

LIMIT THEOREMS FOR SOME FUNCTIONALS WITH HEAVY TAILS OF A DISCRETE TIME MARKOV CHAIN

PATRICK CATTIAUX¹ AND MAWAKI MANOU-ABI¹

Abstract. Consider an irreducible, aperiodic and positive recurrent discrete time Markov chain $(X_n, n \geq 0)$ with invariant distribution μ . We shall investigate the long time behaviour of some functionals of the chain, in particular the additive functional $S_n = \sum_{i=1}^n f(X_i)$ for a possibly non square integrable function f . To this end we shall link ergodic properties of the chain to mixing properties, extending known results in the continuous time case. We will then use existing results of convergence to stable distributions, obtained in [M. Denker and A. Jakubowski, *Stat. Probab. Lett.* **8** (1989) 477–483; M. Tyran-Kaminska, *Stochastic Process. Appl.* **120** (2010) 1629–1650; D. Krizmanic, Ph.D. thesis (2010); B. Basrak, D. Krizmanic and J. Segers, *Ann. Probab.* **40** (2012) 2008–2033] for stationary mixing sequences. Contrary to the usual \mathbb{L}^2 framework studied in [P. Cattiaux, D. Chafai and A. Guillin, *ALEA, Lat. Am. J. Probab. Math. Stat.* **9** (2012) 337–382], where weak forms of ergodicity are sufficient to ensure the validity of the Central Limit Theorem, we will need here strong ergodic properties: the existence of a spectral gap, hyperboundedness (or hypercontractivity). These properties are also discussed. Finally we give explicit examples.

Mathematics Subject Classification. 60F05, 60F17, 60J05, 60E07.

Received December 19, 2012. Revised May 17, 2013.

1. INTRODUCTION

On some probability space (Ω, \mathbb{P}) , consider a discrete time Markov chain $\{X_n, n \geq 0\}$ with transition kernel P , which is irreducible, aperiodic and positive recurrent with invariant probability measure μ , assumed to be ergodic. As usual we denote by \mathbb{P}_x the conditional distribution of \mathbb{P} knowing that $X_0 = x$. Let f be a μ integrable function such that $\int f d\mu = 0$. According to the ergodic theorem the normalized additive functional

$$\frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^n f(X_i)$$

goes to 0 as n goes to infinity, both in $\mathbb{L}^1(\mu)$ and \mathbb{P}_μ almost surely.

When $f \in \mathbb{L}^2(\mu)$, it is expected that S_n/\sqrt{n} (or $S_n/\sqrt{\text{Var}_\mu(S_n)}$) converges in distribution to some gaussian limit (Central Limit Theorem). We refer to [6] for an account of the impressive literature on the subject,

Keywords and phrases. Markov chains, stable limit theorems, stable distributions, log-Sobolev inequality, additive functionals, functional limit theorem.

¹ Institut de Mathématiques de Toulouse. CNRS UMR 5219. Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 09, France. cattiaux@math.univ-toulouse.fr; manoumawaki@gmail.com

as well as results including an anomalous behavior of the Variance (*i.e.* $\text{Var}_\mu(S_n)$ is not of order n), or an anomalous rate of convergence (*i.e.* different of the two aforementioned normalization rates). The functional version (convergence of the full law of the process to some Brownian motion) is also discussed therein. We shall recall some of the relevant results in our situation in Section 3.

As noticed in [6], the key tool is the rate of convergence of $\|P^n f\|_{\mathbb{L}^2(\mu)}$ to 0. This rate of convergence is connected to the mixing properties of the sequence $f(X_n)$ so that one can use the massive literature relying mixing and central limit properties.

In the context of [6], the chain comes from a time continuous Markov process, so that this rate of convergence is studied by using recent results on various functional inequalities. At some point, the derivation of these inequalities strongly uses the time continuous setting, allowing us to consider time derivatives.

In particular, if $f \in \mathbb{L}^p(\mu)$ for some $p > 2$, one can use weak Poincaré inequalities furnishing a quick enough but non exponential decay of the Variance of $P^n f$, controlled by the initial $\mathbb{L}^p(\mu)$ norm of f . When f does not belong to any $\mathbb{L}^p(\mu)$ for $p > 2$, one has to assume the existence of a spectral gap.

The goal of the next Section 2 is, after recalling standard mixing properties, to extend the relevant inequalities to the discrete time setting. In particular we extend the notion of weak Poincaré inequality yielding non exponential decays for the \mathbb{L}^2 norm of $P^n f$. This can be used to mimic what is done in [6] in the situation of a discrete time Markov chain.

If f is not square integrable, the situation for the additive functional, is much more delicate, even in the i.i.d. case. CLT analogues have been proved by Gnedenko, the limiting distribution being a stable one, but require rigid assumptions on the tails of $f(X_i)$. A few extensions to weakly dependent cases, essentially mixing cases, have been obtained in [4, 13, 14, 16, 21], for both the usual and the functional versions. Let us also mention the recent [15] dealing with Markov chains and the very recent [3] which contains some new results on the general stationary case but merely difficult to explicitly explain in our situation.

It is thus natural to show how to apply these mixing results to the case of a Markov chain. Other methods (martingale method, coupling...) are available (see [15]).

Our results for this problem, contained in Section 3, substantially improve upon the ones in [15] on many points, but not on all (see the detailed discussion in Rem. 3.6 and the final section). But, presumably, the main interest here is the method. Actually, once the correspondence between ergodicity and mixing is understood, the only thing to do in order to apply stable convergence results for mixing sequences, is to check the so called “anti-clustering” condition. It is at this point that the contraction property introduced in [15] (*i.e.* Pf belongs to a strictly smaller Lebesgue space than f itself) appears as a natural assumption in order to control covariances.

The price to pay (at least if $f \in \mathbb{L}^p(\mu)$ for $1 \leq p < 2$), as in [15] or for the \mathbb{L}^2 case is that we have to assume the existence of a spectral gap. This is due to the fact that, contrary to the case $p > 2$, we have to precisely control the tails, not only from above but also from below since what is required is an equivalent for these tails; *i.e.* controlling the \mathbb{L}^p norm of $P^n f$ by the stronger $\mathbb{L}^{p'}$ norm of f for some $p' > p$ is not enough to get the correct rate of convergence to a stable distribution. Notice that deviation bounds are studied in [7].

The aforementioned contraction property for f is not easy to prove for one chosen function. A stronger form of this contraction property, is that P maps continuously $\mathbb{L}^p(\mu)$ into $\mathbb{L}^{p'}(\mu)$ for some (or for all) $p' > p$. This is a well known property called “hyperboundedness” (or hypercontractivity). For continuous time semi-groups this property is equivalent to a logarithmic Sobolev inequality. Again in the discrete time setting, these connexions are not so clear. The final section is devoted to describe typical situations where we can show hyperboundedness. The standard situation of the usual birth and death chain on the half line is finally described. In this case, generically, the stable invariance principle does not hold.

2. MIXING AND QUANTITATIVE ERGODICITY FOR MARKOV CHAINS

As said in the introduction we shall recall in this section the relationship between mixing and quantitative ergodic properties of the chain. We then will give some explicit conditions ensuring the control of the decay to equilibrium, in terms of functional inequalities.

2.1. Mixing and quantitative ergodicity

Recall some usual mixing coefficients:

Definition 2.1. Let \mathcal{F}_j (resp. \mathcal{G}_j) be respectively the backward (or the past) and the forward (or the future) σ -fields generated by X_n for $0 \leq n \leq j$ (resp. $j \leq n$).

- The strong mixing coefficient $\alpha(n)$ is defined as:

$$\begin{aligned} \alpha(n) &= \sup_j \left\{ \sup_{A,B} (\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)), A \in \mathcal{F}_j, B \in \mathcal{G}_{j+n} \right\}, \\ &= \frac{1}{4} \sup_j \left\{ \sup_{F,G} \text{Cov}(F, G), F \in \mathcal{F}_j (\text{resp. } \mathcal{G}_{n+j}) \text{ measurable and bounded by } 1 \right\}. \end{aligned}$$

If $\lim_{n \rightarrow \infty} \alpha(n) = 0$, the process is strongly mixing.

- The φ -mixing coefficient $\varphi(n)$ is defined as:

$$\varphi(n) = \sup_j \left\{ \sup_{A,B} (\mathbb{P}(B|A) - \mathbb{P}(B)), A \in \mathcal{F}_j, B \in \mathcal{G}_{j+n} \right\}.$$

If $\lim_{n \rightarrow \infty} \varphi(n) = 0$, the process is φ -mixing or uniformly mixing.

- The ρ -mixing coefficient $\rho(n)$ is defined as the maximal correlation coefficient, *i.e.*

$$\rho(n) = \sup_j \left\{ \sup_{F,G} \text{Corr}(F, G), F \in \mathbb{L}^2(\mathcal{F}_j), G \in \mathbb{L}^2(\mathcal{G}_{n+j}) \right\}.$$

If $\lim_{n \rightarrow \infty} \rho(n) = 0$ the process is ρ -mixing.

Of course for strictly stationary sequences (*i.e.* such that, for all $n \geq 0$, the law of $(X_{n+j})_{j \geq 0}$ is the same as the one of $(X_j)_{j \geq 0}$) the supremum on j is irrelevant.

It is worth noticing that the mixing coefficients of $Z_n = f(X_n)$ are smaller than the corresponding ones for X_n . Also recall that

$$2\alpha(n) \leq \varphi(n), \quad 4\alpha(n) \leq \rho(n), \quad \rho(n) \leq 2\sqrt{\varphi(n)}. \quad (2.1)$$

Let $(X_n)_{n \geq 0}$ be an irreducible, aperiodic and positive recurrent Markov chain with unique invariant probability measure μ . We denote by P the transition kernel of the chain. Recall that P is a bounded operator in all $\mathbb{L}^p(\mu)$'s ($p \geq 1$), with operator norm equal to 1 (*i.e.* a contraction). We also introduce the adjoint operator P^* , defined by

$$\int f P g \, d\mu = \int g P^* f \, d\mu$$

for f and g square integrable w.r.t. μ , which is again a contraction.

The following definition introduces some way to control the ergodic decay to equilibrium.

Definition 2.2. For $r, p \geq 1$ we define

$$\alpha_{r,p}(n) = \text{Sup} \left\{ \| P^n g \|_{\mathbb{L}^p(\mu)}, \forall g : \| g \|_{\mathbb{L}^r(\mu)} = 1 \text{ and } \int g \, d\mu = 0 \right\}.$$

We define similarly $\alpha_{p,r}^*(n)$.

There are of course some relationships between these decay rates. Indeed, if we look at the operator

$$Qf = Pf - \int f \, d\mu$$

it is easily seen that $Q^n f = P^n f - \int f d\mu$. Of course the operator norm of Q acting on \mathbb{L}^p spaces satisfies $\|Q\|_{r,p} \leq 2\alpha_{r,p}$ and $\|Q\|_{p,p} \leq 2$ for all $1 \leq p \leq +\infty$. Hence, according to Riesz–Thorin interpolation theorem

$$\alpha_{r,p} \leq 2\alpha_{r_1,p_1}^t \alpha_{r_2,p_2}^{1-t}$$

if

$$\frac{1}{r} = \frac{t}{r_1} + \frac{1-t}{r_2} \text{ and } \frac{1}{p} = \frac{t}{p_1} + \frac{1-t}{p_2} \text{ for some } 0 \leq t \leq 1.$$

Of course

$$\alpha_{r,p}^* = \alpha_{\frac{p}{p-1}, \frac{r}{r-1}}.$$

Remark 2.1. Though it is not really important in most applications, one cannot directly use the Riesz–Thorin interpolation theorem with P in restriction to the subspace of zero mean functions, contrary to what is done in some papers (see *e.g.* (2.7) in [15]). That is why an extra factor 2 (the operator norm of Q) has to appear. For a more accurate discussion see [9].

In order to compare mixing properties and rate of convergence for the stationary chain (*i.e.* assuming that X_0 is distributed according to μ), consider F and G , respectively \mathcal{F}_j and \mathcal{G}_{j+n} measurable and centered (for instance $F = \mathbb{1}_A - \mu(A)$ and $G = \mathbb{1}_B - \mu(B)$).

Define f (resp. g) by $\mathbb{E}_\mu(F|X_j) = f(X_j)$ (resp. $\mathbb{E}_\mu(G|X_{j+n}) = g(X_{j+n})$). Then

$$\text{Cov}_\mu(F, G) = \int P^n g f d\mu = \int P^{n/2} g (P^*)^{n/2} f d\mu, \tag{2.2}$$

the latter being true provided n is even. Since f and g are still centered we immediately deduce (as in [6])

Proposition 2.1. *For all n , if $[n/2]$ denotes the integer part of $n/2$,*

- (1) $\alpha_{\infty,2}^2(n) \vee (\alpha^*)_{\infty,2}^2(n) \leq 4\alpha(n) \leq \alpha_{\infty,2}([n/2]) \alpha_{\infty,2}^*([n/2])$.
- (2) *Either $\alpha_{2,2}(n) = 1$ for all n or $\alpha_{2,2}(n) \leq c e^{-\lambda n}$ for some $\lambda > 0$. In the second case*

$$\alpha_{2,2}^2(n) = (\alpha^*)_{2,2}^2(n) \leq \rho(n) \leq c\alpha_{2,2}(n).$$

- (3) $\varphi(n) \leq \alpha_{1,\infty}^2([n/2])$.

Proof. For the left hand side of (1) we may choose $F = P^n f(X_0)$ and $G = f(X_n)$ for a centered f bounded by 1. It follows $\int (P^n f)^2 d\mu = \text{Cov}(F, G) \leq 4\alpha(n)$. The right hand side is immediate, thanks to (2.2) and the fact that α is non increasing.

For (2) first note the equality $\alpha_{2,2}(n) = \alpha_{2,2}^*(n)$. Hence $\alpha_{2,2}(k) \alpha_{2,2}^*(k) = \alpha_{2,2}(2k)$. Now, using the Markov property it is immediate that if $\alpha_{2,2}(n_0) = e^{-\lambda} < 1$ for some n_0 , $\alpha_{2,2}(n) \leq c e^{-n\lambda}$ for all n . Hence, $\alpha_{2,2}(k+1) \alpha_{2,2}^*(k) \leq c \alpha_{2,2}(2k+1)$. The right hand side of (2) follows, using (2.2). The left hand side is similar to (1).

For (3) just use the first equality in (2.2). □

One can improve the results in the symmetric case (the proof is left to the reader)

Proposition 2.2. *Assume that in addition μ is symmetric, i.e. $\int f P g d\mu = \int g P f d\mu$ for nice f and g . Then*

- (1) $\alpha_{\infty,2}^2(n+1) \leq 4\alpha(2n+1) \leq 4\alpha(2n) = \alpha_{\infty,2}^2(n)$.
- (2) $\alpha_{2,2}^2(n+1) \leq \rho(2n+1) \leq \rho(2n) = \alpha_{2,2}^2(n)$.

2.2. Functional inequalities in the discrete time setting

To conclude this section we shall give explicit criteria for the estimation of the convergence rate. These criteria (and much more) are well known in the continuous time case. In the latter situation proofs are often based on differentiation in time.

We begin with the renowned spectral gap estimate which is trivial in this situation

Proposition 2.3. *The following are equivalent*

- (1) $\forall f \in \mathbb{L}^2(\mu), \text{Var}_\mu(Pf) \leq e^{-2\lambda} \text{Var}_\mu(f),$
- (2) $\forall f \in \mathbb{L}^2(\mu)$ and all $n, \text{Var}_\mu(P^n f) \leq e^{-2\lambda n} \text{Var}_\mu(f),$
- (3) there exists C (necessarily larger than 1) such that $\forall f \in \mathbb{L}^2(\mu),$

$$\text{Var}_\mu(f) \leq C \langle (I - P^*P)f, f \rangle := \int (f - P^*Pf) f \, d\mu.$$

The inequality in (3) is called a Poincaré inequality, and if it holds, $e^{-\lambda} = \frac{C-1}{C}.$

One of the particular feature of the symmetric situation is that any kind of exponential decay will imply the previous contraction property of the variance. Here is the analogue of Lemma 2.12 in [10].

Lemma 2.1. *Assume that μ is symmetric. Let \mathcal{C} be a dense subset of $\mathbb{L}^2(\mu).$ Suppose that there exists $\beta > 0,$ and, for any $f \in \mathcal{C},$ a constant c_f such that:*

$$\forall n, \quad \text{Var}_\mu(P^n f) \leq c_f e^{-\beta n}.$$

Then

$$\forall f \in \mathbb{L}^2(\mu), \forall n, \quad \text{Var}_\mu(P^n f) \leq e^{-\beta n} \text{Var}_\mu(f).$$

In particular, in the symmetric situation, if $\alpha_{\infty,2}(n) \leq c e^{-\beta n},$ then $\alpha_{2,2}(n) \leq e^{-\beta n}.$

Proof. Thanks to symmetry,

$$\begin{aligned} \langle P^n f, P^n f \rangle &= \langle P^{n+1} f, P^{n-1} f \rangle \\ &\leq \langle P^{n+1} f, P^{n+1} f \rangle^{1/2} \langle P^{n-1} f, P^{n-1} f \rangle^{1/2}. \end{aligned}$$

If $h(n) = \ln(\langle P^n f, P^n f \rangle) + \beta n,$ it follows that

$$h(n) \leq \frac{1}{2} h(n+1) + \frac{1}{2} h(n-1),$$

which is some convexity property similar to Lemma 2.11 in [10]. In particular, if for some n and some $\varepsilon > 0$ it holds $h(n) \geq h(n-1) + \varepsilon,$ this convexity property implies $h(n+1) \geq h(n) + \varepsilon,$ hence h grows linearly at infinity. Replacing $f \in \mathcal{C}$ by $f - \int f \, d\mu,$ we deduce from the hypothesis that h is bounded. Hence, h is non increasing, $h(n) \leq h(0),$ and the lemma is proved. \square

It can be shown that this result is not true in the non-symmetric case.

It can be interesting to obtain some non exponentially decreasing controls of $\alpha_{\infty,2}.$ This can be done in terms of weak Poincaré inequalities similar to the continuous time case where they were introduced in [20].

To understand what happens remark that for a centered $f,$

$$\text{Var}_\mu(f) = \langle (I - P^*P)f, f \rangle + \text{Var}_\mu(Pf). \tag{2.3}$$

In the symmetric case, according to the convexity property in Lemma 2.1,

$$\langle Pf, Pf \rangle \leq \langle f, f \rangle^{1/2} \langle P^2 f, P^2 f \rangle^{1/2} \leq \frac{1}{2} \langle f, f \rangle + \frac{1}{2} \langle P^2 f, P^2 f \rangle$$

so that

$$\langle (I - P^*P)f, f \rangle \geq \langle (I - P^*P)Pf, Pf \rangle.$$

We may thus iterate in (2.3) and obtain

$$\begin{aligned} \text{Var}_\mu(f) &\leq n \langle (I - P^*P)f, f \rangle + \text{Var}_\mu(P^n f) \\ &\leq n \langle (I - P^*P)f, f \rangle + \alpha_{\infty,2}^2(n) \| f - \int f \, d\mu \|_\infty^2. \end{aligned} \tag{2.4}$$

Since $\alpha_{\infty,2}^2$ is non increasing, we may take the linear interpolation of this function between n and $n + 1$, *i.e.*

$$\alpha_{\infty,2}^2(s) = (n + 1 - s)\alpha_{\infty,2}^2(n) + (s - n)\alpha_{\infty,2}^2(n + 1)$$

which is still non increasing, and then consider the left inverse of this function, denoted by β . It thus holds for all $0 < s \leq 1$,

$$\text{Var}_\mu(f) \leq \beta(s) \langle (I - P^*P)f, f \rangle + s \| f - \int f \, d\mu \|_\infty^2. \tag{2.5}$$

This kind of inequality is called a weak Poincaré inequality.

Conversely, assume that (2.5) is true for some non increasing function β . We do no more need any symmetry assumption. Then, thanks to the contraction property of P , for any centered and bounded f it holds

$$\begin{aligned} \langle P^n f, P^n f \rangle &\leq (1 - (1/\beta(s))) \langle P^{n-1} f, P^{n-1} f \rangle + (s/\beta(s)) \| f - \int f \, d\mu \|_\infty^2 \\ &\leq (1 - (1/\beta(s)))^n \langle f, f \rangle + (s/\beta(s)) \| f - \int f \, d\mu \|_\infty^2 \sum_{j=0}^{n-1} (1 - (1/\beta(s)))^j \\ &\leq (1 - (1/\beta(s)))^n \langle f, f \rangle + s (1 - (1 - (1/\beta(s)))^n) \| f - \int f \, d\mu \|_\infty^2 \\ &\leq (1 - (1/\beta(s)))^n \langle f, f \rangle + s \| f - \int f \, d\mu \|_\infty^2. \end{aligned}$$

We thus have for any bounded f ,

$$\text{Var}_\mu(P^n f) \leq (s + (1 - (1/\beta(s)))^n) \| f - \int f \, d\mu \|_\infty^2,$$

for all $0 < s \leq 1$; so that optimizing on s we obtain

$$\alpha_{\infty,2}^2(n) \leq 2s_n \quad \text{where } s_n \text{ is defined by } (1 - (1/\beta(s_n)))^n = s_n. \tag{2.6}$$

Let us summarize what we have obtained

Proposition 2.4. *If (2.5) holds for some non increasing function β , then $\alpha_{\infty,2}(n)$ goes to 0 as $n \rightarrow +\infty$ as explained in (2.6).*

Conversely, in the symmetric situation, any decay to 0 of $\alpha_{\infty,2}(n)$ will imply some weak Poincaré inequality (2.5).

Note that a similar result holds for the adjoint operator P^* which satisfies the same weak Poincaré inequality (since $\langle (I - P^*P)f, f \rangle = \langle (I - PP^*)f, f \rangle$).

In general however, we do not know wether $\alpha_{\infty,2}$ and $\alpha_{\infty,2}^*$ are equal or have the same behaviour. This is discussed in Remark 4.6 in [6]. In particular it is shown therein that if they are slowly decreasing, $\alpha_{\infty,2}$ and $\alpha_{\infty,2}^*$ are of the same order.

Remark 2.2. Instead of looking at P^*P one can use another “symmetrization” *i.e.* $Q = \frac{1}{2}(P + P^*)$ which is often used to introduce the continuous time semi-group $Q_t = e^{t(Q-I)}$ whose generator is $L = Q - I$. In this case, the associated weak Poincaré inequality is

$$\text{Var}_\mu(f) \leq \beta(s) \langle -Lf, f \rangle + s \|f - \int f d\mu\|_\infty^2.$$

Note that this weak Poincaré inequality is weaker than (2.5).

Indeed if (2.5) is satisfied, it holds for all centered f 's,

$$\langle Pf, Pf \rangle \leq \frac{\beta(s) - 1}{\beta(s)} \langle f, f \rangle + \frac{s}{\beta(s)} \|f\|_\infty^2.$$

Hence

$$\begin{aligned} \langle Pf, f \rangle &\leq \langle Pf, Pf \rangle^{\frac{1}{2}} \langle f, f \rangle^{\frac{1}{2}} \\ &\leq \left(\frac{\beta(s) - 1}{\beta(s)} \right)^{\frac{1}{2}} \langle f, f \rangle + \left(\frac{s}{\beta(s)} \right)^{\frac{1}{2}} \|f\|_\infty \langle f, f \rangle^{\frac{1}{2}} \\ &\leq \left(\frac{\beta(s) - 1}{\beta(s)} \right)^{\frac{1}{2}} \langle f, f \rangle + \left(\frac{s}{\beta(s)} \right)^{\frac{1}{2}} \|f\|_\infty^2, \end{aligned}$$

so that

$$\left(1 - \left(\frac{\beta(s) - 1}{\beta(s)} \right)^{\frac{1}{2}} \right) \text{Var}_\mu(f) \leq \langle -Lf, f \rangle + \left(\frac{s}{\beta(s)} \right)^{\frac{1}{2}} \|f\|_\infty^2.$$

We thus obtain that μ satisfies a weak Poincaré inequality, with $\langle -Lf, f \rangle$ as energy form. In particular if the Poincaré inequality in Proposition 2.3 holds for some constant C , a Poincaré inequality holds with $\langle -Lf, f \rangle$ as energy form with constant $C' = \sqrt{C}/(\sqrt{C} - \sqrt{C-1})$.

Nevertheless, in the symmetric situation, both Poincaré inequalities coincide. Indeed

$$\text{Var}_\mu(f) \leq C \langle -Lf, f \rangle = C \langle (I - P)f, f \rangle$$

means that, for all $f \in \mathbb{L}^2(\mu)$ with μ zero mean,

$$\langle Pf, f \rangle \leq \frac{C-1}{C} \langle f, f \rangle,$$

so that the operator norm of P restricted to the closed hyperplane $\int f d\mu = 0$ is less than $(C-1)/C$. It follows that in restriction to this hyperplane,

$$\langle Pf, Pf \rangle \leq \left(\frac{C-1}{C} \right)^2 \langle f, f \rangle,$$

hence the Poincaré inequality

$$\text{Var}_\mu(f) \leq C' \langle (I - P^2)f, f \rangle$$

holds with $C' = C^2/(2C-1)$. So up to the constants, both Poincaré inequalities are the same.

3. CONVERGENCE TO STABLE DISTRIBUTIONS FOR ADDITIVE FUNCTIONALS

In this section we shall study the asymptotic behaviour of $S_n = \sum_{i=1}^n f(X_i)$, where $f \in \mathbb{L}^p(\mu)$ for some $0 \leq p < 2$. For simplicity we shall sometimes write $f(X_i) = Z_i$. In particular, under \mathbb{P}_μ , the sequence $(Z_i)_{i \in \mathbb{N}}$ is strictly stationary. Recall that if $f \in \mathbb{L}^1(\mu)$ we assume in addition that $\int f \, d\mu = 0$.

When $f \in \mathbb{L}^2(\mu)$, it is expected that $n^{-\frac{1}{2}} S_n$ converges in distribution to some gaussian law. Actually the stronger functional CLT, telling that the process

$$t \mapsto n^{-\frac{1}{2}} S_{[nt]}$$

converges in distribution (for the usual topology of continuous paths) to a Brownian motion (with a positive variance) holds as soon as

$$\sum_n n^{-\frac{1}{2}} \|P^n f\|_{\mathbb{L}^2(\mu)} < +\infty. \tag{3.1}$$

For this result and much more sophisticated ones we refer to the survey paper [18] and to [6]. Notice that if $f \in \mathbb{L}^p(\mu)$ for some $p \geq 2$, the previous condition is satisfied as soon as $\sum_n n^{-\frac{1}{2}} \alpha_{p,2}(n) < +\infty$; in particular, according to the discussion after Definition 2.2 as soon as $\sum_n n^{-\frac{1}{2}} \alpha_{\infty,2}^{\frac{p-2}{p}}(n) < +\infty$. If f does not belong to any \mathbb{L}^p space for $p > 2$, the (3.1) is useful under the ρ -mixing assumption, hence if the chain has a spectral gap (in \mathbb{L}^2).

For $p < 2$ the situation is much more delicate. We first have to recall some definitions

Definition 3.3. Z_1 is regularly varying of index $\eta > 0$ if there exists some $c \in [0, 1]$ such that for all $x > 0$,

$$\lim_{u \rightarrow +\infty} \frac{\mathbb{P}(Z_1 > ux)}{\mathbb{P}(|Z_1| > u)} = cx^{-\eta} \quad \text{and} \quad \lim_{u \rightarrow +\infty} \frac{\mathbb{P}(Z_1 < -ux)}{\mathbb{P}(|Z_1| > u)} = (1 - c)x^{-\eta}.$$

If Z_1 is regularly varying, there exists a slowly varying function L (i.e. $\lim_{u \rightarrow +\infty} \frac{L(ut)}{L(u)} = 1$ for all $t > 0$) such that for all $x > 0$,

$$\mathbb{P}(|Z_1| > x) = x^{-\eta} L(x).$$

In particular $Z_1 \in \mathbb{L}^\eta$ if and only if $\int_1^{+\infty} x^{-1} L(x) \, dx < +\infty$. Otherwise $Z_1 \in \mathbb{L}^{\eta'}$ for all $\eta' < \eta$.

If we write $\mathbb{P}(Z_1 > x) = x^{-\eta} L_+(x)$, we see that if Z_1 is regularly varying of index η ,

$$\lim_{u \rightarrow +\infty} \frac{L_+(u)}{L(u)} = c.$$

It follows that L_+ is also slowly varying, the same for the analogous L_- .

From now on we assume that $Z_1 = f(X_1)$ is regularly varying of index η .

Now choose b_n such that

$$\lim_{n \rightarrow +\infty} n \mathbb{P}(|Z_1| > b_n) = 1. \tag{3.2}$$

As we shall recall below, convergence to stable laws (or processes) for mixing sequences require some kind of centering, namely we shall consider

$$T_n = S_n - n \mathbb{E}_\mu [Z_1 \mathbb{1}_{|Z_1| \leq b_n}] = S_n - n c_n \tag{3.3}$$

or

$$T_n(t) = S_{[nt]} - n t c_n,$$

and look at the convergence of T_n/b_n or $T_n(\cdot)/b_n$. It is thus interesting to look at the asymptotic behaviour of

$$\frac{n c_n}{b_n} = \frac{n}{b_n} \mathbb{E}_\mu [Z_1 \mathbb{1}_{|Z_1| \leq b_n}].$$

When $0 < \eta < 1$ this is done in [16] Remark 2.17, where it is shown that

$$\lim_{n \rightarrow +\infty} \frac{n c_n}{b_n} = \frac{\eta}{1 - \eta} (2c - 1).$$

(recall that c is defined in Def. 3.3.) The proof is a simple application of Karamata’s theorem.

Now assume that $1 \leq \eta < 2$ and that $\int f d\mu = 0$. We thus have

$$\frac{n}{b_n} \mathbb{E}_\mu [Z_1 \mathbb{1}_{|Z_1| \leq b_n}] = - \frac{n}{b_n} \mathbb{E}_\mu [Z_1 \mathbb{1}_{|Z_1| > b_n}].$$

But

$$\frac{n}{b_n} \mathbb{E}_\mu [Z_1^+ \mathbb{1}_{|Z_1| > b_n}] = n \mathbb{P}_\mu(Z_1^+ > b_n) + \frac{n}{b_n} \mathbb{E}_\mu [(Z_1^+ - b_n)^+],$$

and

$$\begin{aligned} \frac{n}{b_n} \mathbb{E}_\mu [(Z_1^+ - b_n)^+] &= \int_{b_n}^{+\infty} \frac{n}{b_n} \mathbb{P}_\mu(Z_1^+ > u) du \\ &= \int_1^{+\infty} n \mathbb{P}_\mu(Z_1^+ > u b_n) du \\ &= \int_1^{+\infty} n b_n^{-\eta} L(b_n) u^{-\eta} \frac{L_+(u b_n)}{L(b_n)} du \\ &= (n b_n^{-\eta} L(b_n)) \int_1^{+\infty} u^{-\eta} \frac{L_+(u b_n)}{L(b_n)} du. \end{aligned}$$

Since $\frac{L_+(u b_n)}{L(b_n)} \rightarrow c$ as $n \rightarrow +\infty$, we may use the bounded convergence theorem and obtain that the previous goes to $c/(\eta - 1)$ provided $\eta > 1$. Hence, using Definition 3.3, we obtain

$$\lim_{n \rightarrow +\infty} \frac{n}{b_n} \mathbb{E}_\mu [Z_1^+ \mathbb{1}_{|Z_1| > b_n}] = c + \frac{c}{\eta - 1} = \frac{c\eta}{\eta - 1}.$$

Summing up with the similar estimate for Z_1^- , we thus have shown

Proposition 3.5. *Assume that Z_1 is regularly varying of index $\eta \in (0, 2)$ but $\eta \neq 1$. Assume in addition that $\mathbb{E}_\mu(Z_1) = 0$ if $\eta > 1$. Then*

$$\lim_{n \rightarrow +\infty} \frac{n c_n}{b_n} = \frac{\eta}{1 - \eta} (2c - 1).$$

If Z_1 is symmetric, $c_n = 0$ for all n .

Remark 3.3. The calculation above, for $\eta > 1$, allows us to replace the symmetry assumption in [13] by the weaker $\mathbb{E}_\mu(Z_1) = 0$ (see condition (v) p. 478 of [13]).

Now we recall definitions for stable distributions and processes. We follow the presentation in [4, 16]. We shall say that X has a η -stable distribution with characteristics (b, c_-, c_+) , if its characteristic function satisfies $\mathbb{E}(e^{itX}) = e^{\psi(t)}$ with

$$\psi(t) = ibt + \int (e^{itu} - 1 - itu \mathbb{1}_{|u| \leq 1}) |u|^{-1-\eta} (c_- \mathbb{1}_{u < 0} + c_+ \mathbb{1}_{u > 0}) du. \tag{3.4}$$

A η -stable process Y_t is a cadlag stochastic process starting from 0, with independent and stationary increments, and such that the distribution of Y_1 is η -stable. We refer to [15, 16, 21] (among others) for more details.

The results in [4, 13, 16, 21] deal with the convergence of $\frac{1}{b_n} T_n$ or $\frac{1}{b_n} T_n(\cdot)$ (recall (3.3)) to a stable distribution or a stable process (for the Skorohod topology), with characteristics $(b, c, 1 - c)$.

According to Proposition 3.5, except for $\eta = 1$, we may replace T_n by S_n up to a change in the characteristics, where b is replaced by $b + \frac{\eta}{\eta-1} (2c - 1)$. Recall that in the symmetric case, $c = 1/2$.

The following anti clustering type condition introduced in [12], appears in both [13] (condition (x) p. 483) and [16] (condition \mathcal{D}'):

$$\lim_{k \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \sum_{j=2}^{[n/k]} n \mathbb{P}_\mu (|Z_1| > \varepsilon b_n, |Z_j| > \varepsilon b_n) = 0 \quad , \quad \forall \varepsilon > 0. \tag{3.5}$$

Remark that

$$\begin{aligned} \sum_{j=2}^{[n/k]} n \mathbb{P}_\mu (|Z_1| > \varepsilon b_n) \mathbb{P}_\mu (|Z_j| > \varepsilon b_n) &= \left(\frac{\mathbb{P}_\mu^2 (|Z_1| > \varepsilon b_n)}{\mathbb{P}_\mu^2 (|Z_1| > b_n)} \right) \frac{[n/k]}{n} n^2 \mathbb{P}_\mu^2 (|Z_1| > b_n) \\ &\leq \frac{1}{k} \left(\frac{\mathbb{P}_\mu^2 (|Z_1| > \varepsilon b_n)}{\mathbb{P}_\mu^2 (|Z_1| > b_n)} \right) \end{aligned}$$

so that

$$\limsup_{n \rightarrow +\infty} \sum_{j=2}^{[n/k]} n \mathbb{P}_\mu (|Z_1| > \varepsilon b_n) \mathbb{P}_\mu (|Z_j| > \varepsilon b_n) \leq \varepsilon^{-\eta} \frac{1}{k},$$

and finally

$$\lim_{k \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \sum_{j=2}^{[n/k]} n \mathbb{P}_\mu (|Z_1| > \varepsilon b_n) \mathbb{P}_\mu (|Z_j| > \varepsilon b_n) = 0.$$

Hence we may replace (3.5) by

$$\lim_{k \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \sum_{j=2}^{[n/k]} n \text{Cov}_\mu (\mathbb{1}_{|Z_1| > \varepsilon b_n}, \mathbb{1}_{|Z_j| > \varepsilon b_n}) = 0 \quad , \quad \forall \varepsilon > 0. \tag{3.6}$$

We finally state the convergence result which is contained in Proposition 3 of [13] in the case of variables, and Theorem 3.7 (p. 93) of [16] or Theorem 1.1 of [21] for processes. We formulate the result in our Markov chain framework

Theorem 3.1. *Let $(X_n)_{n \geq 0}$ be an irreducible, aperiodic and positive recurrent Markov chain with unique invariant probability measure μ . Let f be such that $Z_1 = f(X_1)$ is regularly varying of index $\eta \in (0, 2)$. If f is μ integrable we also assume that $\int f d\mu = 0$.*

For b_n defined by (3.2), we first assume that (3.6) holds.

In addition we assume

- (1) *If $\eta \in (0, 1)$, the chain is strongly mixing,*
- (2) *if $\eta \in (1, 2)$ the chain has a spectral gap,*
- (3) *if $\eta = 1$, f is symmetric, i.e. under \mathbb{P}_μ , $f(X_1)$ and $f(-X_1)$ have the same distribution, and the chain has a spectral gap.*

Then $\frac{1}{b_n} S_n(t) = \frac{1}{b_n} \sum_{j=1}^{[nt]} f(X_j)$ converges (under \mathbb{P}_μ), for the Skorohod topology on any time interval $[0, T]$ to some η -stable process.

We already discussed the form of the characteristics of the stable limit.

For $\eta = 1$, if f is not symmetric, we have to replace S_n by T_n .

Remark 3.4. In general the difference between the assumptions in [13, 21] and [16] concerns the case $\eta \in (1, 2)$, where all authors assume that Z is ρ -mixing, but with ρ satisfying $\sum_j \rho(2^j) < +\infty$ in [13, 21], and $\sum_j \rho([2^{j/3}]) < +\infty$ in [16] (see p. 76, Prop. 2.19). Actually both conditions are the same and equivalent to $\sum_n \frac{\rho(n)}{n} < +\infty$ as it is easy to see.

In our situation of course, ρ -mixing is equivalent to the spectral gap property and thus implies that $\rho(n) = e^{-n\lambda}$ as we have seen in Proposition 2.1 so that this condition is satisfied.

Actually one can reinforce Theorem 3.1 as follows

Theorem 3.2. *In Theorem 3.1, one can replace (3.6) by the following weaker condition: there exist $r_n = o(n)$ growing to infinity such that*

$$\lim_{n \rightarrow +\infty} \sum_{j=2}^{r(n)} n \operatorname{Cov}_\mu \left(\mathbb{1}_{|Z_1| > \varepsilon b_n}, \mathbb{1}_{|Z_j| > \varepsilon b_n} \right) = 0, \quad \forall \varepsilon > 0. \tag{3.7}$$

This condition is condition (3.18) in [16].

One can also replace (3.7) by the following weaker condition:

$\forall \varepsilon > 0$, there exists $r(n)$ and $l(n)$ going to infinity, such that $l(n) = o(r(n))$, $r(n) = o(n)$, $n\alpha(l(n)) = o(r(n))$ where α is the mixing coefficient, and

$$\lim_{n \rightarrow +\infty} \mathbb{P}_\mu \left(\max_{2 \leq j \leq r(n)} |Z_j| > \varepsilon b_n \mid |Z_1| > \varepsilon b_n \right) = 0. \tag{3.8}$$

Furthermore as soon as all other assumptions in Theorem 3.1 are satisfied, (3.8) is also necessary for the convergence to a stable process. This is the main Theorem in [21] (also see Thm. 3.4 in [4]).

We conclude this section by giving explicit sufficient conditions for (3.6) (hence Thm. 3.1) to hold.

Theorem 3.3. *In the situation of Theorem 3.1 we assume that $\eta \in (1, 2)$ and that the chain has a spectral gap. If $P|f|$ or $P^*|f|$ belongs to $\mathbb{L}^{\eta+\beta}(\mu)$ for some $\beta > 0$, then condition (3.6) is satisfied and the conclusion of Theorem 3.1 holds true.*

Proof. Write $g(u) = \mathbb{1}_{|u| > \varepsilon b_n} - \mathbb{P}(|Z_1| > \varepsilon b_n)$. We have

$$\begin{aligned} \operatorname{Cov}_\mu \left(\mathbb{1}_{|Z_1| > \varepsilon b_n}, \mathbb{1}_{|Z_j| > \varepsilon b_n} \right) &= \mathbb{E}_\mu \left(g(Z_1) \mathbb{1}_{|Z_j| > \varepsilon b_n} \right) \\ &= \int (P^*)^{j-2} g P(\mathbb{1}_{|f| > \varepsilon b_n}) \, d\mu \\ &\leq \int |(P^*)^{j-2} g| P(|f|/\varepsilon b_n) \, d\mu \\ &\leq \frac{1}{\varepsilon b_n} \left\| (P^*)^{j-2} g \right\|_{\mathbb{L}^{\frac{\eta+\beta}{\eta+\beta-1}}(\mu)} \left\| P|f| \right\|_{\mathbb{L}^{\eta+\beta}(\mu)}, \end{aligned}$$

after using Hölder's inequality. Now since the chain has a spectral gap, we know that $\alpha_{2,2}^*(j) \leq c e^{-\lambda j}$ for some $\lambda > 0$. It follows that $\alpha_{p,p}^*(j) \leq 2c e^{-\lambda j \frac{2(p-1)}{p}}$ for all $+\infty > p > 1$, in particular for $p = \frac{\eta+\beta}{\eta+\beta-1}$. We denote by $\rho_p^*(j)$ this quantity for p as before. We have thus

$$\begin{aligned} \operatorname{Cov}_\mu \left(\mathbb{1}_{|Z_1| > \varepsilon b_n}, \mathbb{1}_{|Z_j| > \varepsilon b_n} \right) &\leq \frac{1}{\varepsilon b_n} \rho_p^*(j-2) \left\| g \right\|_{\mathbb{L}^{\frac{\eta+\beta}{\eta+\beta-1}}(\mu)} \left\| P|f| \right\|_{\mathbb{L}^{\eta+\beta}(\mu)} \\ &\leq \frac{C(f)}{\varepsilon b_n} \rho_p^*(j-2) \left(\mathbb{P}_\mu(|Z_1| > \varepsilon b_n) \right)^{1-\frac{1}{\eta+\beta}} \\ &\leq \frac{C(f)}{\varepsilon b_n} \rho_p^*(j-2) \left(\varepsilon^{-\eta} b_n^{-\eta} L(\varepsilon b_n) \right)^{1-\frac{1}{\eta+\beta}} \\ &\leq \frac{C(f)}{\varepsilon b_n} \rho_p^*(j-2) \left(\varepsilon^{-\eta} b_n^{-\eta} L(b_n) \frac{L(\varepsilon b_n)}{L(b_n)} \right)^{1-\frac{1}{\eta+\beta}}. \end{aligned}$$

Finally, since $\frac{L(\varepsilon b_n)}{L(b_n)}$ is bounded (w.r.t. n for a fixed ε) we have obtained that

$$\text{Cov}_\mu (\mathbb{1}_{|Z_1|>\varepsilon b_n}, \mathbb{1}_{|Z_j|>\varepsilon b_n}) \leq \frac{C(f, \eta, \beta, \varepsilon)}{b_n} (n b_n^{-\eta} L(b_n)) \rho_p^*(j-2) n^{-1+\frac{1}{\eta+\beta}}. \tag{3.9}$$

Finally

$$\sum_{j=2}^{\lfloor n/k \rfloor} n \text{Cov}_\mu (\mathbb{1}_{|Z_1|>\varepsilon b_n}, \mathbb{1}_{|Z_j|>\varepsilon b_n}) \leq (n b_n^{-\eta} L(b_n)) C(f, \eta, \beta, \varepsilon) \frac{n^{\frac{1}{\eta+\beta}}}{b_n} \sum_{j=2}^{+\infty} \rho_p^*(j-2).$$

But b_n satisfies $\lim_{n \rightarrow +\infty} n^{-\frac{1}{\eta}} b_n L^{-\frac{1}{\eta}}(b_n) = 1$, so that, since L is slowly varying, $b_n^{-1} n^\theta \rightarrow 0$ as $n \rightarrow +\infty$ as soon as $\theta < \frac{1}{\eta}$. It follows that

$$\limsup_{n \rightarrow +\infty} \sum_{j=2}^{\lfloor n/k \rfloor} n \text{Cov}_\mu (\mathbb{1}_{|Z_1|>\varepsilon b_n}, \mathbb{1}_{|Z_j|>\varepsilon b_n}) = 0,$$

so that (3.6) is satisfied.

Exchanging P and P^* the same holds if $P^*|f| \in \mathbb{L}^{\eta+\beta}$. □

The next result contains the $\eta < 1$ case.

Theorem 3.4. *In the situation of Theorem 3.1 we assume that $\eta \in (0, 2)$ and that the chain has a spectral gap. If for some $\eta' \leq \eta$ $f \in \mathbb{L}^{\eta'}(\mu)$ and $P(|f|^{\eta'})$ or $P^*(|f|^{\eta'})$ belongs to $\mathbb{L}^{1+\beta}(\mu)$ for some $\beta > 0$ such that $\eta' > \eta/(1 + \beta)$, then condition (3.6) is satisfied and the conclusion of Theorem 3.1 holds true.*

Proof. The proof follows the same lines as the previous one. The first inequality becomes

$$\text{Cov}_\mu (\mathbb{1}_{|Z_1|>\varepsilon b_n}, \mathbb{1}_{|Z_j|>\varepsilon b_n}) \leq \frac{1}{(\varepsilon b_n)^{\eta'}} \| (P^*)^{j-2} g \|_{\mathbb{L}^{\frac{1+\beta}{\beta}}(\mu)} \| (P|f|^{\eta'}) \|_{\mathbb{L}^{1+\beta}(\mu)},$$

yielding

$$\text{Cov}_\mu (\mathbb{1}_{|Z_1|>\varepsilon b_n}, \mathbb{1}_{|Z_j|>\varepsilon b_n}) \leq (n b_n^{-\eta} L(b_n)) \frac{C(f, \eta, \beta, \varepsilon)}{(b_n)^{\eta'}} \rho_p^*(j-2) n^{-1+\frac{1}{\eta}}, \tag{3.10}$$

this time with $p = 1 + \beta$. The conclusion follows, thanks to the assumption $\eta' > \eta/(1 + \beta)$. □

Remark 3.5. If $\eta \in (0, 1)$ in the previous Theorem we do not need the ρ -mixing condition. We can replace it by $\sum_j \alpha_{(1+\beta)/\beta, q}(j) < +\infty$ for some $q > 1 + \beta$ but this time the condition becomes $\eta' > \eta(1 - \frac{1}{q})$.

Remark 3.6. We can compare our results with Theorems 2.4 and 2.8 in [15].

On one hand, first, these authors assume that $L = 1$ and that the spectral gap exists, while for $\eta < 1$ we can relax this assumption as described in the preceding Remark.

Second we do not assume their Condition 2.3. In the sequel we are using the notations in [15]. Notice that if the invariant measure π is symmetric, this condition with $Q = 0$ implies (because of (2.9) therein) that P^2 has a bounded kernel with respect to π . Indeed $P^2(x, dy) = p_2(x, y)\pi(dy)$ with

$$p_2(x, y) = \int p(x, z)p(z, y)\pi(dz) \leq \left(\int p^2(x, z)\pi(dz) \right)^{1/2} \left(\int p^2(y, z)\pi(dz) \right)^{1/2} \leq C(2).$$

Hence the chain generated by P^2 is ultra bounded, i.e. continuously maps \mathbb{L}^1 functions into bounded ones. Hence, all $P^k f$ for $k \geq 2$ will be bounded.

On the other hand our assumption relies on $P|f|$ instead of Pf (this is certainly mainly irrelevant), and for $\eta < 1$ we add a similar assumption on f which is not present in Theorem 2.4 of [15]. The calculation of the characteristics of the limit seem also to slightly differ, but they are not expressed in the same way.

Finally our results cover the functional limit case (invariance principle).

Remark 3.7. One can be disappointed to have to assume the existence of a spectral gap (though the situation in the \mathbb{L}^2 case is similar). Actually the contraction property is a much stronger assumption, as recently shown by Miclo [19],

Theorem 3.5. *If P is symmetric, ergodic and maps continuously $\mathbb{L}^2(\mu)$ into $\mathbb{L}^p(\mu)$ for some $p > 2$, then there exists a spectral gap.*

The continuous time version of this result is contained in [5]. We shall see in the next section that interpolation will then show that, for all $q > 1$, P maps continuously $\mathbb{L}^q(\mu)$ into $\mathbb{L}^{q'}(\mu)$ for some $q' > q$.

4. EXAMPLES

In this section we shall study some examples and counter-examples.

4.1. Hyperboundedness

Let us discuss the contraction property, *i.e.* the fact that P maps continuously \mathbb{L}^p into $\mathbb{L}^{p'}$ for some $p' > p$. In a continuous time setting this property is known as the hyperboundedness of the associated semi-group and is equivalent to a log-Sobolev inequality (see [1]).

In the discrete time setting this property was not really studied for itself, presumably because except for finite state space, there is no chance for the log-Sobolev inequality

$$\int f^2 \ln \left(\frac{f^2}{\int f^2 d\mu} \right) d\mu \leq C_L \langle (I - \frac{1}{2}(P + P^*))f, f \rangle + D_L \int f^2 d\mu$$

to be satisfied.

Hence there are mainly two generic situations in which we can obtain this property

- (1) $P = P_T$ for some $T > 0$, where $(P_t)_{t \geq 0}$ is a μ reversible continuous time hyperbounded Markov process. In this case μ will satisfy a log-Sobolev inequality where $I - \frac{1}{2}(P + P^*)$ is replaced by the infinitesimal generator of P_t . One can of course replace P by $Q = \frac{1}{2}(P + P^*)$.

If P_t is a diffusion process, μ has to be invariant, not necessarily reversible.

In a similar way, we can look at numerical schemes (like the Euler scheme) for approximating hypercontractive diffusion processes, since in many situations the hypercontractivity property is preserved.

- (2) Assume that $P(x, dy) = p(x, y) \mu(dy)$, with $p \in \mathbb{L}^q(\mu \otimes \mu)$ for some large enough q .

Let us give some more precise examples.

For situation (1) consider for simplicity the diffusion process in \mathbb{R}^d with infinitesimal generator $\Delta - \nabla V \cdot \nabla$ and reversible measure $e^{-V} dx$. If V is uniformly convex at infinity (*i.e.* $\xi \cdot HessV(x)\xi \geq C|\xi|^2$ for some $C > 0$ and all $|x| \geq R$) it is known that the associated semi-group is hypercontractive, *i.e.* hyperbounded with operator norm equal to 1. It follows that $P = P_T$ continuously maps \mathbb{L}^2 into $\mathbb{L}^{2+\beta(T)}$ for some $\beta(T) > 0$ and all $T > 0$. Using Riesz–Thorin interpolation theorem, we see that for all $p \geq 2$, P_T maps continuously \mathbb{L}^p into $\mathbb{L}^{p+\beta(p,T)}$ for some $\beta(p, T) > 0$. By duality we thus have that, for all $\eta \in (1, 2)$, P_T is bounded from \mathbb{L}^η into \mathbb{L}^γ for some $\gamma > \eta$, hence we may apply the results of the previous section. The final argument using interpolation is of course available in all hyperbounded situations.

For situation (2), we have for a non-negative $f \in \mathbb{L}^2$ and $\gamma > 2$,

$$\begin{aligned} \|Pf\|_{\mathbb{L}^\gamma}^\gamma &= \int \left(\int p(x, y) f(y) \mu(dy) \right)^\gamma \mu(dx) \\ &\leq \left(\int f^2 d\mu \right)^{\frac{\gamma}{2}} \left(\int \left(\int p^2(x, y) \mu(dy) \right)^{\frac{\gamma}{2}} \mu(dx) \right) \\ &\leq \left(\int f^2 d\mu \right)^{\frac{\gamma}{2}} \left(\int p^\gamma d\mu \otimes d\mu \right). \end{aligned}$$

Hence provided $x \mapsto \int p^2(x, y)\mu(dy)$ belongs to \mathbb{L}^γ , in particular if $p \in \mathbb{L}^\gamma(\mu \otimes \mu)$, we may apply the same argument as in the previous paragraph.

Notice that this assumption is much weaker than assumption (2.9) in [15] where it is assumed that $x \mapsto \int p^2(x, y)\mu(dy)$ is bounded. In the example of Superdiffusion of energy in a lattice dynamics studied in [15] the kernel p is in fact bounded.

In this latter situation P maps continuously \mathbb{L}^1 into \mathbb{L}^∞ (i.e. is ultrabounded). In particular, since we remarked in Proposition 2.1 (3), that $\varphi(n) \leq \alpha_{1,\infty}^2([n/2])$, the chain is φ -mixing.

4.2. Birth and death chains

Here the state space is \mathbb{N} , and the transition matrix is given by

$$P(x, x + 1) = p_x, \quad P(x, x - 1) = q_x, \quad P(x, x) = 1 - p_x - q_x,$$

for all $x \in \mathbb{N}$, of course $q_0 = 0$. This chain is positively recurrent if and only if

$$\sum_{x \in \mathbb{N}^*} \frac{p_0 p_1 \cdots p_{x-1}}{q_1 \cdots q_x} := \sum_{x \in \mathbb{N}^*} \lambda_x < +\infty,$$

in which case the unique invariant (and actually reversible) distribution is given by $\mu(x) = \mu(0) \lambda_x$ for all $x \in \mathbb{N}^*$.

It is well known (see e.g. [11]) that the Poincaré inequality

$$\text{Var}_\mu(f) \leq C \langle (I - P)f, f \rangle = C \langle -Lf, f \rangle \tag{4.1}$$

holds if and only if

$$\sup_{x \geq 1} \mu([x, +\infty[) \sum_{y \leq x-1} \frac{1}{\mu(y) p_y} < +\infty. \tag{4.2}$$

Others characterizations of geometric ergodicity in the discrete time setting are contained in [22] and are based on zeros of the associated birth and death polynomials, or in [2, 8, 10, 17] and are based on Lyapunov functions or exponential moments of hitting times.

It is not difficult to show that the log-Sobolev inequality is not satisfied (consider functions $\mathbb{1}_x$ with x going to infinity), and that $P(x, dy) = p(x, y)\mu(dy)$ but with p that do not satisfy any nice integrability condition.

But let us come back to Theorem 3.2, in particular to the necessary condition (3.8). If $Z_n = f(X_n)$ we have

$$\begin{aligned} \mathbb{P}_\mu(|Z_2| > \varepsilon b_n | |Z_1| > \varepsilon b_n) &\geq \mathbb{P}_\mu(X_2 = X_1 | |Z_1| > \varepsilon b_n) \\ &\geq \inf\{r_x; f(x) > \varepsilon b_n\}. \end{aligned}$$

In particular if $r_x \geq r > 0$ for all x large enough, (3.8) cannot be satisfied.

But even in the case of the reflected random walk ($p_x = p = 1 - q = q_x$ for all $x > 0$, $p_0 = 1$) for which provided $q > p$, ergodicity holds with a spectral gap, (3.8) is generically not satisfied. Indeed in this case, if for simplicity f is increasing,

$$\begin{aligned} \mathbb{P}_\mu(|Z_2| > a | |Z_1| > a) &\geq \mathbb{P}_\mu(X_2 = X_1 - 1, |X_1| > f^{-1}(a + 1) | |X_1| > f^{-1}(a)) \\ &\geq q \mathbb{P}_\mu(|X_1| > f^{-1}(a + 1) | |X_1| > f^{-1}(a)) \geq q \left(\frac{p}{q}\right)^{[f^{-1}(a+1)] - [f^{-1}(a)] + 1}. \end{aligned}$$

Choose e.g. $f(x) = (q/p)^{\alpha x}$ so that $f \in \mathbb{L}^\eta(\mu)$ for $\eta < 1/\alpha$. The previous lower bound does not converge to 0 as $a \rightarrow +\infty$.

This situation is presumably typical of the difficulties to satisfy the stable invariance principle, in particular for discrete valued chains.

Acknowledgements. The authors are indebted to two anonymous referees for their valuable comments and criticism on a first version of this paper.

REFERENCES

- [1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto and G. Scheffer, Sur les inégalités de Sobolev logarithmiques. Vol. 10 of *Panoramas et Synthèses*. Société Mathématique de France, Paris (2000).
- [2] D. Bakry, P. Cattiaux and A. Guillin, Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Func. Anal.* **254** (2008) 727–759.
- [3] K. Bartkiewicz, A. Jakubowski, T. Mikosch and O. Wintenberger, Stable limits for sums of dependent infinite variance random variables. *Probab. Theory Relat. Fields* **150** (2011) 337–372.
- [4] B. Basrak, D. Krizmanic and J. Segers, A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.* **40** (2012) 2008–2033.
- [5] P. Cattiaux, A pathwise approach of some classical inequalities. *Potential Analysis* **20** (2004) 361–394.
- [6] P. Cattiaux, D. Chafaï and A. Guillin, Central Limit Theorem for additive functionals of ergodic Markov Diffusions. *ALEA, Lat. Am. J. Probab. Math. Stat.* **9** (2012) 337–382.
- [7] P. Cattiaux and A. Guillin, Deviation bounds for additive functionals of Markov processes. *ESAIM: PS* **12** (2008) 12–29.
- [8] P. Cattiaux and A. Guillin, Trends to equilibrium in total variation distance. *Ann. Inst. Henri Poincaré. Prob. Stat.* **45** (2009) 117–145.
- [9] P. Cattiaux, A. Guillin and C. Roberto, Poincaré inequality and the L^p convergence of semi-groups. *Elec. Commun. Probab.* **15** (2010) 270–280.
- [10] P. Cattiaux, A. Guillin and P.A. Zitt, Poincaré inequalities and hitting times. *Ann. Inst. Henri Poincaré. Prob. Stat.* **49** (2013) 95–118.
- [11] Mu-Fa Chen, Eigenvalues, inequalities, and ergodic theory. *Probab. Appl.* (New York). Springer-Verlag London Ltd., London (2005).
- [12] R.A. Davis, Stable limits for partial sums of dependent random variables. *Ann. Probab.* **11** (1983) 262–269.
- [13] M. Denker and A. Jakubowski, Stable limit theorems for strongly mixing sequences. *Stat. Probab. Lett.* **8** (1989) 477–483.
- [14] A. Jakubowski, Minimal conditions in p-stable limit theorem. *Stochastic Process. Appl.* **44** (1993) 291–327.
- [15] M. Jara, T. Komorowski and S. Olla, Limit theorems for additive functionals of a Markov chain. *Ann. Appl. Probab.* **19** (2009) 2270–2300.
- [16] D. Krizmanic, Functional limit theorems for weakly dependent regularly varying time series. Ph.D. thesis (2010). Available at <http://www.math.uniri.hr/~dkrizmanic/DKthesis.pdf>.
- [17] S.P. Meyn and R.L. Tweedie, Markov chains and stochastic stability. *Commun. Control Eng. Series*. Springer-Verlag London Ltd., London (1993).
- [18] F. Merlevède, M. Peligrad and S. Utev, Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3** (2006) 1–36.
- [19] L. Miclo, On hyperboundedness and spectrum of Markov operators. Preprint, available on hal-00777146 (2013).
- [20] M. Röckner and F.Y. Wang, Weak Poincaré inequalities and L^2 -convergence rates of Markov semi-groups. *J. Funct. Anal.* **185** (2001) 564–603.
- [21] M. Tyran-Kaminska, Convergence to Lévy stable processes under some weak dependence conditions. *Stochastic Process. Appl.* **120** (2010) 1629–1650.
- [22] E. Van Doorn and P. Schrdner, Geometric ergodicity and quasi-stationarity in discrete time Birth-Death processes. *J. Austral. Math. Soc. Ser. B* **37** (1995) 121–144.