The basic example studied in [29] is when $a(x) := cx^\nu$ and $b(x) := \alpha(\beta - x)$ with $\nu \in [1/2, 1]$, $\alpha, \beta \in \mathbb{R}$. Conditions ensuring the $\mathbb{L}^2(\pi)$-spectral gap property are provided in terms of these parameters. Applications to some classical models in finance are discussed. Note that statistical issues for continuous-time Markov processes as the jump or diffusion processes, are related to the time discretization or sampling schemes of these processes. This often provides discrete-time Markov chains which inherit ergodic properties of the original continuous-time process. Thus we turn to the discussion on the discrete-time case (e.g. see [19] for the jump processes, [29] and the references therein for the (hidden) diffusions). Finally, the context of the stochastic differential equation (2.3) can be generalized to Markov H-valued processes solution to infinite dimensional SDEs, where H is a Hilbert space. A good account of these generalizations can be found in [33], and references therein.

3. The ρ-mixing property and central limit theorems

Let $(X_t, Y_t)_{t\in\mathbb{T}}$ be a MAP taking values in $\mathbb{X} \times \mathbb{R}^d$. $\mathbb{E}_{(x,0)}$, $\mathbb{E}_{\pi,0}$ are the expectation with respect to the initial conditions $(X_0, Y_0) \sim (\delta_x, \delta_0)$ and $(X_0, Y_0) \sim (\pi, \delta_0)$ respectively. First, basic facts for MAPs are proposed. Second, they are used to show that, for a discrete-time MAP, the increment process $(Y_n - Y_{n-1})_{n\in\mathbb{N}^*}$ is exponentially ρ-mixing under (AS1) and (AS2). Then, a CLT is obtained under conditions (AS1) and (AS2) and the expected moment condition (AS3) (i.e. (AS3d)) with $\alpha = 2$.

3.1. *Basic facts on MAPs*

Let $\mathbb{F}_t^{(X,Y)} := \sigma(X_u, Y_u, u \leq t)$, $\mathbb{F}_t^X := \sigma(X_u, u \leq t)$ and $\mathbb{F}_t^Y := \sigma(Y_u, u \leq t)$ be the filtration generated by the processes $(X_t, Y_t)_{t\in\mathbb{T}}$, $(X_t)_{t\in\mathbb{T}}$ and $(Y_t)_{t\in\mathbb{T}}$ respectively.

The additivity property (1.1) for the semigroup $(Q_t)_{t \in \mathbb{T}}$ reads as follows for any measurable ($\mathbb{C}$-valued) function g on $\mathbb{X} \times \mathbb{R}^d$ and any $a \in \mathbb{R}^d$:

$$Q_t(g)_a = Q_t(g_a), \tag{3.1}$$

where $g_a(x, y) := g(x, y + a)$ for every $(x, y) \in \mathbb{X} \times \mathbb{R}^d$. Let us introduce the following notation:

$$\widetilde{Q}_s(x; \mathrm{d}x_1 \times \mathrm{d}y_1) := Q_s(x, 0; \mathrm{d}x_1 \times \mathrm{d}y_1).$$

Then, we have:

Lemma 3.1. *For any $\mathbb{C}$-valued function g on $\mathbb{X} \times \mathbb{R}^d$ such that $\mathbb{E}[|g(X_u, Y_u)|] < \infty$ for every $u \in \mathbb{T}$, we have*:

$$\mathbb{E}\big[g(X_{s+t}, Y_{s+t})|\mathcal{F}_s^{(X,Y)}\big] = Q_t(g_{Y_s})(X_s, 0) = \widetilde{Q}_t(g_{Y_s})(X_s). \tag{3.2}$$

or in terms of the increments of the process $(Y_t)_{t \in \mathbb{T}}$:

$$\mathbb{E}\big[g(X_{s+t}, Y_{s+t} - Y_s)|\mathcal{F}_s^{(X,Y)}\big] = Q_t(g)(X_s, 0) = \widetilde{Q}_t(g)(X_s) = \mathbb{E}_{(X_s,0)}\big[g(X_t, Y_t)\big]. \tag{3.3}$$

Proof. The two formula are derived as follows:

$$\begin{aligned}
\mathbb{E}\big[g(X_{s+t}, Y_{s+t})|\mathcal{F}_s^{(X,Y)}\big] &= \mathbb{E}\big[g(X_{s+t}, Y_{s+t})|X_s, Y_s\big] \quad \text{(Markov property)} \\
&= Q_t(g)(X_s, Y_s) \\
&= Q_t(g_{Y_s})(X_s, 0) \quad \text{(from (3.1))} \\
&= \widetilde{Q}_t(g_{Y_s})(X_s); \\
\mathbb{E}\big[g(X_{s+t}, Y_{s+t} - Y_s)|\mathcal{F}_s^{(X,Y)}\big] &= \mathbb{E}\big[g(X_{s+t}, Y_{s+t} - Y_s)|X_s, Y_s\big] \quad \text{(Markov property)} \\
&= \mathbb{E}\big[g_{-Y_s}(X_{s+t}, Y_{s+t})|X_s, Y_s\big] \\
&= Q_t(g_0)(X_s, 0) = \widetilde{Q}_t(g)(X_s) \quad \text{(from (3.2))} \\
&= \mathbb{E}_{(X_s,0)}\big[g(X_t, Y_t)\big]. \qquad \square
\end{aligned}$$

Lemma 3.2. *For every $n \geq 1$, any $\mathbb{C}$-valued function g such that for every $0 \leq u_1 \leq \cdots \leq u_n$*

$$\mathbb{E}\big[\big|g(X_{u_1}, Y_{u_1}, X_{u_2}, Y_{u_2} - Y_{u_1}, \ldots, X_{u_n}, Y_{u_n} - Y_{u_{n-1}})\big|\big] < \infty$$

we have for any $s \geq 0$ and $t_1, \ldots, t_n \geq 0$:

$$\begin{aligned}
&\mathbb{E}\big[g(X_{s+t_1}, Y_{s+t_1} - Y_s, \ldots, X_{s+\sum_{i=1}^n t_i}, Y_{s+\sum_{i=1}^n t_i} - Y_{s+\sum_{i=1}^{n-1} t_i})|\mathcal{F}_s^{(X,Y)}\big] \\
&= \int \widetilde{Q}_s(X_s; \mathrm{d}x_1 \times \mathrm{d}z_1) \prod_{i=2}^n \widetilde{Q}_s(x_{i-1}; \mathrm{d}x_i \times \mathrm{d}z_i) g(x_1, z_1, \ldots, x_n, z_n) \\
&= \left(\bigotimes_{i=1}^n \widetilde{Q}_{t_i} \right)(g)(X_s). \tag{3.4}
\end{aligned}$$

Proof. Lemma 3.1 gives the case $n = 1$. Let us check that Formula (3.4) is valid for $n = 2$. This can help the reader to follow the induction.

$$\begin{aligned}
&\mathbb{E}\big[g(X_{s+t_1}, Y_{s+t_1} - Y_s, X_{s+t_1+t_2}, Y_{s+t_1+t_2} - Y_{s+t_1})|\mathcal{F}_s^{(X,Y)}\big] \\
&= \mathbb{E}\big[\mathbb{E}\big[g(X_{s+t_1}, Y_{s+t_1} - Y_s, X_{s+t_1+t_2}, Y_{s+t_1+t_2} - Y_{s+t_1})|\mathcal{F}_{s+t_1}^{(X,Y)}\big]|\mathcal{F}_s^{(X,Y)}\big]
\end{aligned}$$

$$= \mathbb{E}\big[\mathbb{E}\big[g(X_{s+t_1}, Y_{s+t_1} - Y_s, X_{s+t_1+t_2}, Y_{s+t_1+t_2} - Y_{s+t_1})|X_{s+t_1}, Y_{s+t_1}\big]|\mathcal{F}_s^{(X,Y)}\big]$$

$$= \mathbb{E}\bigg[\int g(X_{s+t_1}, Y_{s+t_1} - Y_s, x_2, y_2 - Y_{s+t_1})Q_{t_2}(X_{s+t_1}, Y_{s+t_1}; \mathrm{d}x_2 \times \mathrm{d}y_2)|\mathcal{F}_s^{(X,Y)}\bigg]$$

$$= \mathbb{E}\bigg[\int g(X_{s+t_1}, Y_{s+t_1} - Y_s, x_2, z_2)\widetilde{Q}_{t_2}(X_{s+t_1}; \mathrm{d}x_2 \times \mathrm{d}z_2)\big|\mathcal{F}_s^{(X,Y)}\bigg] \quad \text{(using (1.1))}$$

$$= \int \widetilde{Q}_{t_1}(X_s; \mathrm{d}x_1 \times \mathrm{d}z_1) \int \widetilde{Q}_{t_2}(x_1; \mathrm{d}x_2 \times \mathrm{d}z_2)g(x_1, z_1, x_2, z_2) \quad \text{(using (3.3))}$$

$$= (\widetilde{Q}_{t_1} \otimes \widetilde{Q}_{t_2})(g)(X_s).$$

Let us now complete the induction. Assume that Property (3.4) is valid for $n - 1$. Then

$$\mathbb{E}\big[g(X_{s+t_1}, Y_{s+t_1} - Y_s, \ldots, X_{s+\sum_{i=1}^{n} t_i}, Y_{s+\sum_{i=1}^{n} t_i} - Y_{s+\sum_{i=1}^{n-1} t_i})|\mathcal{F}_s^{(X,Y)}\big]$$

$$= \mathbb{E}\big[\mathbb{E}\big[g(X_{s+t_1}, Y_{s+t_1} - Y_s, \ldots, X_{s+\sum_{i=1}^{n} t_i}, Y_{s+\sum_{i=1}^{n} t_i} - Y_{s+\sum_{i=1}^{n-1} t_i})|\mathcal{F}_{s+t_1}^{(X,Y)}\big]|\mathcal{F}_s^{(X,Y)}\big]$$

$$= \mathbb{E}\bigg[\bigg(\bigotimes_{i=2}^{n} \widetilde{Q}_{t_i}\bigg)\big(g(X_{s+t_1}, Y_{s+t_1} - Y_s, \cdot, \ldots, \cdot)\big)(X_{s+t_1})\Big|\mathcal{F}_s^{(X,Y)}\bigg] \quad \text{(induction)}$$

$$= \bigg(\widetilde{Q}_{t_1} \otimes \bigg(\bigotimes_{i=2}^{n} \widetilde{Q}_{t_i}\bigg)\bigg)(g)(X_s) \quad \text{(using (3.3))}. \qquad\qquad\qquad \square$$

Corollary 3.1. *Under* (AS1), *the following properties hold.*

1. *The process* $(Y_t)_{t \in \mathbb{T}}$ *has stationary increments, i.e.*

$$\mathbb{E}_{\pi,0}\big[g(Y_{s+t_1} - Y_s, \ldots, Y_{s+\sum_{i=1}^{n} t_i} - Y_{s+\sum_{i=1}^{n-1} t_i})\big] = \pi\bigg(\bigg(\bigotimes_{i=1}^{n} \widetilde{Q}_{t_i}\bigg)(g)\bigg) \qquad (3.5)$$

   does not depend on s for any function g as in Lemma 3.2.

2. *If* $\mathbb{E}_{\pi,0}[|Y_u|] < \infty$ *for every* $u \in \mathbb{T}$, *then:*

$$\forall (s,t) \in \mathbb{T}^2 \quad \mathbb{E}_{\pi,0}[Y_{s+t}] = \mathbb{E}_{\pi,0}[Y_t] + \mathbb{E}_{\pi,0}[Y_s].$$

3. $(\xi_n := Y_n - Y_{n-1})_{n \in \mathbb{N}^*}$ *is a stationary sequence of* $\mathbb{R}^d$*-valued random variables and if h is a* $\mathbb{C}$*-valued function such that* $\mathbb{E}_{\pi,0}[|h(\xi_1, \ldots, \xi_n)|^2] = 1$, *then* $\widetilde{Q}_1^{\otimes n}(h) \in \mathbb{L}^2(\pi)$ *with*

$$\big\|\widetilde{Q}_1^{\otimes n}(h)\big\|_2 \leq 1, \qquad (3.6)$$

   where $\widetilde{Q}_1^{\otimes n}$ denotes the n-fold kernel product $\bigotimes_{i=1}^{n} \widetilde{Q}_1$.

Proof. Take the expectation of (3.4) with respect to the probability mesure π:

$$\mathbb{E}_{\pi,0}\big[g(Y_{s+t_1} - Y_s, \ldots, Y_{s+\sum_{i=1}^{n} t_i} - Y_{s+\sum_{i=1}^{n-1} t_i})\big]$$

$$= \mathbb{E}_{\pi,0}\bigg[\bigg(\bigotimes_{i=1}^{n} \widetilde{Q}_{t_i}\bigg)(g)(X_s)\bigg] = \mathbb{E}_{\pi P_s, 0}\bigg[\bigg(\bigotimes_{i=1}^{n} \widetilde{Q}_{t_i}\bigg)(g)(X_0)\bigg]$$

$$= \mathbb{E}_{\pi,0}\bigg[\bigg(\bigotimes_{i=1}^{n} \widetilde{Q}_{t_i}\bigg)(g)(X_0)\bigg] \quad \text{(invariance property of π).}$$

The second property is deduced from the stationarity of the increments of $(Y_t)_{t \in \mathbb{T}}$. Indeed, we can write $\mathbb{E}_{\pi,0}[Y_t] = \mathbb{E}_{\pi,0}[Y_{s+t} - Y_s] = \mathbb{E}_{\pi,0}[Y_{s+t}] - \mathbb{E}_{\pi,0}[Y_s]$. That $(\xi_n)_{n \in \mathbb{N}^*}$ is a stationary sequence of random variables follows from (3.5) with $s = 0$, $t_1 = \cdots = t_n = 1$. The last property follows from (3.4) and the Jensen inequality

$$\big\| \widetilde{Q}_1^{\otimes n} h \big\|_2^2 = \big\| \mathbb{E}_{(\cdot,0)} \big[h(\xi_1, \ldots, \xi_n) \big] \big\|_2^2 \leq \mathbb{E}_{\pi,0} \big[\big| h(\xi_1, \ldots, \xi_n) \big|^2 \big] = 1. \qquad \square$$

Lemma 3.3. *Let* $\xi_n := Y_n - Y_{n-1}$ *for* $n \in \mathbb{N}^*$ *(recall that* $Y_0 = 0$ *a.s.). Let* f *and* h *be two* $\mathbb{C}$*-valued functions such that* $\mathbb{E}_{\pi,0}[|f(\xi_1, \ldots, \xi_n)|^2] < \infty$ *and* $\mathbb{E}_{\pi,0}[|h(\xi_{n+t}, \ldots, \xi_{n+t+m})|^2] < \infty$ *for* $(t, n, m) \in (\mathbb{N}^*)^3$. *Under* (AS1), *the covariance has the following form*

$$\text{Cov}_{\pi,0} \big(f(\xi_1, \ldots, \xi_n); h(\xi_{n+t}, \ldots, \xi_{n+t+m}) \big) = \mathbb{E}_{\pi,0} \big[f(\xi_1, \ldots, \xi_n)(P_{t-1} - \Pi)\big(\widetilde{Q}_1^{\otimes m+1}(h)\big)(X_n) \big], \qquad (3.7)$$

where $\widetilde{Q}_1^{\otimes m+1}$ *denotes the* $(m+1)$*-fold kernel product* $\bigotimes_{i=1}^{m+1} \widetilde{Q}_1$.

Proof. Apply formula (3.4) with $\mathbb{E}_{(x,0)}$ to the specific function $g(x_1, z_1, \ldots, x_{n+t+m}, z_{n+t+m}) = f(z_1, \ldots, z_n) \times h(z_{n+t}, \ldots, z_{n+t+m})$ with $t, n, m \geq 1$:

$$\mathbb{E}_{(x,0)} \big[f(\xi_1, \ldots, \xi_n) h(\xi_{n+t}, \ldots, \xi_{n+t+m}) \big]$$

$$= \widetilde{Q}_1^{\otimes n+m+t}(g)(x)$$

$$= \int_{(\mathbb{X} \times \mathbb{R}^d)^{n+t+m}} \widetilde{Q}_1(x; \mathrm{d}x_1 \times \mathrm{d}z_1) \prod_{i=2}^{n+t+m} \widetilde{Q}_1(x_{i-1}; \mathrm{d}x_i \times \mathrm{d}z_i) g(x_1, \ldots, z_{n+t+m})$$

$$= \int_{(\mathbb{X} \times \mathbb{R}^d)^n} \widetilde{Q}_1(x; \mathrm{d}x_1 \times \mathrm{d}z_1) \prod_{i=2}^{n} \widetilde{Q}_1(x_{i-1}; \mathrm{d}x_i \times \mathrm{d}z_i) f(z_1, \ldots, z_n)$$

$$\times \int_{(\mathbb{X} \times \mathbb{R}^d)^{t-1}} \prod_{i=n+1}^{n+t-1} \widetilde{Q}_1(x_{i-1}; \mathrm{d}x_i \times \mathrm{d}z_i)$$

$$\times \int_{(\mathbb{X} \times \mathbb{R}^d)^{m+1}} \prod_{i=n+t}^{n+t+m} \widetilde{Q}_1(x_{i-1}; \mathrm{d}x_i \times \mathrm{d}z_i) h(z_{n+t}, \ldots, z_{n+t+m}).$$

The second term reduces to $\int_{\mathbb{X}^{t-1}} \prod_{i=n+1}^{n+t-1} \widetilde{Q}_1(x_{i-1}; \mathrm{d}x_i \times \mathbb{R}^d) = \int_{\mathbb{X}^{t-1}} \prod_{i=n+1}^{n+t-1} P_1(x_{i-1}; \mathrm{d}x_i) = \int_{\mathbb{X}} P_{t-1}(x_n; \mathrm{d}x_{n+t-1})$. The third is $\widetilde{Q}_1^{\otimes m+1}(h)(x_{n+t-1})$. Then we have

$$\mathbb{E}_{(x,0)}[f(\xi_1, \ldots, \xi_n) h(\xi_{n+t}, \ldots, \xi_{n+t+m})]$$

$$= \int_{(\mathbb{X} \times \mathbb{R}^d)^n} \widetilde{Q}_1(x; \mathrm{d}x_1 \times \mathrm{d}z_1) \prod_{i=2}^{n} \widetilde{Q}_1(x_{i-1}; \mathrm{d}x_i \times \mathrm{d}z_i) f(z_1, \ldots, z_n) P_{t-1}\big(\widetilde{Q}_1^{\otimes m+1}(h)\big)(x_n)$$

$$= \mathbb{E}_{(x,0)} \big[f(\xi_1, \ldots, \xi_n) P_{t-1}\big(\widetilde{Q}_1^{\otimes m+1} h \big)(X_n) \big] \quad \text{(using (3.4) with } \mathbb{E}_{(x,0)}\text{)}.$$

Then, integrating against the probability measure $\pi(\cdot)$ gives

$$\mathbb{E}_{\pi,0} \big[f(\xi_1, \ldots, \xi_n) h(\xi_{n+t}, \ldots, \xi_{n+t+m}) \big] = \mathbb{E}_{\pi,0} \big[f(\xi_1, \ldots, \xi_n) P_{t-1}\big(\widetilde{Q}_1^{\otimes m+1} h \big)(X_n) \big]. \qquad (3.8)$$

Since $\Pi(\widetilde{Q}_1^{\otimes m+1}(h))(x) = \pi(\widetilde{Q}_1^{\otimes m+1}(h))$ for every $x \in \mathbb{X}$, we obtain

$$\mathbb{E}_{\pi,0} \big[f(\xi_1, \ldots, \xi_n) \Pi\big(\widetilde{Q}_1^{\otimes m+1}(h)\big)(X_n) \big] = \mathbb{E}_{\pi,0} \big[f(\xi_1, \ldots, \xi_n) \big] \pi \big(\widetilde{Q}_1^{\otimes m+1}(h) \big)$$

$$= \mathbb{E}_{\pi,0} \big[f(\xi_1, \ldots, \xi_n) \big] \mathbb{E}_{\pi,0} \big[h(\xi_{n+t}, \ldots, \xi_{n+t+m}) \big],$$

where the last equality follows from (3.5). $\qquad \square$

Remark 3.1. *We can prove a continuous-time counterpart of Lemma* 3.3. *But, we restrict ourself to the discrete-time version because this is the version we need in the paper.*

3.2. ρ-mixing property of $(Y_n - Y_{n-1})_{n \geq 1}$ for discrete-time stationary MAPs

Let us recall some basic facts on the ρ-mixing of a (strictly) stationary sequence of random variables $(\xi_n)_{n \in \mathbb{N}^*}$ (e.g. see [11]). For any $p \in \mathbb{N}^*$ and $q \in \mathbb{N}^* \cup \{\infty\}$ with $p \leq q$, $\mathcal{G}_p^q := \sigma(\xi_p, \ldots, \xi_q)$ denotes the σ-algebra generated by random variables $\xi_p, \ldots, \xi_q$. The ρ-mixing coefficient at horizon $t > 0$, denoted by $\rho(t)$, is defined by

$$\rho(t) := \sup_{n \in \mathbb{N}^*} \sup\{|\mathrm{Corr}(f; h)| \, f \in \mathbb{L}^2(\mathcal{G}_1^n), h \in \mathbb{L}^2(\mathcal{G}_{n+t}^\infty)\}, \tag{3.9}$$

where $\mathrm{Corr}(f; h)$ is the correlation coefficient of the two random variables f and g. In fact, ρ-mixing coefficient may be computed as follows from [11], Proposition 3.18: for $t > 0$

$$\rho(t) = \sup_{n \in \mathbb{N}^*} \sup_{m \in \mathbb{N}^*} \sup\{|\mathrm{Corr}(f; h)| \, f \in \mathbb{L}^2(\mathcal{G}_1^n), h \in \mathbb{L}^2(\mathcal{G}_{n+t}^{n+t+m})\}. \tag{3.10}$$

The stationary sequence $(\xi_n)_{n \in \mathbb{N}^*}$ is said to be ρ-*mixing* if

$$\lim_{t \to +\infty} \rho(t) = 0.$$

We know from condition (AS2) that $(X_n)_{n \in \mathbb{N}}$ is ρ-mixing [81]. In the special case when $Y_n := \sum_{k=1}^n \xi(X_k)$, it is clear that $(Y_n - Y_{n-1})_{n \geq 1}$ is also ρ-mixing from [11], p. 28. We extend this fact to general $(Y_n)_{n \in \mathbb{N}}$ in the next proposition.

Proposition 3.1 (ρ-mixing). *Under* (AS1) *and* (AS2)*, the stationary sequence $(\xi_n := Y_n - Y_{n-1})_{n \in \mathbb{N}^*}$ is ρ-mixing at an exponential rate*: *there exists $\varepsilon > 0$ such that*

$$\rho(t) = \mathrm{O}\big(\exp(-\varepsilon t)\big).$$

Proof. For the sake of simplicity, assume that $d = 1$. First, note that the random variables f and h in (3.10) can be assumed to be of $\mathbb{L}^2$-norm 1. Thus, we just have to deal with covariances. Recall that $(\xi_n := Y_n - Y_{n-1})_{n \in \mathbb{N}^*}$ is known to be stationary under (AS1) from Corollary 3.1. The σ-algebra $\mathcal{G}_1^n$ and $\mathcal{G}_{n+t}^{n+t+m}$ in the mixing coefficients will be relative to the stationary sequence $(\xi_n)_{n \in \mathbb{N}^*}$. Second, let us consider two $\mathbb{L}^2$-normed random variables $f(\xi_1, \ldots, \xi_n) \in \mathbb{L}^2(\mathcal{G}_1^n)$, $h(\xi_{n+t}, \ldots, \xi_{n+t+m}) \in \mathbb{L}^2(\mathcal{G}_{n+t}^{n+t+m})$. For any $m \geq 1$, the map $x \mapsto \widetilde{Q}_1^{\otimes m+1}(h)(x)$ belongs to $\mathbb{L}^2(\pi)$ and we have $\|\widetilde{Q}_1^{\otimes m+1}(h)\|_2 \leq 1$ from Corollary 3.1 (see (3.6)). Since P_t and Π are contractions on $\mathbb{L}^2(\pi)$, we have $(P_{t-1} - \Pi)(\widetilde{Q}_1^{\otimes m+1}(h)) \in \mathbb{L}^2(\pi)$.

The Cauchy–Schwarz inequality and the last comments allow us to write from (3.7)

$$\begin{aligned}
\mathrm{Cov}(f; h)^2 &\leq \mathbb{E}_{\pi,0}\big[\big|f(\xi_1, \ldots, \xi_n)\big|^2\big] \mathbb{E}_{\pi,0}\big[\big|(P_{t-1} - \Pi)(\widetilde{Q}_1^{\otimes m+1}(h))(X_n)\big|^2\big] \\
&= \mathbb{E}_{\pi,0}\big[\big|(P_{t-1} - \Pi)(\widetilde{Q}_1^{\otimes m+1}(h))(X_0)\big|^2\big] \quad (\pi \text{ is } P_n\text{-invariant}) \\
&= \big\|(P_{t-1} - \Pi)(\widetilde{Q}_1^{\otimes m+1}(h))\big\|_2^2 \\
&\leq \|P_{t-1} - \Pi\|_2^2 \big\|\widetilde{Q}_1^{\otimes m+1}(h)\big\|_2^2 \quad (\text{since } \widetilde{Q}_1^{\otimes m+1}(h) \in \mathbb{L}^2(\pi)) \\
&\leq \|P_{t-1} - \Pi\|_2^2 \quad (\text{since } \|\widetilde{Q}_1^{\otimes m+1}(h)\|_2 \leq 1).
\end{aligned}$$

Therefore, it follows that for every $t \geq 1$:

$$\sup\{|\mathrm{Corr}(f; h)| \, f \in \mathbb{L}^2(\mathcal{G}_1^n), h \in \mathbb{L}^2(\mathcal{G}_{n+t}^{n+t+m})\} \leq \|P_{t-1} - \Pi\|_2.$$

The right-hand side term in the inequality above does not depend on m and n, so that we obtain from (3.10)

$$\rho(t) \le \|P_{t-1} - \Pi\|_2.$$

The proof is completed by using the exponential estimate (2.2) of $\|P_{t-1} - \Pi\|_2$ under (AS2). □

3.3. *Central limit theorem for the Markov additive processes*

In a first step, we consider a discrete-time $\mathbb{X} \times \mathbb{R}^d$-valued MAP, $(X_n, Y_n)_{n \in \mathbb{N}}$, for which the driving Markov chain $(X_n)_{n \in \mathbb{N}}$ is assumed to satisfy (AS1) and (AS2). Recall that condition (AS3d) for $\alpha = 2$ is

$$\mathbb{E}_{\pi,0}\big[|Y_1|^2\big] < \infty.$$

This condition implies that $\mathbb{E}_{\pi,0}[|Y_1|] < \infty$, and we suppose that $\mathbb{E}_{\pi,0}[Y_1] = 0$ for convenience (if not, replace Y_n by $Y_n - \mathbb{E}_{\pi,0}[Y_n] = Y_n - n\mathbb{E}_{\pi,0}[Y_1]$ from Corollary 3.1).

We know from Proposition 3.1 that $(Y_{n+1} - Y_n)_{n \in \mathbb{N}}$ is stationary and is exponentially ρ-mixing when (AS1) and (AS2) hold. Then, under the expected moment condition $\mathbb{E}_{\pi,0}[|Y_1|^2] < \infty$, the CLT for $(Y_n)_{n \in \mathbb{N}^*}$ follows from [52, 73] (e.g. see [11], Theorem 11.4). To the best of our knowledge, Theorem 3.1 for general MAP is new. The notation $\mathcal{N}(0, 0)$ stands for the Dirac distribution at 0.

Theorem 3.1. *Under* (AS1) *and* (AS2) *and* (AS3d) *for $\alpha = 2$, $(Y_n/\sqrt{n})_{n \in \mathbb{N}}$ converges in distribution when $n \to +\infty$ to the d-dimensional Gaussian law $\mathcal{N}(0, \Sigma)$, where Σ is the asymptotic covariance $d \times d$-matrix*

$$\Sigma := \lim_n \frac{1}{n}\mathbb{E}_{\pi,0}\big[Y_n Y_n^*\big],$$

where the symbol $*$ denotes the transpose operator.

That $(Y_n/\sqrt{n})_{n \in \mathbb{N}}$ satisfies the CLT under the condition $\mathbb{E}_{\pi,0}[|Y_1|^2] < \infty$ was known in some cases. Such standard situations are recalled in the two next remarks (with $d = 1$ to simplify).

Remark 3.2 (Martingale method). *If there exists a measurable function $\xi : \mathbb{X} \to \mathbb{R}$ such that $Y_n - Y_{n-1} = \xi(X_n)$ and $\mathbb{E}_{\pi,0}[|Y_1|^2] = \pi(\xi^2) < \infty$, then $(Y_n/\sqrt{n})_{n \in \mathbb{N}}$ converges in distribution to the Gaussian law $\mathcal{N}(0, \sigma^2)$ where $\sigma^2 = \pi(\xi^2) + 2\sum_{\ell=1}^{+\infty}\pi(\xi P^\ell \xi) \in [0, +\infty)$. This result follows from the Gordin–Lifsic theorem [34]. Indeed, (AS2) implies that $(X_n)_{n \in \mathbb{N}}$ is ergodic and that there is a solution $\tilde{\xi} \in \mathbb{L}^2(\pi)$ to the Poisson equation: $\tilde{\xi} - P\tilde{\xi} = \xi$. Then, the difference martingale method of [34] can be used to obtain the CLT.*

Remark 3.3 (Uniform ergodicity). *Recall that the Markov chain $(X_n)_{n \in \mathbb{N}}$ is said to be uniformly ergodic if $\lim_{n \to +\infty}\|P^n - \Pi\|_\infty = 0$. This property implies (AS2) (but is stronger) and is fulfilled if and only if $(X_n)_{n \in \mathbb{N}}$ is ergodic, aperiodic and satisfies the Doeblin condition (D_0). In addition, for an aperiodic and ergodic Markov chain $(X_n)_{n \in \mathbb{N}}$, Doeblin's condition is equivalent to the uniform mixing (or ϕ-mixing) property, and then, the ϕ-mixing coefficients go to 0 at least exponentially fast (see [10,81]).*

Set $\xi_n := Y_n - Y_{n-1}$. If $(X_n)_{n \in \mathbb{N}}$ is uniformly ergodic and if $\mathbb{E}_{\pi,0}[Y_1^2] < \infty$, then the real number $\sigma^2 = \mathbb{E}_{\pi,0}[\xi_1^2] + 2\sum_{\ell=1}^{+\infty}\mathbb{E}_{\pi,0}[\xi_1 \xi_\ell]$ is well-defined in $[0, +\infty)$. If $\sigma^2 > 0$, then the sequence $(Y_n/\sqrt{n})_{n \in \mathbb{N}}$ converges in distribution to $\mathcal{N}(0, \sigma^2)$ [38]. This CLT is established as follows: the stationarity and the uniform ergodicity of $(X_n)_{n \in \mathbb{N}}$ extend to the sequence $(\xi_n)_{n \in \mathbb{N}}$, and the ϕ-mixing coefficients of $(\xi_n)_{n \in \mathbb{N}}$ also go to 0 at an exponential rate (see [38], Remark 4, Lemma 1). The proof is completed using [53], Theorem 18.5.2.

The CLT for a continuous-time MAP $(X_t, Y_t)_{t \ge 0}$ is deduced from the discrete-time statement.

Theorem 3.2. *Under* (AS1) *and* (AS2) *and* (AS3) *with $\alpha = 2$, $(Y_t/\sqrt{t})_{t > 0}$ satisfies a CLT.*

Proof. For $t \in [0, +\infty)$, we set $t = n + v$, where n is the integer part of t and $v \in [0, 1)$. We can write:

$$\frac{Y_t}{\sqrt{t}} = \frac{(Y_t - Y_n)}{\sqrt{t}} + \frac{\sqrt{n}}{\sqrt{t}} \frac{Y_n}{\sqrt{n}}. \tag{3.11}$$

Recall that $(Q_t)_{t \geq 0}$ is the transition semigroup of $(X_t, Y_t)_{t \geq 0}$. It is easily checked that the MAP $(X_t, Y_t)_{t \geq 0}$ "sampled" at discrete instants, $(X_n, Y_n)_{n \in \mathbb{N}}$, is a discrete-time MAP with transition kernel $Q := Q_1$ which satisfies (AS1), (AS2) and (AS3d). Therefore, $(Y_n/\sqrt{n})_{n \in \mathbb{N}}$ satisfies a CLT thanks to Theorem 3.1. Finally, the sequence $((Y_t - Y_n)/\sqrt{t})_{t > 0}$ converges in probability to 0 from the Tchebychev inequality and condition (AS3d):

$$\mathbb{P}_{\pi,0}\{|Y_t - Y_n| > \sqrt{t}\varepsilon\} = \mathbb{P}_{\pi,0}\{|Y_v| > \sqrt{t}\varepsilon\} \quad \text{(stationary increments)}$$

$$\leq \frac{\mathbb{E}_{\pi,0}[|Y_v|^2]}{t\varepsilon^2} \leq \frac{\sup_{v \in (0,1]} \mathbb{E}_{\pi,0}[|Y_v|^2]}{t\varepsilon^2} \xrightarrow[t \to +\infty]{} 0.$$

Therefore, $(Y_t/\sqrt{t})_{t>0}$ satisfies a CLT from (3.11). $\qquad\square$

Remark 3.4 (FCLT). *Proposition* 3.1 *allows us to deduce from* [8], *Theorem* 19.2, *that a functional central limit theorem also holds* $(d = 1)$. *That is, under the assumptions of Theorem* 3.1, *we have*:

$$\left(\frac{Y_{\lfloor nt \rfloor}}{\sqrt{n}}\right)_{t \geq 0} \xrightarrow[n \to +\infty]{\mathcal{L}} (B_t)_{t \geq 0} \tag{3.12}$$

as random elements of $D[0, \infty)$, *the Skorokhod space of cadlag functions on* $\mathbb{R}_+$, *and where* $(B_t)_{t \geq 0}$ *is a Brownian motion with zero drift and some variance parameter. Let us give some comments on the FCLT relevant to our context.*

(a) *The case of a discrete-time MAP* $(X_n, Y_n)_{n \in \mathbb{N}}$ *with* $(X_n)_{n \in \mathbb{N}}$ *satisfying the Doeblin condition is covered by* [38]. $(Y_n)_{n \in \mathbb{N}}$ *is shown to be* ϕ-*mixing and a FCLT for* ϕ-*mixing sequences is used. We extend their approach to our case of* $\mathbb{L}^2$-*spectral gap.*

(b) *Under* (AS1) *and* (AS2) *and the expected moment condition of order* 2, *Maigret* [65] *has established a FCLT for* $(Y_n := \xi(X_{n-1}, X_n))_{n \in \mathbb{N}^*}$ *in the specific case where* $(X_n)_{n \in \mathbb{N}}$ *is Harris-recurrent. It is worth noticing that condition* (AS2) *cannot be compared with the Harris-recurrence property.*

(c) *If* $(X_t)_{t \geq 0}$ *is a stationary ergodic Markov process with a strongly continuous transition semigroup* $(P_t)_{t \geq 0}$ *on* $\mathbb{L}^2(\pi)$, *the following convergence holds for any* $f \in \mathbb{L}^2(\pi)$ *such that* $\pi(f) = 0$ [7], *Theorem* 2.1, *Proposition* 2.3, (*see also* [84] *in the Harris-recurrent case*):

$$\left(\frac{1}{\sqrt{n}} \int_0^{nt} \xi(X_s) \, ds\right)_{t \geq 0} \xrightarrow[n \to +\infty]{\mathcal{L}} (B_t)_{t \geq 0},$$

where $(B_t)_{t \geq 0}$ *is a Brownian motion with zero drift and some variance parameter. Set* $Y_t := \int_0^t \xi(X_s) \, ds$. *Since* $\xi \in \mathbb{L}^2(\pi)$, *we have* $\mathbb{E}_{\pi,0}[|Y_t|^2] \leq \pi(|\xi|^2)$ *for every* $t \in (0, 1]$, *thus* (AS3) *with* $\alpha = 2$ *is true. Then, the convergence result above is easily deduced from* (3.12) *using the discrete-time stationary MAP* $(X_n, Y_n)_{n \geq 1}$ *introduced in the proof of Theorem* 3.2.

(d) *Glynn and Whitt deal with the integral functional of a regenerative process in* [30,31]. *Their results apply to a Markov process which is a specific instance of a regenerative process. Conditions for the CLT* (FCLT) *to hold are expressed in terms of a second moment on the increments* $Y_{T_1} := \int_0^{T_1} \xi(X_s) \, ds$ *of the process* $(Y_t)_{t > 0}$ *over a regeneration cycle of length* T_1 (*and an additional condition of negligeability in probability of* $\sup_{0 \leq s \leq T_1} |Y_s|$). *The fact that we only consider the Markov case makes our conditions easier to check.*

4. Refinements of the central limit theorem for MAPs

Let $(X_t, Y_t)_{t \in \mathbb{T}}$ be a MAP taking values in $\mathbb{X} \times \mathbb{R}^d$. The canonical scalar product on $\mathbb{R}^d$ is denoted by $\langle \cdot, \cdot \rangle$. The Fourier operators associated with $(X_t, Y_t)_{t \in \mathbb{T}}$ are introduced in the next subsection and are shown to satisfy a semigroup

property. In the discrete-time case, precise expansions of the characteristic function of the additive component Y_t can be deduced from [49] under (AS2). They are central to the derivation of our limit theorems in this section. Limit theorems are first considered for discrete-time MAPs. A local limit theorem, a Berry–Esseen bound and a first-order Edgeworth expansion are obtained. The continuous-time case is addressed thanks to the basic reduction to the discrete-time case used for the CLT.

4.1. *Fourier operators. A semigroup property*

For any $t \in \mathbb{T}$ and $\zeta \in \mathbb{R}^d$, we consider the linear operator $S_t(\zeta)$ acting (in a first step) on the space of bounded measurable functions $f : \mathbb{X} \to \mathbb{C}$ as follows:

$$\forall x \in \mathbb{X} \quad (S_t(\zeta)f)(x) := \mathbb{E}_{(x,0)}\big[\mathrm{e}^{\mathrm{i}\langle\zeta,Y_t\rangle}f(X_t)\big]. \tag{4.1}$$

Note that $S_t(0) = P_t$. In the discrete-time case, $S_1(\zeta)$ corresponds to the Fourier operator which was first introduced by Nagaev [68] in the special case when $Y_n = \sum_{k=1}^n \xi(X_k)$ (see [42,44] and the reference therein), and was extended to discrete-time MAPs in [4,40] to prove local limit theorems and renewals theorems (see also [26]). All these works are based on the following formula (see Proposition 4.1 below):

$$\forall \zeta \in \mathbb{R}^d, \forall n \in \mathbb{N} \quad (S_n(\zeta)f)(x) := \mathbb{E}_{(x,0)}\big[\mathrm{e}^{\mathrm{i}\langle\zeta,Y_n\rangle}f(X_n)\big] = \big(S_1(\zeta)^n f\big)(x). \tag{4.2}$$

This formula clearly reads as the semigroup property: $S_{m+n}(\zeta) = S_m(\zeta) \circ S_n(\zeta)$. In the continuous-time, it seems that the operators $S_t(\zeta)$ were first introduced in [50] for investigating AFs of continuous-time Markov processes on a compact metric state space $\mathbb{X}$. In [50], P_t was assumed to have a spectral gap on the space of all continuous $\mathbb{C}$-valued functions on $\mathbb{X}$, and $(S_t(\zeta))_{t>0}$ was thought of as a semigroup (see (SG) below) on this space.

Here, in view of (AS2), the above mentioned semigroup property has to be considered on the Lebesgue spaces $\mathbb{L}^p(\pi)$ $(1 \le p \le \infty)$.

Proposition 4.1. *For all $t \in \mathbb{T}$ and $\zeta \in \mathbb{R}^d$, $S_t(\zeta)$ defines a linear contraction on $\mathbb{L}^p(\pi)$, and we have*:

$$\forall \zeta \in \mathbb{R}^d, \forall (s,t) \in \mathbb{T}^2 \quad S_{t+s}(\zeta) = S_t(\zeta) \circ S_s(\zeta). \tag{SG}$$

In particular, relation (4.2) *holds for all $f \in \mathbb{L}^p(\pi)$.*

Proof. The first assertion is easy to prove. Next, for any $\zeta \in \mathbb{R}^d$ and $f \in \mathbb{L}^p(\pi)$, let us set: $g(x,y) := f(x)\mathrm{e}^{\mathrm{i}\langle\zeta,y\rangle}$ with $x \in \mathbb{X}$ and $y \in \mathbb{R}^d$. Then, using the Markov property and Lemma 3.1:

$$\begin{aligned}
\big(S_{t+s}(\zeta)f\big)(x) &:= \mathbb{E}_{(x,0)}\big[\mathrm{e}^{\mathrm{i}\langle\zeta,Y_{t+s}\rangle}f(X_{t+s})\big] \\
&= \mathbb{E}_{(x,0)}\big[\mathbb{E}_{(x,0)}\big[\mathrm{e}^{\mathrm{i}\langle\zeta,Y_{t+s}\rangle}f(X_{t+s})|\mathcal{F}_s^{(X,Y)}\big]\big] = \mathbb{E}_{(x,0)}\big[(Q_t g_{Y_s})(X_s,0)\big] \\
&= \mathbb{E}_{(x,0)}\big[\mathrm{e}^{\mathrm{i}\langle\zeta,Y_s\rangle}(Q_t g)(X_s,0)\big] = \mathbb{E}_{(x,0)}\big[\mathrm{e}^{\mathrm{i}\langle\zeta,Y_s\rangle}\mathbb{E}_{X_s,0}\big[f(X_t)\mathrm{e}^{\mathrm{i}\langle\zeta,Y_t\rangle}\big]\big] \\
&= \mathbb{E}_{(x,0)}\big[\mathrm{e}^{\mathrm{i}\langle\zeta,Y_s\rangle}\big(S_t(\zeta)f\big)(X_s)\big] = \big(S_s(\zeta)\big(S_t(\zeta)f\big)\big)(x)
\end{aligned}$$

the third equality results from: $g_{Y_s}(x,y) = f(x)\mathrm{e}^{\mathrm{i}\langle\zeta,(y+Y_s)\rangle} = \mathrm{e}^{\mathrm{i}\langle\zeta,Y_s\rangle}g(x,y)$. This gives the semigroup property (SG). The last assertion is obvious. □

4.2. *Expansions of the characteristic function of the additive component*

Here we assume that $(X_n, Y_n)_{n\in\mathbb{N}}$ is a discrete-time MAP taking values in $\mathbb{X} \times \mathbb{R}^d$ (possibly derived from a continuous-time MAP) such that the driving Markov chain $(X_n)_{n\in\mathbb{N}}$ is stationary and satisfies (AS2). This last property ensures that $S_1(0)$ has good spectral properties, and the iterates $S_1(\zeta)^n$ occurring in (4.2) are studied using the Nagaev–Guivarc'h spectral method which consists in applying the perturbation theory to the Fourier operators $S_1(\zeta)$ for small ζ. However using the standard perturbation theorem requires strong assumptions on Y_1. Here we shall appeal to the

weak spectral method introduced in [45] and based on the Keller–Liverani perturbation theorem [60]. This method is fully developed in the Markov framework in [49], see references therein. In the sequel, $F^{(\ell)}$ denotes the partial derivative of order ℓ of a $\mathbb{C}$-valued function F defined on an open subset of $\mathbb{R}^d$.

Conditions (AS1) and (AS2) are assumed to hold throughout the subsection.

Proposition 4.2. *Let $m_0 \in \mathbb{N}^*$. Under condition* (AS3d) *for some $\alpha > m_0$, there exists a bounded open neighborhood $\mathcal{O}$ of $\zeta = 0$ in $\mathbb{R}^d$ such that we have for all $f \in \mathbb{L}^s(\pi)$ with any $s > \frac{\alpha}{\alpha - m_0}$:*

$$\forall n \in \mathbb{N}, \forall \zeta \in \mathcal{O} \quad \mathbb{E}_{\pi,0}\big[e^{i\langle \zeta, Y_n \rangle} f(X_n)\big] = \lambda(\zeta)^n L(\zeta, f) + R_n(\zeta, f), \tag{4.3}$$

where $\lambda(\cdot), L(\cdot, f), R_n(\cdot, f)$ are $\mathbb{C}$-valued functions of class $\mathcal{C}^{m_0}$ on $\mathcal{O}$, with $\lambda(0) = 1$ and $L(0, f) = \pi(f)$. Moreover, we have the following properties for $\ell = 0, \ldots, m_0$:

$$- \quad \sup_{\zeta \in \mathcal{O}} \big|L^{(\ell)}(\zeta, f)\big| < \infty, \tag{4.4a}$$

$$- \quad \exists \kappa \in (0, 1), \quad \sup_{\zeta \in \mathcal{O}} \big|R_n^{(\ell)}(\zeta, f)\big| = \mathrm{O}(\kappa^n). \tag{4.4b}$$

If $f := 1_{\mathbb{X}}$, we have $R_n(0, 1_{\mathbb{X}}) = 0$.

When $Y_n = \sum_{k=1}^n \xi(X_k)$, the above properties are proved in [49], Section 7.3, by using (4.2) and some operator-type derivation arguments. For a general additive component Y_n, the method is the same[1] using Lemmas 4.1 and 4.2 below which slightly extend [49], Lemmas 4.2, 7.4. Mention that, by using the same lemmas, Proposition 4.2 can also be deduced from [35] which specifies the method introduced in [36,45] to prove Taylor expansions of $\lambda(\cdot), L(\cdot, f), R_n(\cdot, f)$.[2]

The operator norm in the space $\mathcal{L}(\mathbb{L}^p, \mathbb{L}^{p'})$ of the linear bounded operators from $\mathbb{L}^p(\pi)$ to $\mathbb{L}^{p'}(\pi)$ is denoted by $\|\cdot\|_{p, p'}$.

Lemma 4.1. *If $1 \le p' < p$, then the map $\zeta \mapsto S_1(\zeta)$ is continuous from $\mathbb{R}^d$ to $\mathcal{L}(\mathbb{L}^p, \mathbb{L}^{p'})$.*

Proof. We have for $\zeta \in \mathbb{R}^d$, $\zeta_0 \in \mathbb{R}^d$ and $f \in \mathbb{L}^p(\pi)$, thanks to Hölder's inequality

$$
\begin{aligned}
\big|\big(S_1(\zeta) - S_1(\zeta_0)\big) f(x)\big|^{p'} &= \big|\mathbb{E}_{(x,0)}\big[e^{i\langle \zeta, Y_1 \rangle} f(X_1)\big] - \mathbb{E}_{(x,0)}\big[e^{i\langle \zeta_0, Y_1 \rangle} f(X_1)\big]\big|^{p'} \\
&\le \mathbb{E}_{(x,0)}\big[\big|e^{i\langle \zeta - \zeta_0, Y_1 \rangle} - 1\big|^{p'} \big|f(X_1)\big|^{p'}\big] \\
&\le 2^{p'} \mathbb{E}_{(x,0)}\big[\min\big\{1, \big|\langle \zeta - \zeta_0, Y_1 \rangle\big|\big\}^{p'} \big|f(X_1)\big|^{p'}\big],
\end{aligned}
$$

the last inequality resulting from the classic inequality $|e^{ia} - 1| \le 2 \min\{1, |a|\}$. An integration with respect to π and the use of Hölder's inequality give

$$
\begin{aligned}
\pi\big(\big|\big(S_1(\zeta) - S_1(\zeta_0)\big) f\big|^{p'}\big) &\le 2^{p'} \mathbb{E}_{\pi,0}\big[\min\big\{1, \big|\langle \zeta - \zeta_0, Y_1 \rangle\big|\big\}^{(pp')/(p-p')}\big]^{(p-p')/p} \mathbb{E}_{\pi,0}\big[\big|f(X_1)\big|^p\big]^{p'/p} \\
&\le 2^{p'} \big\|\min\big\{1, \big|\langle \zeta - \zeta_0, Y_1 \rangle\big|\big\}\big\|_{(pp')/(p-p')}^{p'} \|f\|_p^{p'},
\end{aligned}
$$

since π is invariant. Thus, we deduce that $\|S_1(\zeta) - S_1(\zeta_0)\|_{p,p'} \le 2\|\min\{1, |\langle \zeta - \zeta_0, Y_1 \rangle|\}\|_{(pp')/(p-p')}$ goes to 0 when $|\zeta - \zeta_0| \to 0$ from Lebesgue's theorem. $\qquad \square$

---

[1]See the beginning of the Appendix. In particular, mention that $\lambda(\zeta)$ is the dominant eigenvalue of $S_1(\zeta)$, $L(\zeta, \cdot)$ is related to the associated eigenprojection, and κ can be chosen as $\kappa = (e^{-\varepsilon} + 1)/2$ where $\varepsilon > 0$ is defined in (2.2).

[2]As observed in [35], the passage from the Taylor expansions to the differentiability properties can be derived from [13].

Lemma 4.2. *Assume that* (AS3d) *holds for some* $\alpha > m_0$ *($m_0 \in \mathbb{N}^*$), and let* $1 \leq j \leq m_0$. *If* $p > 1$ *and* $p_j := \alpha p/(\alpha + jp) \geq 1$, *then* $\zeta \mapsto S_1(\zeta)$ *is* j-*times continuously differentiable from* $\mathbb{R}^d$ *to* $\mathcal{L}(\mathbb{L}^p, \mathbb{L}^{p_j})$, *and* $\sup_{\zeta \in \mathbb{R}^d} \|S_1^{(j)}(\zeta)\|_{p, p_j} \leq \mathbb{E}_{\pi,0}[|Y_1|^\alpha]^{j/\alpha}$.

Proof. For the sake of simplicity, we suppose that $d = 1$. Below we consider any $\zeta \in \mathbb{R}$, $\zeta_0 \in \mathbb{R}$ and $f \in \mathbb{L}^p(\pi)$. For $1 \leq j \leq m_0$, define (formally) the following linear operator:

$$\forall x \in \mathbb{X} \quad \left(S_1^{(j)}(\zeta)f\right)(x) := \mathbb{E}_{(x,0)}\left[(\mathrm{i}Y_1)^j \mathrm{e}^{\mathrm{i}\zeta Y_1} f(X_1)\right].$$

First we have:

$$\left|S_1^{(j)}(\zeta)f(x)\right|^{p_j} \leq \mathbb{E}_{(x,0)}\left[|Y_1|^{jp_j}|f(X_1)|^{p_j}\right]$$

so that, from Hölder's inequality,

$$\left\|S_1^{(j)}(\zeta)\right\|_{p, p_j} \leq \mathbb{E}_{\pi,0}\left[|Y_1|^\alpha\right]^{j/\alpha}.$$

Second, define $\Delta := S_1^{(j-1)}(\zeta) - S_1^{(j-1)}(\zeta_0) - (\zeta - \zeta_0)S_1^{(j)}(\zeta_0)$. Then, for $j \in \{1, \ldots, m_0 - 1\}$, we have thanks to the classic inequality $|\mathrm{e}^{\mathrm{i}a} - 1 - \mathrm{i}a| \leq 2|a|\min\{1, |a|\}$

$$\left|\Delta f(x)\right|^{p_j} \leq 2^{p_j}|\zeta - \zeta_0|^{p_j}\mathbb{E}_{(x,0)}\left[\min\{1, |(\zeta - \zeta_0)Y_1|\}^{p_j}|Y_1|^{jp_j}|f(X_1)|^{p_j}\right].$$

It follows from Hölder's inequality that the operator norm satisfies

$$\|\Delta\|_{p, p_j} \leq 2|\zeta - \zeta_0|\left\|\min\{1, |(\zeta - \zeta_0)Y_1|\}|Y_1|^j\right\|_{\alpha/j}.$$

This proves that $S_1^{(j-1)}(\cdot)$ is differentiable from $\mathbb{R}$ to $\mathcal{L}(\mathbb{L}^p, \mathbb{L}^{p_j})$, and that its derivatives is $S_1^{(j)}$. Finally, we obtain:

$$\left|\left(S_1^{(j)}(\zeta) - S_1^{(j)}(\zeta_0)\right)f(x)\right|^{p_j} \leq 2^{p_j}\mathbb{E}_{(x,0)}\left[\min\{1, |(\zeta - \zeta_0)Y_1|\}^{p_j}|Y_1|^{jp_j}|f(X_1)|^{p_j}\right]$$

from which we deduce that the operator norm satisfies

$$\left\|S_1^{(j)}(\zeta) - S_1^{(j)}(\zeta_0)\right\|_{p, p_j} \leq 2\left\|\min\{1, |(\zeta - \zeta_0)Y_1|\}|Y_1|^j\right\|_{\alpha/j}.$$

Thus $S_1(\cdot)$ is j-times continuously differentiable from $\mathbb{R}$ to $\mathcal{L}(\mathbb{L}^p, \mathbb{L}^{p_j})$. $\square$

Next, let us return to our probabilistic context. Let ∇ and Hess denote the gradient and the Hessian operators respectively. In the following proposition, the d-dimensional vector $\nabla\lambda(0)$ and the symmetric $d \times d$-matrix Hess$\lambda(0)$ are related to the mean vector $\mathbb{E}_{\pi,0}[Y_1]$ and the asymptotic covariance matrix associated with the sequence $(Y_n - n\mathbb{E}_{\pi,0}[Y_1])/\sqrt{n}$.

Proposition 4.3.

(i) *If* (AS3d) *holds for some* $\alpha > 1$, *then* $\nabla\lambda(0) = i\mathbb{E}_{\pi,0}[Y_1]$.
(ii) *If* (AS3d) *holds for some* $\alpha > 2$, *then the following limit exists in the set of the non-negative symmetric* $d \times d$-*matrices*:

$$\Sigma := \lim_n \frac{1}{n}\mathbb{E}_{\pi,0}\left[\left(Y_n - n\mathbb{E}_{\pi,0}[Y_1]\right)\left(Y_n - n\mathbb{E}_{\pi,0}[Y_1]\right)^*\right] = -\text{Hess}\,\lambda(0).$$

Proof. Assume that $d = 1$ for the sake of simplicity (for $d \geq 2$, the proof is similar by using partial derivatives). By differentiating at $\zeta = 0$ the equality $\mathbb{E}_{\pi,0}[\mathrm{e}^{\mathrm{i}\zeta Y_n}] = \lambda(\zeta)^n L(\zeta, 1_\mathbb{X}) + R_n(\zeta, 1_\mathbb{X})$ of Proposition 4.2, we obtain:

$$\mathrm{i}\mathbb{E}_{\pi,0}[Y_n] = n\lambda^{(1)}(0) + L^{(1)}(0, 1_\mathbb{X}) + R_n^{(1)}(0, 1_\mathbb{X}).$$

Since $\mathbb{E}_{\pi,0}[Y_n] = n\mathbb{E}_{\pi,0}[Y_1]$ (from Corollary 3.1), we deduce that $\lambda^{(1)}(0) = \mathrm{i}$ and $\lim_n \mathbb{E}_{\pi,0}[Y_n]/n = \mathrm{i}\mathbb{E}_{\pi,0}[Y_1]$ from (4.4b). To prove (ii), assume for convenience that $\mathbb{E}_{\pi,0}[Y_1] = 0$. Then $\lambda^{(1)}(0) = 0$, and differentiating twice the above equality at $\zeta = 0$ gives: $-\mathbb{E}_{\pi,0}[Y_n^2] = n\lambda^{(2)}(0) + L^{(2)}(0, 1_{\mathbb{X}}) + R_n^{(2)}(0, 1_{\mathbb{X}})$. We obtain the desired property by using again (4.4b). $\qquad\square$

4.3. *Refinements of the CLT for discrete-time MAPs*

In this subsection, $(X_n, Y_n)_{n\in\mathbb{N}}$ is a MAP taking values in $\mathbb{X} \times \mathbb{R}^d$, with a driving Markov chain $(X_n)_{n\in\mathbb{N}}$ satisfying (AS1) and (AS2). The assumptions below imply that $\mathbb{E}_{\pi,0}[|Y_1|] < \infty$, and for convenience we suppose that $\mathbb{E}_{\pi,0}[Y_1] = 0$ (if not, replace Y_n by $Y_n - n\mathbb{E}_{\pi,0}[Y_1]$).

Theorems 4.1 to 4.3 below have been established in [49] for additive components of the form $Y_n = \sum_{k=1}^n \xi(X_k)$. To the best of our knowledge, the present extensions to general MAP are new.

4.3.1. *A local limit theorem*
The classical Markov nonlattice condition is needed to state the local limit theorem (LLT):

Nonlattice condition. *There is no $a \in \mathbb{R}^d$, no closed subgroup H in $\mathbb{R}^d$, $H \neq \mathbb{R}^d$, and no bounded measurable function $\beta : \mathbb{X} \to \mathbb{R}^d$ such that*: $Y_1 + \beta(X_1) - \beta(X_0) \in a + H$ $\mathbb{P}_{\pi,0}$-*a.s.*

This condition is equivalent to the following operator-type property. For each $p \in (1, \infty)$ and for all compact subset K of $\mathbb{R}^d \setminus \{0\}$, there exists $\rho \in (0, 1)$ such that:

$$\sup_{\zeta \in K} \| S_1(\zeta)^n \|_p = \mathrm{O}(\rho^n). \tag{4.5}$$

This result is established in [49], Section 5, for additive functionals. The proof for general MAPs is similar. Since $\mathbb{E}_{\pi,0}[e^{\mathrm{i}\langle\zeta, Y_n\rangle}] = \pi(S_1(\zeta)^n 1_{\mathbb{X}})$ by (SG), it follows that

$$\sup_{\zeta \in K} \left|\mathbb{E}_{\pi,0}\left[e^{\mathrm{i}\langle\zeta, Y_n\rangle}\right]\right| = \mathrm{O}(\rho^n).$$

Theorem 4.1. *The assumptions of Theorem 3.1 are supposed to be satisfied, so that $(Y_n/\sqrt{n})_{n\in\mathbb{N}^*}$ converges in distribution to a d-dimensional Gaussian vector with covariance matrix Σ. Let us assume that Σ is a definite positive matrix. Finally, suppose that the nonlattice condition is true. Then, we have for all compactly supported continuous function $g : \mathbb{R}^d \to \mathbb{R}$:*

$$\lim_{n\to+\infty} \sqrt{\det\Sigma}\,(2\pi n)^{d/2}\mathbb{E}_{\pi,0}\left[g(Y_n)\right] = \int_{\mathbb{R}^d} g(x)\,\mathrm{d}x.$$

Proof. Thanks to (4.3) with $f := 1_{\mathbb{X}}$, Theorem 4.1 can be established as in the i.i.d. case: use Proposition 4.2 to control $L(\cdot, 1_{\mathbb{X}})$ and $R_n(\zeta, 1_{\mathbb{X}})$ and, as in [12], use the nonlattice condition and the following second-order Taylor expansion of $\lambda(\cdot)$, which follows from Theorem 3.1 and from [46], Lemma 4.2.

Lemma 4.3. *Assume that conditions (AS2) and (AS3d) with $\alpha = 2$ hold and that $\mathbb{E}_{\pi,0}[Y_1] = 0$. Then the function $\lambda(\cdot)$ in equality (4.3) satisfies the following second-order Taylor expansion*[3]:

$$\lambda(\zeta) = 1 - \langle\zeta, \Sigma\zeta\rangle/2 + \mathrm{o}(|\zeta|^2). \qquad\square$$

Remark 4.1. *We mention that a local limit theorem has been obtained in [66] for the process $(Y_n := \sum_{k=1}^n Z_k)_{n\in\mathbb{N}^*}$ associated with a stationary hidden Markov chain $(X_n, Z_n)_{n\in\mathbb{N}}$. In [66], $(X_n)_{n\in\mathbb{N}}$ is only assumed to be an ergodic stationary Markov chain so that the additional conditions for the local limit theorem to hold are more involved than those of Theorem 4.1.*

---

[3]A direct application of Proposition 4.2 gives this expansion, but under condition (AS3d) with $\alpha > 2$.

4.3.2. *Rate of convergence in the one-dimensional CLT*

Here we suppose that $d = 1$. Under the condition $\mathbb{E}_{\pi,0}[|Y_1|^{2+\varepsilon}] < \infty$, the asymptotic variance σ^2 of Proposition 4.3 is defined by $\sigma^2 := \lim_n \mathbb{E}_{\pi,0}[Y_n^2]/n$.

Theorem 4.2. *Under conditions* (AS1) *and* (AS2) *and* (AS3d) *for some* $\alpha > 3$ *and if* $\sigma^2 > 0$, *then there exists some constant* $B > 0$ *such that*

$$\forall n \geq 1 \quad \sup_{a \in \mathbb{R}} \left| \mathbb{P}_{\pi,0} \left\{ \frac{Y_n}{\sigma \sqrt{n}} \leq a \right\} - \Phi(a) \right| \leq \frac{B}{\sqrt{n}}, \tag{4.6}$$

where $\Phi(\cdot)$ *is the distribution function of the Gaussian distribution* $\mathcal{N}(0, 1)$.

Proof. Here, the functions $\lambda(\cdot)$, $L(\cdot) := L(\cdot, 1_{\mathbb{X}})$ and $R_n(\cdot) := R_n(\cdot, 1_{\mathbb{X}})$ in Proposition 4.2 are three times continuously differentiable on $\mathcal{O}$ and satisfy the following properties:

$\sup_{u \in \mathcal{O}} |L(u) - 1|/|u| < \infty$ (from (4.4a) and $L(0) = 1$),
$\sup_{u \in \mathcal{O}} |R_n(u)/u| = \mathrm{O}(\kappa^n)$ (from (4.4b) and $R_n(0) = 0$),
$\lambda(u) = 1 - \sigma^2 u^2/2 + \mathrm{O}(u^3)$ for u small enough (since $\lambda^{(1)}(0) = 0$ and $\lambda^{(2)}(0) = -\sigma^2$ from Proposition 4.3).

Then, we can borrow the proof of the Berry–Esseen theorem of the i.i.d. case (see [23]). □

Remark 4.2. *The details of the previous proof are reported in* [48], *Theorem* 2, *for the additive functional* $Y_n = \sum_{k=1}^n \xi(X_k, X_{k-1})$ *of a* V-*geometrically Markov chain. They are the same in our context. In fact, by writing out the arguments of* [48], *Theorem* 2, *we can derive the following more precise property: the constant* B *in* (4.6) *depends on the sequence* $(Y_n)_{n \in \mathbb{N}}$, *but only through* σ^2 *and* $\mathbb{E}_{\pi,0}[|Y_1|^{3+\varepsilon}]$. *Of course, this control is not as precise as in the i.i.d. case* [23], *but it is enough to obtain interesting statistical properties as in* [24,48] *or in Section* 5.

Remark 4.3. *Let us consider the specific case* $Y_n - Y_{n-1} = \xi(X_n)$ *for some real-valued measurable function* ξ. *Under conditions* (AS1) *and* (AS2), *if the real number* σ^2 *defined in Remark* 3.2 *is positive, then we have* (4.6) *under the expected moment condition* $\pi(|\xi|^3) < \infty$. *This follows from* [47], *Corollary* 3.1, *which is based on the spectral method and martingale difference arguments (see also* [49], *Section* 6). *Note that the moment condition* $\pi(|\xi|^3) < \infty$ *is optimal according to the i.i.d. case* [23].

Remark 4.4. *Let* $(X_n)_{n \in \mathbb{N}}$ *be a* ρ-*mixing Markov chain. The additive functionals of* $(X_n)_{n \in \mathbb{N}}$ *involved in the* M-*estimation of Markov models (see* (5.5)) *are of the form* $Y_n = \sum_{k=1}^n \xi(X_{k-1}, X_k)$. *Since* $(X_n, Y_n)_{n \in \mathbb{N}}$ *is a MAP, Theorem* 4.2 *applies provided that* $\xi : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ *is a measurable function such that* $\mathbb{E}_{\pi,0}[\xi(X_0, X_1)] = 0$ *and* $\mathbb{E}_{\pi,0}[|\xi(X_0, X_1)|^{3+\varepsilon}] < \infty$ *for some* $\varepsilon > 0$. *This will be supported by the statistical result of Section* 5.

Finally let us state a first-order Edgeworth expansion.

Theorem 4.3. *Assume that conditions* (AS1) *and* (AS2) *and* (AS3d) *hold for some* $\alpha > 3$, *that* σ^2 *is positive and the nonlattice condition is true. Then, there exists* $\mu_3 \in \mathbb{R}$ *such that*:

$$\mathbb{P}_{\pi,0} \left\{ \frac{Y_n}{\sigma \sqrt{n}} \leq a \right\} = \Phi(a) + \frac{\mu_3}{6\sigma^3 \sqrt{n}} (1 - a^2) \eta(a) + \mathrm{o}\left(\frac{1}{\sqrt{n}} \right), \tag{4.7}$$

where $\eta(\cdot)$ *is the density of the Gaussian distribution* $\mathcal{N}(0, 1)$.

Other limit theorems can be stated under condition (AS2) as, for instance, a multidimensional Berry–Esseen theorem in the Prohorov metric (see [49], Section 9), and the multidimensional renewal theorems (see [39]). Although Proposition 4.2 extends to the case when the order of regularity m_0 is not integer, it does not allow to deal with the convergence of Y_n (properly normalized) to stable laws, since we assume $\alpha > m_0$ (in place of the expected condition $\alpha = m_0$). For an additive functional $Y_n = \sum_{k=1}^n \xi(X_k)$, a careful examination of the proof of Lemmas 4.1 and 4.2 shows that this limitation could be overcame under a condition of the type: $\xi \in \mathbb{L}^\beta(\pi) \implies P\xi \in \mathbb{L}^{\beta'}(\pi)$ with $\beta' > \beta$.

Anyway mention that, under condition (AS2) and the previous condition on ξ, convergence to stable laws is obtained in [54], Section 2.3, by using a "martingale approximation" approach. A natural question is to ask whether the last condition on ξ is necessary.

4.4. *The non-stationary case*

Under (AS2), we discuss the extension of the previous results to the non-stationary case. Let μ be the initial distribution of $(X_n)_{n\in\mathbb{N}}$. The careful use of [49], Proposition 7.3, allows us to extend Proposition 4.2 as follows. Under condition (AS3d)[4] with $\alpha > m_0$, and under the following assumption on μ

(NS) μ *is a bounded linear form on* $\mathbb{L}^r(\pi)$ *with r such that* $1 < r < \alpha s/(\alpha + m_0 s)$,

where $s > \alpha/(\alpha - m_0)$, all the conclusions of Proposition 4.2 remain true when π is replaced by μ, namely: for some bounded open neighborhood $\mathcal{O}$ of $\zeta = 0$ in $\mathbb{R}^d$, we have for $f \in \mathbb{L}^s(\pi)$

$$\forall n \in \mathbb{N}, \forall \zeta \in \mathcal{O} \quad \mathbb{E}_{\mu,0}\big[e^{i\langle\zeta,Y_n\rangle}f(X_n)\big] = \lambda(\zeta)^n L(\zeta, f, \mu) + R_n(\zeta, f, \mu), \tag{4.8}$$

with $\mathbb{C}$-valued functions $\lambda(\cdot)$, $L(\cdot, f, \mu)$, $R_n(\cdot, f, \mu)$ satisfying the same properties as in Proposition 4.2. It is worth noticing that $\lambda(\cdot)$ is the same function as in (4.3), contrary to $L(\cdot, f, \mu)$ and $R_n(\cdot, f, \mu)$ which both depend on μ.

Condition (NS) means that μ is absolutely continuous with respect to π with density $\phi \in \mathbb{L}^{r'}(\pi)$ where $r' = r/(r-1)$ is the conjugate number of r. It is easily checked that $r' > \alpha s/((\alpha - m_0)s - \alpha) > 1$. Note that the bigger is the exponent α in condition (AS3d), the closer to 1 is the allowed value of r'.

Proposition 4.3 extends to the non-stationary case as follows.

(i) If (AS3d) and (NS) hold with $m_0 = 1$, then $\nabla\lambda(0) = i\lim_n \mathbb{E}_{\mu,0}[Y_n]/n$.
(ii) If (AS3d) and (NS) hold with $m_0 = 2$, then the conclusions of Proposition 4.3(ii) remain true with μ in place of π.

Using the decomposition (4.8) (with $f := 1_{\mathbb{X}}$), we obtain as in the stationary case the following statements.

1. Under (AS2), (AS3d) with $\alpha = 2$ and μ satisfying condition (NS) with $m_0 = 1$: the CLT, and the LLT under the additional non-lattice condition.
2. Under (AS2), (AS3d) with $\alpha > 3$ and μ satisfying condition (NS) with $m_0 = 3$: the Berry–Esseen bound, and under the non-lattice condition, the first order Edgeworth expansion (4.7) with the additional term $-b_\mu \eta(u)/(\sigma\sqrt{n})$, where b_μ is the asymptotic bias: $b_\mu = \lim_n \mathbb{E}_{\mu,0}[Y_n]$ (see [49] for details).

For instance, let us sketch the proof of the CLT. Equality (4.8) with $f := 1_{\mathbb{X}}$ gives

$$\mathbb{E}_{\mu,0}\big[e^{i\langle\zeta,Y_n/\sqrt{n}\rangle}\big] = \lambda\big(\zeta/\sqrt{n}\big)^n L\big(\zeta/\sqrt{n}, 1_{\mathbb{X}}, \mu\big) + R_n\big(\zeta/\sqrt{n}, 1_{\mathbb{X}}, \mu\big).$$

Since $m_0 = 1$ we have $\lim_n L(\zeta/\sqrt{n}, 1_{\mathbb{X}}, \mu) = 1$ and $\lim_n R_n(\zeta/\sqrt{n}, 1_{\mathbb{X}}, \mu) = 0$. Finally, the second-order Taylor expansion of Lemma 4.3 shows that $\lim_n \lambda(\zeta/\sqrt{n}) = \exp(-\langle\zeta, \Sigma\zeta\rangle/2)$.

In general, the previous statements 1. and 2. do not apply to the case when the initial distribution μ is a Dirac mass (which is not defined on $\mathbb{L}^r(\pi)$). However, when the state space $\mathbb{X}$ of the driving Markov chain is discrete, these statements are valid with any initial distribution δ_x provided that $\pi(x) > 0$ (because δ_x is then a continuous linear form on each $\mathbb{L}^p(\pi) \equiv \ell^p(\pi)$).

4.5. *The continuous-time case*

In this section, we consider the case where $\mathbb{T} = (0, +\infty)$. The process $(X_t)_{t>0}$ is assumed to satisfy conditions (AS1) and (AS2). Let us mention that the moment condition (AS3) reduces to

$$\forall v \in (0, 1] \quad \mathbb{E}_{\pi,0}\big[|Y_v|^\alpha\big] < \infty$$

---

[4]In this non-stationary case, we only require condition (AS3d) with the stationary distribution π and the mean vector remains $\mathbb{E}_{\pi,0}[Y_1]$.

when the semigroup $(Q_t)_{t\geq 0}$ is strongly continuous on $\mathbb{L}^2((\pi, 0))$ (so is $(P_t)_{t>0}$ on $\mathbb{L}^2(\pi)$).

All the theorems of the previous subsection are extended to $(Y_t)_{t>0}$. Recall that Theorems 4.1–4.3 concern the multidimensional local limit theorem, the one-dimensional Berry–Esseen theorem, the one-dimensional first-order Edgeworth expansion respectively. For the sake of simplicity, we still assume that $\mathbb{E}_{\pi,0}[Y_1] = 0$.

Theorem 4.4. *The conclusions of Theorems* 4.1–4.3 *are valid for* $(Y_t/\sqrt{t})_{t>0}$ *under the same assumptions*, *up to the following change*: *the moment condition* (AS3d) *is reinforced* (*with the same condition on* α) *in* (AS3):

$$\sup_{v\in(0,1]} \mathbb{E}_{\pi,0}\big[|Y_v|^\alpha\big] < \infty.$$

Note that the extensions to the non-stationary case presented in Section 4.4 can be adapted to the continuous-time case.

When Y_t is defined by $Y_t := \int_0^t \xi(X_s)\,\mathrm{d}s$, any moment condition of the type $\sup_{v\in[0,1]} \mathbb{E}_{\pi,0}[|Y_v|^\alpha] < \infty$ $(\alpha \geq 1)$ is fulfilled if we have $\pi(|\xi|^\alpha) < \infty$. Indeed:

$$\forall v \in [0,1] \quad \mathbb{E}_{\pi,0}\big[|Y_v|^\alpha\big] \leq \mathbb{E}_{\pi,0}\bigg[\int_0^1 \big|\xi(X_s)\big|^\alpha\,\mathrm{d}s\bigg] = \int_0^1 \mathbb{E}_{\pi,0}\big[\big|\xi(X_s)\big|^\alpha\big]\,\mathrm{d}s = \pi\big(|\xi|^\alpha\big).$$

Note that the nonlattice condition used in Theorem 4.4 is the same as in the discrete-time case (see Section 4.3.1) and plays the same role. Indeed, writing $t = n + v$ where n is the integer part of t, we know that $\mathbb{E}_{\pi,0}[e^{\mathrm{i}\langle\zeta,Y_t\rangle}] = \pi(S_1(\zeta)^n(S_v(\zeta)1_{\mathbb{X}}))$. Using (4.5) and the fact that S_v is a contraction on $\mathbb{L}^p(\pi)$ $(p \in (1, +\infty))$, it follows that

$$\sup_{\zeta\in K}\big|\mathbb{E}_{\pi,0}\big[e^{\mathrm{i}\langle\zeta,Y_t\rangle}\big]\big| = \mathrm{O}\big(\rho^n\big). \tag{4.9}$$

We prove Proposition 4.4 below which is the continuous-time version of Proposition 4.2. Then, combining Proposition 4.4 with relation (4.9), the Fourier techniques of the i.i.d. case can be used to extend Theorems 4.1–4.3 to $(Y_t/\sqrt{t})_{t>0}$

Proposition 4.4. *Let* $m_0 \in \mathbb{N}^*$. *Write time* t *as* $t = n + v$ *where* n *is the integer part of* t. *Under condition* (AS3) *for some* $\alpha > m_0$, *there exists a bounded open neighborhood* $\mathcal{O}$ *of* $\zeta = 0$ *in* $\mathbb{R}^d$ *such that we have for all* $f \in \mathbb{L}^s(\pi)$ *with any* $s > \alpha/(\alpha - m_0)$:

$$\forall t \in (0, +\infty), \forall \zeta \in \mathcal{O} \quad \mathbb{E}_{\pi,0}\big[e^{\mathrm{i}\langle\zeta,Y_t\rangle}f(X_t)\big] = \lambda(\zeta)^n L\big(\zeta, S_v(\zeta)f\big) + R_n\big(\zeta, S_v(\zeta)f\big),$$

where $\lambda(\cdot)$, $L(\cdot,\cdot)$ *and* $R_n(\cdot,\cdot)$ *are the functions of Proposition* 4.2. *Moreover, the* $\mathbb{C}$-*valued functions* $L_{v,f}(\zeta) := L(\zeta, S_v(\zeta)f)$ *and* $R_{n,v,f}(\zeta) := R_n(\zeta, S_v(\zeta)f)$ *are of class* $\mathcal{C}^{m_0}$ *on* $\mathcal{O}$, *and we have the following properties for* $\ell = 0, \ldots, m_0$:

$$\sup_{\zeta\in\mathcal{O},v\in[0,1]} \big|L_{v,f}^{(\ell)}(\zeta)\big| < \infty,$$

$$\exists\kappa \in (0,1), \quad \sup_{\zeta\in\mathcal{O},v\in[0,1]} \big|R_{n,v,f}^{(\ell)}(\zeta)\big| = \mathrm{O}\big(\kappa^n\big).$$

Note that we have $\lambda(0) = 1$, $L_{v,f}(0) = \pi(f)$, *and* $R_{n,v,1_{\mathbb{X}}}(0) = 0$.

Proof. From (4.1) and (SG), we obtain for any $\zeta \in \mathbb{R}^d$, $f \in \mathbb{L}^p$ $(1 \leq p \leq \infty)$:

$$\mathbb{E}_{\pi,0}\big[e^{\mathrm{i}\langle\zeta,Y_t\rangle}f(X_t)\big] = \pi\big(S_{n+v}(\zeta)f\big) = \pi\big(S_1(\zeta)^n\big(S_v(\zeta)f\big)\big) = \mathbb{E}_{\pi,0}\big[e^{\mathrm{i}\langle\zeta,Y_n\rangle}\big(S_v(\zeta)f\big)(X_n)\big], \tag{4.10}$$

and the desired expansion then follows from Proposition 4.2. The two following (straightforward) extensions of Lemmas 4.1–4.2 are needed to establish the others assertions. Let $t \in (0, +\infty)$.

Lemma 4.4. *If* $1 \leq p' < p$, *then the map* $\zeta \mapsto S_t(\zeta)$ *is continuous from* $\mathbb{R}^d$ *to* $\mathcal{L}(\mathbb{L}^p, \mathbb{L}^{p'})$.

Lemma 4.5. *Assume that* $\mathbb{E}_{\pi,0}[|Y_t|^\alpha] < \infty$ *for some* $\alpha > m_0$ ($m_0 \in \mathbb{N}^*$), *and let* $1 \leq j \leq m_0$. *If* $p > 1$ *and* $p_j := \alpha p/(\alpha + jp) \geq 1$, *then* $\zeta \mapsto S_t(\zeta)$ *is* j-*times continuously differentiable from* $\mathbb{R}^d$ *to* $\mathcal{L}(\mathbb{L}^p, \mathbb{L}^{p_j})$, *and* $\sup_{\zeta \in \mathbb{R}^d} \|S_t^{(j)}(\zeta)\|_{p,p_j} \leq \mathbb{E}_{\pi,0}[|Y_t|^\alpha]^{j/\alpha}$.

The regularity properties (in ζ) of the functions $L(\zeta, S_v(\zeta)f)$ and $R_n(\zeta, S_v(\zeta)f)$ are not a direct consequence of those stated in Proposition 4.2 because of the additional term $S_v(\zeta)f$. To that effect we need a careful use of the operator-type derivation procedure. This part is postponed to the Appendix on the basis of [49]. □

5. A Berry–Esseen theorem for the *M*-estimators of ρ-mixing Markov chains

The *M*-estimators are a general class of estimators in parametric statistics. This covers the special cases of maximum likelihood estimators, the least square estimators and the minimum contrast estimators. In the i.i.d. case, a modern treatment on *M*-estimation is reported in [85], Chapter 5, and a Berry–Esseen bound for *M*-estimators is obtained in [74]. In a statistical framework, such a bound has to be uniform in the parameters. Pfanzagl's method, which is applied to Markov data in [48], requires a preliminary result on the rate of convergence in the CLT for additive functionals, with a precise control of the constants with respect to the functional (cf. Remark 5.1). Earlier extensions of [74] to the Markov context are discussed in [48]. For ρ-mixing Markov chains, the closest work to ours is [76]. Our main improvement is on the moment conditions which are now close to those of the i.i.d. case. A detailed comparison is presented at the end of the section.

Let Θ be any nonempty parameter set. For a Markov chain $(X_n)_{n \in \mathbb{N}}$ with state space $\mathbb{X}$ and transition kernel P_θ which depends on $\theta \in \Theta$, we introduce the following uniform $\mathbb{L}^2$-spectral gap (i.e. the uniform ρ-mixing) property.

($\mathcal{M}$) *The Markov chain* $(X_n)_{n \in \mathbb{N}}$ *has a uniform* $\mathbb{L}^2$-*spectral gap with respect to the parameter set* Θ *if*
1. *for all* $\theta \in \Theta$, $(X_n)_{n \in \mathbb{N}}$ *has a unique* P_θ-*invariant distribution* π_θ;
2. *for all* $\theta \in \Theta$, $(X_n)_{n \in \mathbb{N}}$ *is stationary (i.e.* $X_0 \sim \pi_\theta$);
3. *its transition kernel satisfies condition* (AS2) *in a uniform way with respect to* θ, *namely there exist* $C > 0$ *and* $\kappa \in (0,1)$ *such that*

$$\forall \theta \in \Theta, \forall n \geq 1 \quad \|P_\theta^n - \Pi_\theta\|_{2,\theta} \leq C\kappa^n,$$

where $\Pi_\theta(f) := \pi_\theta(f)1_\mathbb{X}$ *for* $f \in \mathbb{L}^2(\pi_\theta)$ *and* $\|\cdot\|_{2,\theta}$ *denotes the operator norm on* $\mathbb{L}^2(\pi_\theta)$.

In order to derive a Berry–Esseen bound for the *M*-estimators of $(X_n)_{n \in \mathbb{N}}$ satisfying ($\mathcal{M}$), we need a uniform Berry–Esseen bound for some specific additive functionals of the Markov chain $(X_n)_{n \in \mathbb{N}}$. In the next subsection, we propose such a uniform Berry–Esseen bound for the second component of a general parametric MAP. This result will be applied to the MAPs associated with these specific additive functionals (see Remark 5.1).

5.1. *A uniform Berry–Essen bound for the second component of a parametric MAP*

Here we propose a refinement of Proposition 4.3 and Theorem 4.2. Let us introduce the following condition.

(A) *For every* $\theta \in \Theta$, $(X_n, Y_n)_{n \in \mathbb{N}}$ *is a* $\mathbb{X} \times \mathbb{R}$-*valued MAP*, Y_1 *is* $\mathbb{P}_\theta$-*integrable and centered (i.e.* $\mathbb{E}_\theta[Y_1] = 0$).

Below, the driving Markov chain $(X_n)_{n \in \mathbb{N}}$ is assumed to satisfy condition ($\mathcal{M}$). Thus, the notation $\mathbb{P}_\theta$ stands for the underlying probability measure, which depends on θ through the transition kernel Q_θ of $(Y_n, X_n)_{n \in \mathbb{N}}$ and the initial (stationary) distribution $(\pi_\theta, 0)$. $\mathbb{E}_\theta[\cdot]$ denotes the associated expectation.

Theorem 5.1. *Assume that condition* (A) *is true for the MAP* $(X_n, Y_n)_{n \in \mathbb{N}}$ *and that the driving Markov chain* $(X_n)_{n \in \mathbb{N}}$ *satisfies condition* ($\mathcal{M}$). *If* $M_2 := \sup_{\theta \in \Theta} \mathbb{E}_\theta[|Y_1|^{2+\varepsilon}] < \infty$ *with some* $\varepsilon > 0$, *then* $\sigma^2(\theta) := \lim_n \mathbb{E}_\theta[Y_n^2]/n$ *is well-defined and is finite for each* $\theta \in \Theta$, *the function* $\sigma^2(\cdot)$ *is bounded on* Θ, *and there exists a positive constant* C_Y *such that*

$$\forall n \geq 1 \quad \sup_{\theta \in \Theta} \left| \sigma^2(\theta) - \frac{\mathbb{E}_\theta[Y_n^2]}{n} \right| \leq \frac{C_Y}{n}. \tag{5.1}$$

The constant C_Y depends on the sequence $(Y_n)_{n\in\mathbb{N}}$, but only through the constant M_2.

If the two following additional conditions hold true

$$\exists \varepsilon > 0, \quad M_3 := \sup_{\theta\in\Theta} \mathbb{E}_\theta\big[|Y_1|^{3+\varepsilon}\big] < \infty, \tag{5.2}$$

$$\sigma_0 := \inf_{\theta\in\Theta} \sigma(\theta) > 0, \tag{5.3}$$

then there exists a positive constant B_Y such that

$$\forall \theta \in \Theta, \forall n \geq 1 \quad \sup_{a\in\mathbb{R}} \left| \mathbb{P}_\theta \left\{ \frac{Y_n}{\sigma(\theta)\sqrt{n}} \leq a \right\} - \Phi(a) \right| \leq \frac{B_Y}{\sqrt{n}}. \tag{5.4}$$

The constant B_Y depends on the sequence $(Y_n)_{n\in\mathbb{N}}$ but only through σ_0 and the constant M_3.

Recall that the proofs of Proposition 4.3 and Theorem 4.2 are based on Proposition 4.2. Here, for any fixed $\theta \in \Theta$, Proposition 4.2 applies and gives an expansion of $\mathbb{E}_\theta[\mathrm{e}^{\mathrm{i}\zeta Y_n} f(X_n)]$, but for a neighbourhood $\mathcal{O}_\theta$ of $\zeta = 0$, some $\mathbb{C}$-valued functions $\lambda_\theta(\cdot)$, $L_\theta(\cdot, f)$, $R_{\theta,n}(\cdot, f)$, and some $\kappa_\theta \in (0, 1)$, which all may depend on θ. Consequently, in order to prove Theorem 5.1, we must establish that, under conditions $(\mathcal{M})$, (A) and the moment condition $\sup_{\theta\in\Theta} \mathbb{E}_\theta[|Y_1|^{m_0+\varepsilon}] < \infty$, all the conclusions of Proposition 4.2 are fulfilled in a uniform way with respect to $\theta \in \Theta$. This job has been done in [48], Section III.2, in the context of V-geometrically ergodic Markov chains. The arguments in the present setting are the same up to the following changes: replace the uniform V-geometrical ergodicity assumption of [48] by assumption $(\mathcal{M})$, and replace the domination condition (D_{m_0}) of [48] by the moment condition $\sup_{\theta\in\Theta} \mathbb{E}_\theta[|Y_1|^{m_0+\varepsilon}] < \infty$. The previous assumptions allow us to extend Lemmas 4.1 and 4.2, and so Proposition 4.2, in a uniform way in $\theta \in \Theta$.

Remark 5.1. *In the next subsection, Theorem 5.1 will be applied as follows. Given a Markov chain $(X_n)_{n\in\mathbb{N}}$ satisfying condition $(\mathcal{M})$ with respect to Θ, we consider the MAP $(X_n, Y_n(p))_{n\in\mathbb{N}}$ where $Y_n(p)$ depends on some parameter $p \in \mathcal{P}$ and is of the form*

$$Y_n(p) := \sum_{k=1}^{n} g(p, X_{k-1}, X_k).$$

The property of the constant C_Y in Theorem 5.1 ensures that inequality (5.1) is uniform in p and θ when $M_2 := \sup_{p\in\mathcal{P}, \theta\in\Theta} \mathbb{E}_\theta[|Y_1(p)|^{2+\varepsilon}] < \infty$ (of course, the asymptotic variance in (5.1) is replaced by some $\sigma^2(\theta, p)$). In the same way, the Berry–Esseen bound (5.4) is uniform in p and θ when $M_3 := \sup_{p\in\mathcal{P}, \theta\in\Theta} \mathbb{E}_\theta[|Y_1(p)|^{3+\varepsilon}] < \infty$ and $\inf_{p\in\mathcal{P}, \theta\in\Theta} \sigma(\theta, p) > 0$.

Note that these comments extend to a general MAP $(X_n, Y_n)_{n\in\mathbb{N}}$ which may depend on some parameter γ via its probability distribution and its functional form, provided that the bounds M_2, M_3, σ_0 in Theorem 5.1 are uniform in γ.

Remark 5.2. *The conclusions of Theorem 5.1 are also valid when $X_0 \sim \mu_\theta$ with μ_θ of the form $\mu_\theta = \phi_\theta \, \mathrm{d}\pi_\theta$, provided that $\sup_{\theta\in\Theta} \|\phi_\theta\|_{r'} < \infty$, with r' defined as in Section 4.4 (case $m_0 = 3$).*

5.2. A Berry–Esseen bound for the M-estimators of ρ-mixing Markov chains

Throughout this subsection, Θ is some general parameter space and $(X_n)_{n\geq0}$ is a Markov chain with state space $\mathbb{X}$ satisfying the uniform $\mathbb{L}^2$-spectral gap condition $(\mathcal{M})$. The underlying probability measure and the associated expectation are denoted by $\mathbb{P}_\theta$ and $\mathbb{E}_\theta[\cdot]$. Recall that $(X_n)_{n\in\mathbb{N}}$ is assumed to be stationary under $(\mathcal{M})$. Let us introduce the additive functional of $(X_n)_{n\geq0}$

$$M_n(\alpha) = \frac{1}{n} \sum_{k=1}^{n} F(\alpha, X_{k-1}, X_k), \tag{5.5}$$

where $\alpha \equiv \alpha(\theta) \in \mathcal{A}$ is the parameter of interest, $F(\cdot, \cdot, \cdot)$ is a real-valued measurable function on $\mathcal{A} \times \mathbb{X}^2$ and $\mathcal{A}$ is an open interval on the real line. Function F is assumed to satisfy the following moment condition

$$\sup\{\mathbb{E}_\theta[|F(\alpha, X_0, X_1)|], \theta \in \Theta, \alpha \in \mathcal{A}\} < \infty. \tag{5.6}$$

Set $M_\theta(\alpha) := \mathbb{E}_\theta[F(\alpha, X_0, X_1)]$. We assume 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 we have $M_\theta(\alpha) > M_\theta(\alpha_0), \forall \alpha \neq \alpha_0$. To estimate α_0, we consider the M-estimator $\widehat{\alpha}_n$ defined by

$$M_n(\widehat{\alpha}_n) = \min_{\alpha \in \mathcal{A}} M_n(\alpha).$$

Also assume that, for all $(x, y) \in \mathbb{X}^2$, the map $\alpha \mapsto F(\alpha, x, y)$ is twice continuously differentiable on $\mathcal{A}$. Let $F^{(1)}$ and $F^{(2)}$ be the first and second order partial derivatives of F with respect to α. Then

$$M_n^{(1)}(\alpha) = \frac{1}{n}\sum_{k=1}^n F^{(1)}(\alpha, X_{k-1}, X_k), \qquad M_n^{(2)}(\alpha) = \frac{1}{n}\sum_{k=1}^n F^{(2)}(\alpha, X_{k-1}, X_k). \tag{5.7}$$

We shall appeal to the following assumptions.

(V0) *There exists some real constant $\varepsilon > 0$ such that*

$$\sup_{\theta \in \Theta, \alpha \in \mathcal{A}} \mathbb{E}_\theta[|F^{(1)}(\alpha, X_0, X_1)|^{3+\varepsilon} + |F^{(2)}(\alpha, X_0, X_1)|^{3+\varepsilon}] < \infty.$$

(V1) $\forall \theta \in \Theta$, $\mathbb{E}_\theta[F^{(1)}(\alpha_0, X_0, X_1)] = 0$ *and $\alpha_0 \equiv \alpha_0(\theta)$ is the unique parameter value for which this property is true*;
(V2) $m(\theta) := \mathbb{E}_\theta[F^{(2)}(\alpha_0, X_0, X_1)]$ *satisfies $\inf_{\theta \in \Theta} m(\theta) > 0$*;
(V3) $\forall n \geq 1$, $M_n^{(1)}(\widehat{\alpha}_n) = 0$.

Notice that (V0) gives $\sup_{\theta \in \Theta} m(\theta) < \infty$. Set $Y_n^{(1)}(\alpha) := n M_n^{(1)}(\alpha)$ and $Y_n^{(2)}(\alpha) := n M_n^{(2)}(\alpha)$. Then, thanks to Theorem 5.1 applied to MAPs $(X_n, Y_n^{(1)}(\alpha))_{n \in \mathbb{N}}$ and $(X_n, Y_n^{(2)}(\alpha))_{n \in \mathbb{N}}$, the conditions (V0)–(V2) enable us to define the asymptotic variances:

$$\sigma_1^2(\theta) := \lim_n \frac{1}{n}\mathbb{E}_\theta[Y_n^{(1)}(\alpha_0)^2], \qquad \sigma_2^2(\theta) := \lim_n \frac{1}{n}\mathbb{E}_\theta[(Y_n^{(2)}(\alpha_0) - nm(\theta))^2],$$

and we know that $\sup_{\theta \in \Theta} \sigma_j(\theta) < \infty$ for $j = 1, 2$. The following additional conditions are also required:

(V4) $\inf_{\theta \in \Theta} \sigma_j(\theta) > 0$ *for $j = 1, 2$.*
(V5) *There exist $\eta > 2$ and a measurable function $W > 0$ such that $\sup_{\theta \in \Theta} \mathbb{E}_\theta[W^\eta] < \infty$ and*

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

(V6) *There exists a sequence $\gamma_n \to 0$ such that*

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

with $d := \inf_{\theta \in \Theta} m(\theta)/(4(\mathbb{E}_\theta[W(X_0)] + 1))$.

Theorem 5.2. *Assume that the Markov chain $(X_n)_{n \in \mathbb{N}}$ satisfies condition $(\mathcal{M})$, that F satisfies condition (5.6), that the M-estimator $\widehat{\alpha}_n$ is defined as above, and finally that conditions (V0)–(V6) are fulfilled. Set $\tau(\theta) := \sigma_1(\theta)/m(\theta)$. Then there exists a positive constant C such that*

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

Thanks to Theorem 5.1, the proof of Theorem 5.2 borrows the adaptation of Pfanzgal's method given in [48]. One of the main difficulties in this method is to obtain a Berry–Esseen bound for the additive functionals $Y_n(p) := \sum_{k=1}^n g(p, X_{k-1}, X_k)$ with $p := (v, q, \alpha_0)$ and

$$g(p, X_{k-1}, X_k) := F^{(1)}(\alpha_0, X_{k-1}, X_k) + \frac{v}{\sqrt{q}} \frac{\sigma_1(\theta)}{m(\theta)} \big(F^{(2)}(\alpha_0, X_{k-1}, X_k) - m(\theta)\big)$$

for $|v| \le 2\sqrt{\ln q}$. Observe that we have from (V0)–(V2):

$$\sup_{\{(v,q): |v| \le 2\sqrt{\ln q}\}, \theta \in \Theta} \mathbb{E}_\theta\big[\big|g(p, X_0, X_1)\big|^{3+\varepsilon}\big] < \infty.$$

Then, Remark 5.1 gives the desired Berry–Esseen bound for $(Y_n(p))_{n\in\mathbb{N}}$ in a uniform way over the parameter (θ, p).

When the X_n's are i.i.d., Theorem 5.2 corresponds to Pfanzgal's theorem [74] up to the following changes: in [74], π_θ is the common law of the X_n's; the additive functional is $M_n(\alpha) = (1/n)\sum_{k=1}^n F(\alpha, X_k)$; we simply have $\sigma_1^2(\theta) = \mathbb{E}_\theta[F^{(1)}(\theta, X_0)^2]$ and $\sigma_2^2(\theta) = \mathbb{E}_\theta[(F^{(2)}(\theta, X_0) - m(\theta))^2]$, and finally assumption (V0) is replaced by the weaker (and optimal) moment condition: $\sup_{\theta\in\Theta} \mathbb{E}_\theta[|F^{(1)}(\theta, X_0)|^3 + |F^{(2)}(\theta, X_0)|^3] < \infty$.

Earlier extensions of [74] to the Markov context are discussed in [48]. Let us compare our result with that of [76], in which the family of transition probabilities P_θ is assumed to satisfy a uniform Doeblin condition with respect to $\theta \in \Theta$. This condition corresponds to a uniform $\mathbb{L}^\infty$-spectral gap condition with respect to Θ which is stronger than our condition $(\mathcal{M})$ (see Section 2.1). Let us mention that the moment condition on $F^{(1)}$ and $F^{(2)}$ in [76] is the following ($\alpha(\theta) = \theta$ in [76]):

$$\sup_{x\in\mathbb{X}, \theta\in\Theta} \mathbb{E}_\theta\big[\big|F^{(1)}(\theta, X_0, X_1)\big|^3 + \big|F^{(2)}(\theta, X_0, X_1)\big|^3 | X_0 = x\big] < \infty.$$

Because of the supremum over $x \in \mathbb{X}$, this condition is in general much stronger than our moment condition (V0) (despite the order $3 + \varepsilon$ in (V0) instead of 3). To see that, neglect the role of θ and consider a functional f on $\mathbb{X}$. Then the difference between the condition used in [76] and (V0) is comparable to that between $\sup_{x\in\mathbb{X}} \mathbb{E}[|f(X_1)|^3 | X_0 = x]$ and $\mathbb{E}_\pi[|f(X_1)|^{3+\varepsilon}]$ (or, equivalently, between the supremum norm $\|P(|f|^3)\|_\infty$ and the norm $\|f\|_{3+\varepsilon}$ of f in $\mathbb{L}^{3+\varepsilon}(\pi)$). Consequently, Theorem 5.2 applies to the models considered in [76] but requires weaker moment conditions.

Remark 5.3. *The conclusion of Theorem* 5.2 *holds true when $X_0 \sim \mu_\theta$ and μ_θ satisfies the condition given in Remark* 5.2. *In this case, if F is such that*

$$\sup_{\theta\in\Theta, \alpha\in\mathcal{A}} \mathbb{E}_\theta\big[\big|F(\alpha, X_0, X_1)\big|^{1+\varepsilon}\big] < \infty$$

for some $\varepsilon > 0$, then $M_\theta(\alpha) = \mathbb{E}_\theta[F(\alpha, X_0, X_1)]$ can also be defined by (see Section 4.4):

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

6. Conclusion

In this paper, we propose limit theorems for the second component $(Y_t)_{t\in\mathbb{T}}$ of a discrete or continuous-time Markov Additive Process (MAP) $(X_t, Y_t)_{t\in\mathbb{T}}$ when $(X_t)_{t\in\mathbb{T}}$ has a $\mathbb{L}^2(\pi)$-spectral gap. The derivation of the CLT is based on a ρ-mixing condition strongly connected to the $\mathbb{L}^2(\pi)$-spectral gap property. The results related to the convergence rate in the CLT are developed from the weak spectral method of [49]. Note that here the discrete and continuous-time cases are covered in a unified way. In this context, the semigroup property (SG) for the family of operators $(S_t(\zeta))_{t\in\mathbb{T}}$ defined by $(S_t(\zeta)f)(x) := \mathbb{E}_{(x,0)}[e^{i\langle\zeta, Y_t\rangle} f(X_t)]$ ($\zeta \in \mathbb{R}^d$, $x \in X$) has a central role. We mention that this semigroup property is essentially true only for MAPs. The impact of the results is expected to be high for models involving a $\mathbb{L}^2(\pi)$-spectral gap, since the limit theorems are valid for general (discrete and continuous time) MAPs, and under optimal (or almost optimal) moment conditions. This is illustrated in Section 5 where a Berry–Esseen bound for the M-estimator associated with ρ-mixing Markov chains, is derived under the (almost) expected moment condition.

Appendix: Additional material for the proof of Proposition 4.4

Here, we study the regularity properties of the functions $\zeta \mapsto L(\zeta, S_v(\zeta)f)$ and $\zeta \mapsto R_n(\zeta, S_v(\zeta)f)$ involved in the decomposition of Proposition 4.4.

(1) Let us recall that we have (see (4.2))

$$\forall \zeta \in \mathbb{R}^d, \forall n \in \mathbb{N} \quad \mathbb{E}_{\pi,0}\big[e^{i\langle \zeta, Y_n \rangle} f(X_n)\big] = \pi\big(S_1(\zeta)^n f\big)$$

and, for t in some open neighbourhood $\mathcal{O}$ of $\zeta = 0$, (see [49], Section 7.2)

$$S_1(\zeta)^n = \lambda(\zeta)^n \Pi(\zeta) + N(\zeta)^n,$$

where $\lambda(\zeta)$ is the dominant eigenvalue of $S_1(\zeta)$, $\Pi(\zeta)$ is the associated rank-one eigenprojection and $N(\zeta)$ is a bounded linear operator on each $\mathbb{L}^p(\pi)$ $1 < p < \infty$. Both equalities imply that

$$\mathbb{E}_{\pi,0}\big[e^{i\langle \zeta, Y_n \rangle} f(X_n)\big] = \lambda(\zeta)^n \pi\big(\Pi(\zeta)f\big) + \pi\big(N(\zeta)^n f\big). \tag{A.1}$$

Furthermore the eigenprojection $\Pi(\zeta)$ and the operators $N(\zeta)^n$ are defined as in the standard perturbation theory by

$$\Pi(\zeta) = \frac{1}{2i\pi} \oint_{\Gamma_1} \big(z - S_1(\zeta)\big)^{-1} \mathrm{d}z, \qquad N(\zeta)^n = \frac{1}{2i\pi} \oint_{\Gamma_0} z^n \big(z - S_1(\zeta)\big)^{-1} \mathrm{d}z,$$

where these line integrals are considered respectively on some oriented circle Γ_1 centered at $z = 1$, and on some oriented circle Γ_0 centered at $z = 0$, with radius $\kappa < 1$ where κ is (for instance) $(1 + \exp(-\varepsilon))/2$ with ε defined in (2.2).

(2) Let us return to the continuous-time case. We obtain from (4.10) and (A.1)

$$\forall t \in \mathcal{O} \quad \mathbb{E}_{\pi,0}\big[e^{i\langle \zeta, Y_t \rangle} f(X_t)\big] = \lambda(\zeta)^n \pi\big(\Pi(\zeta)\big(S_v(\zeta)f\big)\big) + \pi\big(N(\zeta)^n\big(S_v(\zeta)f\big)\big).$$

Thus, we can write with the notations introduced in Proposition 4.4

$$\forall t \in \mathcal{O} \quad L\big(\zeta, S_v(\zeta)f\big) := \pi\big(\Pi(\zeta)\big(S_v(\zeta)f\big)\big), \qquad R_n\big(\zeta, S_v(\zeta)f\big) := \pi\big(N(\zeta)^n\big(S_v(\zeta)f\big)\big).$$

Therefore, we only need to study the regularity of the map $\zeta \mapsto (z - S_1(\zeta))^{-1} \circ S_v(\zeta)$ on $\mathcal{O}$ for controlling that of the map $\zeta \mapsto \mathbb{E}_{\pi,0}[e^{i\langle \zeta, Y_t \rangle} f(X_t)]$ on $\mathcal{O}$ (and, as a result, proving Proposition 4.4).

(3) Recall that $\| \cdot \|_{p,p'}$ denotes the operator norm in the space $\mathcal{L}(\mathbb{L}^p, \mathbb{L}^{p'})$ of the linear bounded operators from $\mathbb{L}^p(\pi)$ to $\mathbb{L}^{p'}(\pi)$. The notation $W(\cdot) \in \mathcal{C}^j(\theta, \theta')$ means that there exists a bounded open neighborhood $\mathcal{V}$ of $\zeta = 0$ in $\mathbb{R}^d$ such that:

$$\forall \zeta \in \mathcal{V} \quad W(\zeta) \in \mathcal{L}\big(\mathbb{L}^\theta, \mathbb{L}^{\theta'}\big) \quad \text{and} \quad W : \mathcal{V} \mapsto \mathcal{L}\big(\mathbb{L}^\theta, \mathbb{L}^{\theta'}\big) \quad \text{has a continuous } j\text{-order differential on } \mathcal{V}.$$

Let us introduce the maps $U : \zeta \mapsto (z - S_1(\zeta))^{-1}$ and $V : \zeta \mapsto S_v(\zeta)$. We are going to apply the next obvious regularity property. Let $1 \le \theta_{2m_0+2} < \theta_{2m_0+1} < \cdots < \theta_1 < \theta_0 < \infty$ (note that $\mathbb{L}^{\theta_0} \subset \mathbb{L}^{\theta_1} \subset \cdots \subset \mathbb{L}^{\theta_{2m_0+1}} \subset \mathbb{L}^{\theta_{2m_0+2}}$), and assume that we have:

$$U \in \mathcal{C}^0(\theta_{2m_0+1}, \theta_{2m_0+2}) \cap \mathcal{C}^1(\theta_{2m_0-1}, \theta_{2m_0+2}) \cap \cdots \cap \mathcal{C}^{m_0-1}(\theta_3, \theta_{2m_0+2}) \cap \mathcal{C}^{m_0}(\theta_1, \theta_{2m_0+2}),$$

$$V \in \mathcal{C}^0(\theta_0, \theta_1) \cap \mathcal{C}^1(\theta_1, \theta_3) \cap \cdots \cap \mathcal{C}^{m_0-1}(\theta_1, \theta_{2m_0-1}) \cap \mathcal{C}^{m_0}(\theta_1, \theta_{2m_0+1}).$$

Then $UV \in \mathcal{C}^{m_0}(\theta_0, \theta_{2m_0+2})$.

(4) Let us introduce the following (non-increasing) maps from $[1, +\infty)$ to $\mathbb{R}$:

$$T_0(\theta) := \frac{\alpha\theta}{\alpha + \varepsilon_0\theta} \quad \text{and} \quad T_1(\theta) := \frac{\alpha\theta}{\alpha + \theta},$$

where ε_0 will be defined in (A.4). Let $\theta > 1$. Lemma 4.4 and the continuous inclusions between the Lebesgue spaces show that

$$T_0(\theta) \geq 1 \quad \Longrightarrow \quad \forall \theta' \in [1, T_0(\theta)] \quad S_v(\cdot) \in \mathcal{C}^0(\theta, \theta'). \tag{A.2}$$

On the same way, Lemma 4.5 gives for $j = 1, \ldots, m_0$:

$$T_1^j(\theta) \geq 1 \quad \Longrightarrow \quad \forall \theta' \in [1, T_1^j(\theta)] \quad S_v(\cdot) \in \mathcal{C}^j(\theta, \theta'), \tag{A.3}$$

and the derivatives in the last property are uniformly bounded in $v \in [0, 1]$ on any bounded open neighborhood of $\zeta = 0$.

Now set $\theta_0 := s$, $\theta_1 := T_0(s)$, and observe that the assumption on s (i.e. $s > \alpha/(\alpha - m_0)$) is equivalent to $T_1^{m_0}(\theta_0) = \alpha\theta_0/(\alpha + m_0\theta_0) > 1$, so that there exists $\varepsilon_0 > 0$ such that

$$(T_0 T_1)^{m_0} T_0(\theta_1) = (T_0 T_1)^{m_0} T_0\big(T_0(\theta_0)\big) = \frac{\alpha\theta_0}{\alpha + (m_0 + (m_0 + 2)\varepsilon_0)\theta_0} = 1. \tag{A.4}$$

Define

$$\theta_2 := T_0(\theta_1), \qquad \theta_3 := T_1 T_0(\theta_1), \qquad \theta_4 := T_0 T_1 T_0(\theta_1), \qquad \ldots, \qquad \theta_{2m_0+2} := (T_0 T_1)^{m_0} T_0(\theta_1),$$

namely: $\theta_{2j} := (T_0 T_1)^{j-1} T_0(\theta_1)$ for $j = 1, \ldots, m_0 + 1$, and $\theta_{2j+1} := T_1(T_0 T_1)^{j-1} T_0(\theta_1)$ for $j = 1, \ldots, m_0$. Note that $\theta_{2m_0+2} = 1$. From (A.2) and (A.3), $V(\cdot) := S_v(\cdot)$ satisfies the regularity properties stated in part (3), and the corresponding derivatives (on any bounded open neighborhood of $\zeta = 0$) are uniformly bounded in $v \in [0, 1]$.

Next, setting $I := \{\theta_1, \theta_2, \ldots, \theta_{2m_0+2}\}$, it follows from (A.2) and (A.3) (with $v = 1$) that condition $\mathcal{C}(m_0)$ of [49], Section 7.1, holds, so that the conclusions reported in [49], p. 48, are true:

(H_0) *if $\theta \in I$ and $T_0(\theta) \in I$, then $\zeta \mapsto (z - S_1(\zeta))^{-1} \in \mathcal{C}^0(\theta, T_0(\theta))$ uniformly in $z \in \Gamma_0 \cup \Gamma_1$.*

and for $\ell = 1, \ldots, m_0$:

(H_ℓ) *If $\theta \in \bigcap_{k=0}^{\ell} [T_0^{-1}(T_0 T_1)^{-k}(I) \cap (T_1 T_0)^{-k}(I)]$, then $\zeta \mapsto (z - S_1(\zeta))^{-1} \in \mathcal{C}^\ell(\theta, (T_0 T_1)^\ell T_0(\theta))$ uniformly in $z \in \Gamma_0 \cup \Gamma_1$.*

Therefore $U(\cdot) := (z - S_1(\cdot))^{-1}$ satisfies the regularity properties stated in part (3).

(5) Finally, we deduce from the property of part (3) that there exists a neighbourhood $\mathcal{V}$ of $\zeta = 0$ in $\mathbb{R}^d$ such that the map $\zeta \mapsto (z - S_1(\zeta))^{-1} \circ S_v(\zeta)$ is m_0-times continuously differentiable from $\mathcal{V}$ to $\mathcal{L}(\mathbb{L}^s(\pi), \mathbb{L}^1(\pi))$ uniformly in $z \in \Gamma_0 \cup \Gamma_1$ and furthermore we have for $\ell = 0, \ldots, m_0$:

$$\sup\big\{\big\|\big((z - S_1(\zeta))^{-1} \circ S_v(\zeta)\big)^{(\ell)}\big\|_{s,1}; z \in \Gamma_0 \cup \Gamma_1, \zeta \in \mathcal{O}, v \in (0, 1]\big\} < \infty.$$

References

[1] S. Asmussen. *Ruin Probabilities*. World Sci. Publishing Co. Inc., River Edge, NJ, 2000. MR1794582

[2] S. Asmussen. *Applied Probability and Queues*, Vol. 51, 2nd edition. Springer-Verlag, New York, 2003. MR1978607

[3] S. Asmussen, F. Avram and M. R. Pistorius. Russian and American put options under exponential phase-type Lévy models. *Stochastic Process. Appl.* **109** (2004) 79–111. MR2024845

[4] M. Babillot. Théorie du renouvellement pour des chaînes semi-markoviennes transientes. *Ann. Inst. H. Poincaré Probab. Statist.* **24** (1988) 507–569. MR0978023

[5] A. Benveniste and J. Jacod. Systèmes de Lévy des processus de Markov. *Invent. Math.* **21** (1973) 183–198. MR0343375

[6] J. Bergh and J. Löfström. *Interpolation Spaces. An Introduction*. Springer-Verlag, Berlin, 1976.

[7] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Probab. Theory Related Fields* **60** (1982) 185–201. MR0663900

[8] P. Billingsley. *Probability and Measure*, 3rd edition. John Wiley & Sons Inc., New York, 1995. MR1324786

[9] M. Bladt, B. Meini, M. F. Neuts and B. Sericola. Distributions of reward functions on continuous-time Markov chains. In *Matrix-Analytic Methods* 39–62. World Sci. Publishing, Adelaide, 2002. MR1923878

[10] R. C. Bradley. Basic properties of strong mixing conditions. a survey and some open questions. *Probab. Surv.* **2** (2005) 107–144. MR2178042

[11] R. C. Bradley. Introduction to strong mixing conditions (Volume I). Technical report, Indiana Univ., 2005.

[12] L. Breiman. *Probability*. SIAM, Philadelphia, PA, 1993. MR1163370

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

[14] O. Cappé, E. Moulines and T. Rydén. *Inference in Hidden Markov Models*. Springer, New York, 2005. MR2159833

[15] E. Çinlar. Markov additive processes Part II. *Probab. Theory Related Fields* **24** (1972) 95–121.

[16] E. Çinlar. *Introduction to Stochastic Processes*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1975. MR0380912

[17] E. Çinlar. Shock and wear models and Markov additive processes. In *The Theory and Applications of Reliability, with Emphasis on Bayesian and Nonparametric Methods*, Vol. I 193–214. Academic Press, New York, 1977.

[18] M.-F. Chen. *From Markov Chains to Non-equilibrium Particle Systems*, 2nd edition. World Sci. Publishing Co. Inc., River Edge, NJ, 2004. MR2091955

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

[20] J. L. Doob. *Stochastic Processes*. John Wiley & Sons, New York, 1953. MR0058896

[21] Ī. Ī. Ezhov and A. V. Skorohod. Markov processes which are homogeneous in the second component. I. *Theory Probab. Appl.* **14** (1969) 1–13.

[22] Ī. Ī. Ezhov and A. V. Skorohod. Markov processes which are homogeneous in the second component. II. *Theory Probab. Appl.* **14** (1969) 652–667.

[23] W. Feller. *An Introduction to Probability Theory and Its Applications*, Vol. II. John Wiley and Sons, New York, 1971. MR0270403

[24] D. Ferré. Développement d'Edgeworth d'ordre 1 pour des M-estimateurs dans le cas de chaînes V-géométriquement ergodiques. *CRAS* **348** (2010) 331–334. MR2600134

[25] G. Fort, E. Moulines, G. O. Roberts and J. S. Rosenthal. On the geometric ergodicity of hybrid samplers. *J. Appl. Probab.* **40** (2003) 123–146. MR1953771

[26] C.-D. Fuh and T. L. Lai. Asymptotic expansions in multidimensional Markov renewal theory and first passage times for Markov random walks. *Adv. in Appl. Probab.* **33** (2001) 652–673. MR1860094

[27] M. Fukushima and M. Hitsuda. On a class of Markov processes taking values on lines and the central limit theorem. *Nagoya Math. J.* **30** (1967) 47–56. MR0216568

[28] H. Ganidis, B. Roynette and F. Simonot. Convergence rate of some semi-groups to their invariant probability. *Stochastic Process. Appl.* **79** (1999) 243–263. MR1671843

[29] V. Genon-Catalot, T. Jeantheau and C. Larédo. Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli* **6** (2000) 1051–1079. MR1809735

[30] P. W. Glynn and W. Whitt. Limit theorems for cumulative processes. *Stochastic Process. Appl.* **47** (1993) 299–314. MR1239842

[31] P. W. Glynn and W. Whitt. Necessary conditions in limit theorems for cumulative processes. *Stochastic Process. Appl.* **98** (2002) 199–209. MR1887533

[32] B. Goldys and B. Maslowski. Exponential ergodicity for stochastic reaction-diffusion equations. In *Stochastic Partial Differential Equations and Applications – VII* 115–131. Chapman & Hall/CRC, Boca Raton, FL, 2006. MR2227225

[33] B. Goldys and B. Maslowski. Lower estimates of transition densities and bounds on exponential ergodicity for stochastic PDE's. *Ann. Probab.* **34** (2006) 1451–1496. MR2257652

[34] M. I. Gordin. On the central limit theorem for stationary Markov processes. *Soviet Math. Dokl.* **19** (1978) 392–394.

[35] S. Gouëzel. Characterization of weak convergence of Birkhoff sums for Gibbs–Markov maps. Preprint, 2008.

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

[37] J.-B. Gravereaux and J. Ledoux. Poisson approximation for some point processes in reliability. *Adv. in Appl. Probab.* **36** (2004) 455–470. MR2058145

[38] S. Grigorescu and G. Oprican. Limit theorems for $J - X$ processes with a general state space. *Probab. Theory Related Fields* **35** (1976) 65–73. MR0413274

[39] D. Guibourg and L. Hervé. A renewal theorem for strongly ergodic Markov chains in dimension $d \geq 3$ and in the centered case. *Potential Anal.* **34** (2011) 385–410.

[40] Y. Guivarc'h. Application d'un théorème limite local à la transcience et à la récurrence de marches aléatoires. In *Théorie du potentiel (Orsay, 1983)* 301–332. *Lecture Notes in Math.* **1096**. Springer, Berlin, 1984. MR0890364

[41] Y. Guivarc'h. Limit theorems for random walks and products of random matrices. In *Proceedings of the CIMPA-TIFR School on Probability Measures on Groups, Mumbai 2002* 257–332. *TIFR Studies in Mathematics Series*. Tata Institute of Fundamental Research, Mumbai, India, 2002.

[42] Y. Guivarc'h and J. Hardy. Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov. *Ann. Inst. H. Poincaré Probab. Statist.* **24** (1988) 73–98. MR0937957

[43] O. Häggström. Acknowledgement of priority concerning "On the central limit theorem for geometrically ergodic Markov chains." *Probab. Theory Related Fields* **135** (2006) 470. MR2240696

[44] H. Hennion and L. Hervé. *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness*. *Lecture Notes in Math.* **1766**. Springer, Berlin, 2001. MR1862393

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

[46] L. Hervé. Théorème local pour chaînes de Markov de probabilité de transition quasi-compacte. Applications aux chaînes v-géométriquement ergodiques et aux modèles itératifs. *Ann. Inst. H. Poincaré Probab. Statist.* **41** (2005) 179–196. MR2124640

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

[48] L. Hervé, J. Ledoux and V. Patilea. A Berry–Esseen theorem on *M*-estimators for geometrically ergodic Markov chains. *Bernoulli* (2012). To appear.

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

[50] M. Hitsuda and A. Shimizu. The central limit theorem for additive functionals of Markov processes and the weak convergence to Wiener measure. *J. Math. Soc. Japan* **22** (1970) 551–566. MR0273688

[51] H. Holzmann. Martingale approximations for continuous-time and discrete-time stationary Markov processes. *Stochastic Process. Appl.* **115** (2005) 1518–1529. MR2158018

[52] I. A. Ibragimov. A note on the central limit theorem for dependent random variables. *Theory Probab. Appl.* **20** (1975) 135–141. MR0362448

[53] I. A. Ibragimov and Y. V. Linnik. *Independent and Stationary Sequences of Random Variables*. Walters-Noordhoff, the Netherlands, 1971. MR0322926

[54] M. Jara, T. Komorowski and S. Olla. Limit theorems for additive functionals of a Markov chain. *Ann. Appl. Probab.* **19** (2009) 2270–2300. MR2588245

[55] S. F. Jarner and E. Hansen. Geometric ergodicity of Metropolis algorithms. *Stochastic Process. Appl.* **85** (2000) 341–361. MR1731030

[56] A. Jobert and L. C. G. Rogers. Option pricing with Markov-modulated dynamics. *SIAM J. Control Optim.* **44** (2006) 2063–2078. MR2248175

[57] G. L. Jones. On the Markov chain central limit theorem. *Probab. Surv.* **1** (2004) 299–320. MR2068475

[58] N. V. Kartashov. Determination of the spectral ergodicity exponent for the birth and death process. *Ukrain. Math. J.* **52** (2000) 1018–1028. MR1817318

[59] J. Keilson and D. M. G. Wishart. A central limit theorem for processes defined on a finite Markov chain. *Proc. Cambridge Philos. Soc.* **60** (1964) 547–567. MR0169271

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

[61] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** (1986) 1–19. MR0834478

[62] P. Lezaud. Chernoff and Berry–Esseen inequalities for Markov processes. *ESAIM Probab. Stat.* **5** (2001) 183–201. MR1875670

[63] T. M. Liggett. Exponential L_2 convergence of attractive reversible nearest particle systems. *Ann. Probab.* **17** (1989) 403–432. MR0985371

[64] N. Limnios and G. Oprican. *Semi-Markov Processes and Reliability.* Birkhauser Boston Inc., Boston, MA, 2001. MR1843923

[65] N. Maigret. Théorème de limite centrale fonctionnel pour une chaîne de Markov récurrente au sens de Harris et positive. *Ann. Inst. H. Poincaré Probab. Statist.* **14** (1978) 425–440. MR0523221

[66] M. Maxwell and M. Woodroofe. A local limit theorem for hidden Markov chains. *Statist. Probab. Lett.* **32** (1997) 125–131. MR1436857

[67] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, London, 1993. MR1287609

[68] S. V. Nagaev. Some limit theorems for stationary Markov chains. *Theory Probab. Appl.* **11** (1957) 378–406.

[69] J. Neveu. Une généralisation des processus à accroissements positifs indépendants. *Abh. Math. Sem. Univ. Hambourg* **25** (1961) 36–61. MR0130714

[70] S. Özekici and R. Soyer. Reliability modeling and analysis in random environments. In *Mathematical Reliability: An Expository Perspective* 249–273. Kluwer Acad. Publ., Boston, MA, 2004. MR2065488

[71] A. Pacheco and N. U. Prabhu. Markov-additive processes of arrivals. In *Advances in Queueing* 167–194. CRC, Boca Raton, FL, 1995. MR1395158

[72] A. Pacheco, L. C. Tang and N. U. Prabhu. *Markov-Modulated Processes & Semiregenerative Phenomena.* World Sci. Publishing, Hackensack, NJ, 2009. MR2484312

[73] M. Peligrad. On the central limit theorem for ρ-mixing sequences of random variables. *Ann. Probab.* **15** (1987) 1387–1394. MR0905338

[74] J. Pfanzagl. The Berry–Esseen bound for minimum contrast estimates. *Metrika* **17** (1971) 81–91. MR0295467

[75] M. Pinsky. Differential equations with a small parameter and the central limit theorem for functions defined on a finite Markov chain. *Probab. Theory Related Fields* **9** (1968) 101–111. MR0228067

[76] B. L. S. P. Rao. On the rate of convergence of estimators for Markov processes. *Probab. Theory Related Fields* **26** (1973) 141–152. MR0339420

[77] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, 3rd edition. Springer-Verlag, Berlin, 1999.

[78] G. O. Roberts and J. S. Rosenthal. Geometric ergodicity and hybrid Markov chains. *Electron. Commun. Probab.* **2** (1997) 13–25. MR1448322

[79] G. O. Roberts and J. S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probab. Surv.* **1** (2004) 20–71. MR2095565

[80] G. O. Roberts and R. L. Tweedie. Geometric L^2 and L^1 convergence are equivalent for reversible Markov chains. *J. Appl. Probab.* **38A** (2001) 37–41. MR1915532

[81] M. Rosenblatt. *Markov Processes. Structure and Asymptotic Behavior.* Springer-Verlag, New York, 1971. MR0329037

[82] V. T. Stefanov. Exact distributions for reward functions on semi-Markov and Markov additive processes. *J. Appl. Probab.* **43** (2006) 1053–1065. MR2274636

[83] J. L. Steichen. A functional central limit theorem for Markov additive processes with an application to the closed Lu–Kumar network. *Stoch. Models* **17** (2001) 459–489. MR1871234

[84] A. Touati. Théorèmes de limite centrale fonctionnels pour les processus de Markov. *Ann. Inst. H. Poincaré Probab. Statist.* **19** (1983) 43–55. MR0699977

[85] A. W. van der Vaart. *Asymptotic Statistics.* Cambridge Univ. Press, Cambridge, 1998. MR1652247

[86] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probab. Theory Related Fields* **128** (2004) 255–321. MR2031227