J. Math. Anal. Appl. 434 (2016) 52–83

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Limit theorems for Markov chains by the symmetrization method

Christophe Cuny [a,*], Michael Lin [b]

[a] *Laboratoire MAS, Ecole Centrale de Paris, Grande Voie des Vignes, 92295 Chatenay-Malabry cedex, France*

[b] *Department of Mathematics, Ben-Gurion University, Beer-Sheva, Israel*

A R T I C L E I N F O

Article history:
Received 21 April 2015
Available online 4 August 2015
Submitted by U. Stadtmueller

Dedicated to the memory of Mikhail (Misha) Gordin

Keywords:
Forward–backward martingale decomposition
Sector condition
Numerical range and Stolz region
Normal dilation of Markov operators
Almost sure CLT
Poincaré's inequality

A B S T R A C T

Let P be a Markov operator with invariant probability m, ergodic on $L^2(\mathbb{S}, m)$, and let $(W_n)_{n\geq 0}$ be the Markov chain with state space $\mathbb{S}$ and transition probability P on the space of trajectories $(\Omega, \mathbb{P}_m)$, with initial distribution m. Following Wu and Olla we define the symmetrized operator $P_s = (P + P^*)/2$, and analyze the linear manifold $\mathcal{H}_{-1} := \sqrt{I - P_s}L^2(\mathbb{S}, m)$. We obtain for real $f \in \mathcal{H}_{-1}$ an explicit forward–backward martingale decomposition with a coboundary remainder. For such f we also obtain some maximal inequalities for $S_n(f) := \sum_{k=0}^{n} f(W_k)$, related to the law of iterated logarithm. We prove an almost sure central limit theorem for $f \in \mathcal{H}_{-1}$ when P is normal in $L^2(\mathbb{S}, m)$, or when P satisfies the sector condition. We characterize the sector condition by the numerical range of P on the complex $L^2(\mathbb{S}, m)$ being in a sector with vertex at 1. We then show that if P has a real normal dilation which satisfies the sector condition, then $\mathcal{H}_{-1} = \sqrt{I - P}L^2(\mathbb{S}, m)$. We use our approach to prove that P is L^2-uniformly ergodic if and only if it satisfies (the discrete) Poincaré's inequality.

1. Introduction

Let P be a transition probability on a measurable space $(\mathbb{S}, \mathcal{S})$, leaving invariant a probability m on $\mathcal{S}$. We denote also by P the Markov operator induced on $L^2(m) := L^2(\mathbb{S}, m)$, which is known [37, p. 65] to be a contraction of $L^2(m)$, and we denote by P^* its adjoint. We assume that P is ergodic (i.e. $Pf = f \Rightarrow f = cte$).

Let $(W_n)_{n\geq 0}$ be the canonical Markov chain with state space $(\mathbb{S}, \mathcal{S})$ associated with P, defined on the canonical space of trajectories $(\Omega, \mathcal{A}, \mathbb{P}_m) = (\mathbb{S}^{\mathbb{N}}, \mathcal{S}^{\otimes\mathbb{N}}, \mathbb{P}_m)$ with initial distribution m. We denote by θ the shift on Ω, which leaves $\mathbb{P}_m$ invariant.

For $f \in L^2(m)$ we denote $S_n(f) := \sum_{k=0}^{n} f(W_k)$. The purpose of this paper is to obtain limit theorems for $(S_n(f))_{n\geq 0}$ when $f \in L^2(m)$ is centered and belongs to an appropriate subspace.

* Corresponding author.
E-mail addresses: christophe.cuny@ecp.fr (C. Cuny), lin@math.bgu.ac.il (M. Lin).

http://dx.doi.org/10.1016/j.jmaa.2015.07.061

The work of Kipnis and Varadhan [35] on the central limit theorem (CLT) for reversible Markov chains ($P = P^*$) inspired L. Wu [51] and Olla [45] to approach the problem for non-symmetric P by looking at the symmetrized operator $P_s := \frac{1}{2}(P+P^*)$. They introduced a certain linear manifold $\mathcal{H}_{-1}$, which turns out to be precisely $\sqrt{I-P_s}L^2(m)$ (see Section 2). For any $f \in \mathcal{H}_{-1}$, Wu and Olla obtained a decomposition of $S_n(f)$ into a forward and a backward martingale plus a coboundary term. When P satisfies the so-called *sector condition*, they proved that if $f \in \mathcal{H}_{-1}$, then the CLT holds for $(S_n(f))_{n\geq 1}$. Wu also obtained Donsker's invariance principle and, for symmetric operators he proved Strassen's invariance principle in [52].

Our purpose is two-fold. Firstly, we revisit the symmetrization approach of Wu [51] and Olla [45] (used also by Cattiaux, Chafaï and Guillin [9] for continuous time Markov processes), but we define the linear manifold $\mathcal{H}_{-1}$ directly as $\sqrt{I-P_s}L^2(m)$. This allows us to obtain an explicit forward–backward martingale decomposition. Using recent results on martingales and backward martingales (see [10,18] or [15]), we use the symmetrization method to obtain other limit theorems, such as Strassen's functional law of the iterated logarithm or the almost sure central limit theorem, for functions in a certain subspace $\mathbf{B}$ of $\mathcal{H}_{-1}$ (see below).

Secondly, we provide additional examples of operators to which the method applies, by characterizing the sector condition, and by studying conditions under which $\mathcal{H}_{-1} = \sqrt{I-P}L^2(m)$ (which is important, since even for P normal $\mathcal{H}_{-1}$ may be only $(I-P)L^2(m)$ [23]).

The paper is organized as follows. In Section 2 we define $\mathcal{H}_{-1}$ and derive its properties, most of them given in [52] or [45] (sometimes implicitly). We show that $\mathcal{H}_{-1}$ is invariant under P, and define $\mathbf{B}$ as the $\mathcal{H}_{-1}$ closure of $(I-P)\mathcal{H}_{-1}$. In Section 3 we obtain an explicit forward–backward martingale decomposition, and derive some tightness results for $(S_n(f))$ when $f \in \mathcal{H}_{-1}$. In Section 4 we obtain Donsker's invariance principle and Strassen's invariance principle for $(S_n(f))$ when $f \in \mathbf{B}$. In Section 5 we prove the almost sure CLT, when $f \in \mathbf{B}$. In Section 6 we characterize the sector condition (which yields $\mathbf{B} = \mathcal{H}_{-1}$ [45]) in terms of the numerical range of the action of P on the complex $L^2(m)$. In Section 7 we study the relationship between $\mathcal{H}_{-1}$ and $\sqrt{I-P}L^2(m)$. The main tool here is the existence of a real normal dilation of P, such that its complexification has its spectrum in a sector with vertex at 1. This yields the equality $\mathbf{B} = \mathcal{H}_{-1} = \sqrt{I-P}L^2(m)$ and several other characterizations. In particular, if the dilation has its spectrum in a Stolz region, then $f \in \mathcal{H}_{-1}$ if and only if $\sum \|P^n f\|^2 < \infty$. In Section 8 we give several examples in which our results imply the limit theorems for every $f \in \sqrt{I-P}L^2(m)$. In Section 9 we discuss some counter-examples, and show that for any Markov operator, convergence of $\sum_{n\geq 0} \|P^n f\|^2$ implies that $f \in \sqrt{I-P}L^2(m)$, but f need not be in $\mathbf{B}$. In Section 10 we use the results of Sections 2 and 3 to obtain the equivalence of (the discrete) Poincaré's inequality and uniform ergodicity of P.

We would like to mention that most of our results may be generalized to the case where f takes its values in a separable Hilbert space, see Jiang–Wu [34], where the symmetrization method is applied in that setting, with applications. Now, since few auxilliary Hilbert spaces are necessary in our analysis, we prefered not to consider that generality here. We shall only mention from time to time the potential extensions.

2. Symmetrization and the space $\mathcal{H}_{-1}$

In this section we study the symmetrization procedure, used (independently) by L. Wu [52] and (in continuous time only) by Olla [45] (see also [46,34] or [36]).

Recall that m is an invariant probability for P and that P is ergodic.

We shall denote by $\|\cdot\|_0$ the norm in $L^2(m)$.

Define $P_s = (P + P^*)/2$, which is symmetric (self-adjoint) on $L^2(m)$. The P-invariant probability m is invariant for P_s. By the next lemma P_s is also ergodic.

Lemma 2.1. *The operators P, P^* and P_s have the same fixed points. Consequently P_s is ergodic if and only if P is.*

Proof. By a result of Riesz, the L^2-contractions P and P^* have the same fixed points, so $Pf = f \in L^2(m) \Longrightarrow P_s f = f$. Conversely, if $P_s f = f \in L^2(m)$, then by uniform convexity, since $\max\{\|Pf\|_0, \|P^*f\|_0\} \le \|f\|_0$, we see that $Pf = P^*f = f$. $\square$

Denote $L_0^2(m) = L_0^2(\mathbb{S}, m) := \{f \in L^2(\mathbb{S}, m) : \int f\, dm = 0\}$. By ergodicity of P and P_s and the mean ergodic theorem, we have the (orthogonal) ergodic decomposition

$$L^2(m) = \{constants\} \oplus \overline{(I-P)L^2(m)} = \{constants\} \oplus \overline{(I-P_s)L^2(m)}, \tag{1}$$

which yields

$$L_0^2(m) = \overline{(I-P)L^2(m)} = \overline{(I-P)L_0^2(m)} = \overline{(I-P_s)L^2(m)} = \overline{(I-P_s)L_0^2(m)}.$$

Following Derriennic and Lin [21], for a contraction T on a Banach space we define the operator $\sqrt{I-T}$ by

$$\sqrt{I-T} := I - \sum_{n\ge 1} \alpha_n T^n = \sum_{n\ge 1} \left((I-T)\alpha_n \sum_{k=0}^{n-1} T^k\right), \tag{2}$$

where $\sqrt{1-t} = 1 - \sum_{n\ge 1} \alpha_n t^n$, $0 \le |t| \le 1$, with $\alpha_n > 0$ and $\sum_{n\ge 1} \alpha_n = 1$.

Define also a sequence $(\beta_n)_{n\ge 0}$, by the power series expansion

$$\frac{1}{\sqrt{1-t}} = \sum_{n\ge 0} \beta_n t^n \qquad 0 \le |t| < 1\,. \tag{3}$$

Then $\beta_n \ge 0$ for $n \ge 0$ [21].

The ergodic decomposition shows that $\sqrt{I-P_s}L^2(m) = \sqrt{I-P_s}L_0^2(m)$. We then define $\mathcal{H}_{-1} = \sqrt{I-P_s}L_0^2(\mathbb{S}, m)$. By [21], $\overline{\mathcal{H}_{-1}} = \overline{\sqrt{I-P_s}L^2(m)} = L_0^2(m)$. Let us recall the following description of $\mathcal{H}_{-1}$ from [21], see their Theorem 2.11 and their Corollary 2.12.

Lemma 2.2. *The following are equivalent:*

(i) $f \in \mathcal{H}_{-1}$.
(ii) $\sum_{n\ge 0} \beta_n P_s^n f$ *converges in* $L^2(\mathbb{S}, m)$.
(iii) $\liminf_{N\to\infty} \|\sum_{n=0}^{N} \beta_n P_s^n f\|_0 < \infty$.

If any of the above hold, then $g := \sum_{n\ge 0} \beta_n P_s^n f$ *is the unique (in* $L_0^2(\mathbb{S}, m)$*) solution of the equation* $f = \sqrt{I-P_s}g$, *and*

$$g = \sum_{n\ge 0} \left(\beta_n P_s^n \sqrt{I-P_s}g\right) = \sum_{n\ge 1} \left(\alpha_n \sum_{k=0}^{n-1} P_s^k \sqrt{I-P_s}g\right). \tag{4}$$

By the uniqueness stated in Lemma 2.2, we may define an inner-product on $\mathcal{H}_{-1}$ as follows.

Definition. Let $f_1, f_2 \in \mathcal{H}_{-1}$, and g_1, g_2 be the unique elements of $L_0^2(\mathbb{S}, m)$, such that $f_i = \sqrt{I-P_s}g_i$. Then, define

$$\langle f_1, f_2\rangle_{-1} := \langle g_1, g_2\rangle_0,$$

where $\langle\cdot,\cdot\rangle_0$ stands for the inner-product on $L^2(\mathbb{S}, m)$.

It is not difficult to see that $\langle\cdot,\cdot\rangle_{-1}$ is positive-definite and that

$$\|f\|_0 = \|\sqrt{I-P_s}g\|_0 \le 2\|g\|_0 = 2\|f\|_{-1}\,. \tag{5}$$

If $\{f_n = \sqrt{I-P_s}g_n\}$ is Cauchy in $\mathcal{H}_{-1}$, then $\{g_n\}$ is Cauchy in $L^2_0(m)$, so $\|g_n - g\|_0 \to 0$ implies $\|f_n - \sqrt{I-P_s}g\|_{-1} \to 0$, so $(\mathcal{H}_{-1}, \|\cdot\|_{-1})$ is a (real) Hilbert space.

Remark. Unlike [52] and [45], we have defined $\mathcal{H}_{-1}$ directly, without reference to the space $\mathcal{H}_1$ obtained from the Dirichlet form $\langle f,g\rangle_1 := \langle (I-P_s)f,g\rangle_0$. The fact that both definitions coïncide follows from Lemma 2.3 below.

We will need the following lemma from Wu [51]. We give the proofs, for the sake of completeness and because some arguments are needed in the sequel.

Lemma 2.3.

(i) *For every $f \in \mathcal{H}_{-1}$, we have*

$$\|f\|_{-1} = \inf\{C \ge 0 \ :\ |\langle f,h\rangle_0|^2 \le C^2\langle (I-P)h,h\rangle_0,\ \forall h \in L^2_0(\mathbb{S},m)\}. \tag{6}$$

Moreover, if for $f \in L^2(\mathbb{S},m)$ the right-hand-side of (6) is finite, then $f \in \mathcal{H}_{-1}$, and (6) holds.

(ii) *$I-P$ and $I-P^*$ are bounded operators from $L^2(\mathbb{S},m)$ to $\mathcal{H}_{-1}$. More precisely, for every $f \in L^2(\mathbb{S},m)$, taking $Q \in \{P,P^*\}$,*

$$\|(I-Q)f\|_{-1} \le \sqrt{2}\|f\|_0.$$

Proof. We first note (remember that we are on a real Hilbert space) that for every $h \in L^2$,

$$\langle h,(I-P)h\rangle_0 = \langle h,(I-P^*)h\rangle_0 = \langle h,(I-P_s)h\rangle_0 = \|\sqrt{I-P_s}h\|_0^2\,. \tag{7}$$

Let us prove (i). Assume that $f = \sqrt{I-P_s}g$. Then, using Cauchy–Schwarz, for every $h \in L^2$,

$$|\langle f,h\rangle_0|^2 = |\langle g,\sqrt{I-P_s}h\rangle_0|^2 \le \|g\|_0^2\langle (I-P_s)h,h\rangle_0 = \|f\|_{-1}^2\langle (I-P)h,h\rangle\,, \tag{8}$$

which proves that $\|f\|_{-1}$ is not smaller than the right-hand-side of (6). To prove the equality, we exhibit a sequence $(g_n)_n \subset L^2(\mathbb{S},m)$, such that $\langle f,g_n\rangle_0 \to \|g\|_0^2$ and $\langle (I-P_s)g_n,g_n\rangle_0 \to \|g\|_0^2$. By Theorem 2.7 of [21] (see (4)), $g_n = \sum_{k=0}^n \beta_k P_s^k g$ satisfies those requirements.

Let $f \in L^2(\mathbb{S},m)$ be such that the right-hand-side of (6) is finite. Then, using (7) and [21, Theorem 2.13] we see that $f \in \mathcal{H}_{-1}$.

Proof of (ii). For every $h \in L^2(\mathbb{S},m)$, and for $Q \in \{P,P^*\}$, we have, by (7),

$$2\langle h,(I-P_s)h\rangle_0 - \|(I-Q)h\|_0^2 = 2\langle h,(I-Q)h\rangle_0 - \|(I-Q)h\|_0^2 = \|h\|_0^2 - \|Qh\|_0^2\,. \tag{9}$$

Let $f \in L^2(m)$. Then for $h \in L^2(m)$ we have, by positivity of (9) and (7),

$$|\langle (I-Q)f,h\rangle_0|^2 = |\langle f,(I-Q^*)h\rangle_0|^2 \le \|f\|_0^2\|(I-Q^*)h\|_0^2 \le 2\|f\|_0^2\langle h,(I-P)h\rangle_0.$$

By (i), $(I-Q)f \in \mathcal{H}_{-1}$ with $\|(I-Q)f\|_{-1} \le \sqrt{2}\|f\|_0$. $\square$

By the definitions, $\mathcal{H}_{-1}$ is invariant under P_s, and P_s is a contraction of $\mathcal{H}_{-1}$. Since P_s has no fixed points in $L_0^2(m) \supset \mathcal{H}_{-1}$, we have that $(I-P_s)\mathcal{H}_{-1}$ is $\|\cdot\|_{-1}$-dense in $\mathcal{H}_{-1}$.

Lemma 2.3(ii) yields that $\mathcal{H}_{-1}$ is invariant under P and P^*, with $\|I-Q\|_{-1} \le 2\sqrt{2}$ for $Q \in \{P, P^*\}$ and $(I-P)\mathcal{H}_{-1} \subset (I-P)L^2(\mathbb{S}, m) \subset \mathcal{H}_{-1}$.

Lemma 2.4. *We have*

$$\overline{(I-P)L^2(\mathbb{S},m)}^{\mathcal{H}_{-1}} = \overline{(I-P)\mathcal{H}_{-1}}^{\mathcal{H}_{-1}} \tag{10}$$

where $\overline{A}^{\mathcal{H}_{-1}}$ denotes the closure of $A \subset \mathcal{H}_{-1}$ with respect to the $\|\cdot\|_{-1}$-norm.

When P is normal (i.e. $PP^* = P^*P$) we have $\mathcal{H}_{-1} = \overline{(I-P)\mathcal{H}_{-1}}^{\mathcal{H}_{-1}}$.

Proof. One inclusion is clear. Let us show the converse inclusion. We have $(I-P)L^2(\mathbb{S},m) = (I-P)\overline{(I-P)L^2(\mathbb{S},m)}^{L^2(\mathbb{S},m)}$. Hence, if $f \in (I-P)L^2(\mathbb{S},m)$, there exists $g \in \overline{(I-P)L^2(\mathbb{S},m)}$, $(g_n) \subset L^2(\mathbb{S},m)$ such that $f = (I-P)g$ and $(I-P)g_n \to g$ in $L^2(\mathbb{S},m)$. By item (ii) of Lemma 2.3, $(I-P)(I-P)g_n \to f$ in $\mathcal{H}_{-1}$ and $(I-P)g_n \in \mathcal{H}_{-1}$, which proves (10).

When P is a normal operator, then P commutes with P_s and therefore with $\sqrt{I-P_s}$. Thus, in that case, P is a contraction of $\mathcal{H}_{-1}$, with no fixed points. Hence, by von Neumann's mean ergodic theorem (see e.g. [37, p. 1]) $\mathcal{H}_{-1} = \overline{(I-P)\mathcal{H}_{-1}}^{\mathcal{H}_{-1}}$. $\square$

In view of (10), we define $\mathbf{B} := \overline{(I-P)\mathcal{H}_{-1}}^{\mathcal{H}_{-1}}$. The relevance of the space $\mathbf{B}$ will be made clear in Section 4.

Note that P has no fixed points in $\mathcal{H}_{-1} \subset L_0^2(m)$. If P is power-bounded on $\mathcal{H}_{-1}$, i.e. $\sup_{n\ge1} \|P^n\|_{-1} < \infty$, then by the mean ergodic theorem (see e.g. [37, p. 73] and Lemma 2.4), $\mathbf{B} = \mathcal{H}_{-1}$. In general, if P is not normal, there is no reason why P should be power-bounded, but we know the following (see Corollary 3.3 below for an improvement).

Lemma 2.5. *For every $f \in \mathcal{H}_{-1}$ we have $\frac{\|P^n f\|_{-1}}{n} \to 0$. Hence $\sup_n \frac{\|P^n\|_{-1}}{n} < \infty$.*

Proof. We use Lemma 2.3(ii) and the mean ergodic theorem in $L_2^0(m)$ to obtain

$$\frac{\|P^n f\|_{-1}}{n} \le \frac{\|f\|_{-1}}{n} + \Big\|\frac{1}{n}\sum_{k=0}^{n-1}(I-P)P^k f\Big\|_{-1} \le \sqrt{2}\Big\|\frac{1}{n}\sum_{k=0}^{n-1}P^k f\Big\|_0 + \frac{\|f\|_{-1}}{n} \to 0.$$

The above inequalities, together with (5), yield $\frac{1}{n}\|P^n f\|_{-1} \le (2\sqrt{2}+1)\|f\|_{-1}$.

Similarly, $\|P^n(I-P)f\|_{-1} \le 2\sqrt{2}\|f\|_{-1}$, so $\|P^n(I-P)\|_{-1} \le 2\sqrt{2}$. $\square$

Definition 1. We say that an operator T on a Banach space X, such that $\limsup_{n\to\infty} \|T^n\|^{1/n} \le 1$, is *Abel bounded* if

$$\sup_{0\le\lambda<1} (1-\lambda)\|\sum_{n\ge0} \lambda^n T^n\| < \infty\,. \tag{11}$$

The *Abel means* $A_\lambda := (1-\lambda)\sum_{n\ge0} \lambda^n T^n$ satisfy

$$f = \lambda(I-T)(\sum_{n\ge0} \lambda^n T^n f) + A_\lambda f\,. \tag{12}$$

Let us recall the following well-known lemma see for instance [27, Theorem 2.1] (see also [33, Sections 18.4–18.6] for the continuous time case).

Lemma 2.6. *Let X be a reflexive Banach space. Let T be Abel bounded on X. Then T has the ergodic decomposition*

$$X = \{x \in X,\, Tx = x\} \oplus \overline{(I-T)X}\,.$$

Remark. Actually, for any bounded T on a reflexive Banach space X, if $f \in X$ satisfies $\liminf_{\lambda\to 1^-}(1-\lambda)\|\sum_{n\ge 0}\lambda^n T^n f\| < \infty$, then $f \in \{x \in X,\, Tx = x\} \oplus \overline{(I-T)X}$.

The following result may be found in Olla [45] in the continuous time case; see also Jiang–Wu [34] (Lemma 4.3). We include the proof for the sake of completeness.

Proposition 2.7. *(See Olla [45], sector condition.) Assume that there exists $K > 0$ for which*

$$|\langle (I-P)g, h\rangle_0|^2 \le K\langle (I-P)g, g\rangle_0 \langle (I-P)h, h\rangle_0 \qquad \forall g, h \in L^2_0(\mathbb{S}, m)\,. \tag{13}$$

Then P is Abel-bounded on $\mathcal{H}_{-1}$, hence $\mathbf{B} = \mathcal{H}_{-1}$.

Proof. By Lemma 2.5, $\sup_n \frac{1}{n}\|P^n\|_{-1} < \infty$. Hence, for every $0 \le \lambda < 1$, the Abel average operator $A_\lambda := (1-\lambda)\sum_{n\ge 0}\lambda^n P^n$ is well defined on $\mathcal{H}_{-1}$.

Let $f \in \mathcal{H}_{-1}$. We will use item (i) of Lemma 2.3 to bound $\|A_\lambda f\|_{-1}$. Let $g \in L^2_0(m)$. By (12) and (13), we have, writing $f_\lambda := \sum_{n\ge 0}\lambda^n P^n f$.

$$\begin{aligned} |\langle A_\lambda f, g\rangle_0| &\le |\langle f, g\rangle_0| + |\langle (I-P)f_\lambda, g\rangle_0| \\ &\le \|f\|_{-1}|\langle (I-P)g, g\rangle_0|^{1/2} + K|\langle (I-P)f_\lambda, f_\lambda\rangle_0|^{1/2}|\langle (I-P)g, g\rangle_0|^{1/2}\,. \end{aligned} \tag{14}$$

On the other hand, using (12) again and the fact that $\langle f_\lambda, A_\lambda f\rangle_0 \ge 0$, we infer that

$$0 \le \lambda\langle (I-P)f_\lambda, f_\lambda\rangle_0 \le \langle f, f_\lambda\rangle_0 \le \|f\|_{-1}\langle (I-P)f_\lambda, f_\lambda\rangle_0^{1/2}$$

Hence,

$$\lambda\langle (I-P)f_\lambda, f_\lambda\rangle_0^{1/2} \le \|f\|_{-1}\,.$$

To conclude we use the latter bound combined with (14). □

The sector condition (13) means that the operator P is "close" to being symmetric. This heuristic statement will be made more precise in Section 6.

3. The forward–backward martingale decomposition

For $f \in L^2(\mathbb{S}, m)$, write $S_n = f(W_1) + \cdots + f(W_n)$.

Our goal here is to show that whenever $f \in \mathcal{H}_{-1}$, $f(W_1)$ may be written as the sum of a martingale increment, a reverse martingale increment, and a coboundary (for the shift). Since martingales and sums of reverse martingale increments enjoy good properties, such as maximal inequalities or the law of the iterated logarithm, these properties shall "transfer" to (S_n), allowing us to obtain limit theorems for $f \in \mathbf{B}$ or even $f \in \mathcal{H}_{-1}$.

The novelty here is, firstly, that we obtain an explicit form of the (not necessarily unique) forward–backward martingale decomposition. This explicit form, inspired by Derriennic–Lin [22], is not really needed in the sequel, but it might be useful for other problems. Secondly, we obtain a maximal inequality related the law of the iterated logarithm.

Let $(\mathcal{F}_n)_{n\geq 0}$ be the natural filtration, i.e. $\mathcal{F}_n = \sigma\{W_0,\ldots,W_n\}$, and denote the tail σ-fields by $\mathcal{G}_n = \sigma\{W_n, W_{n+1},\ldots\}$.

Proposition 3.1. *Let $f \in \mathcal{H}_{-1}$ and denote $g_n = \sum_{k=0}^{n-1} P_s^k g$, where g is the unique centered $g \in L^2(m)$ such that $f = \sqrt{I-P_s}\,g$. Then the series*

$$\begin{aligned}\mathcal{D}^+ f &:= \frac{1}{2}\sum_{n\geq 1}\alpha_n(g_n(W_1) - Pg_n(W_0)),\\ \mathcal{D}^- f &:= \frac{1}{2}\sum_{n\geq 1}\alpha_n(g_n(W_0) - P^*g_n(W_1)),\\ \mathcal{R} f &:= \frac{1}{2}\sum_{n\geq 1}\alpha_n(I-P)g_n(W_1))\end{aligned}$$

converge in $L^2(\Omega,\mathcal{F},\mathbb{P}_m)$ and

$$\max(\|\mathcal{D}^+ f\|_{L^2(\mathbb{P}_m)}, \|\mathcal{D}^- f\|_{L^2(\mathbb{P}_m)}, \|\mathcal{R} f\|_{L^2(\mathbb{P}_m)}) \leq (\sqrt{2}/2)\|f\|_{-1}\,.$$

In particular, $\mathcal{D}^+$ is a continuous operator from $\mathcal{H}_{-1}$ to $L^2(\Omega,\mathcal{F}_1,\mathbb{P}_m)\ominus L^2(\Omega,\mathcal{F}_0,\mathbb{P}_m)$, $\mathcal{D}^-$ is a continuous operator from $\mathcal{H}_{-1}$ to $L^2(\Omega,\mathcal{G}_0,\mathbb{P}_m)\ominus L^2(\Omega,\mathcal{G}_1,\mathbb{P}_m)$ and $\mathcal{R}$ is a continuous operator from $\mathcal{H}_{-1}$ to $L^2(\Omega,\mathcal{F}_1,\mathbb{P}_m)$. Moreover,

$$f(W_1) = \mathcal{D}^+ f + \mathcal{D}^- f + \mathcal{R} f - (\mathcal{R} f)\circ\theta^{-1}. \tag{15}$$

Proof. It follows from (9), that for every $h \in L^2(\mathbb{S}, m)$ and for $Q \in \{P, P^*\}$, we have

$$\|h\|_0^2 - \|Qh\|_0^2 \leq 2\|\sqrt{I-P_s}\,h\|_0^2\,; \tag{16}$$

$$\|(I-Q)h\|_0^2 \leq 2\|\sqrt{I-P_s}\,h\|_0^2\,. \tag{17}$$

Let $q > p \geq 1$. We have, using (16),

$$\|\sum_{n=p}^{q}\alpha_n(g_n(W_1) - Pg_n(W_0))\|_{L^2(\mathbb{P}_m)}^2 \tag{18}$$

$$= \|\sum_{n=p}^{q}\alpha_n g_n\|_0^2 - \|P\sum_{n=p}^{q}\alpha_n g_n\|_0^2 \leq 2\|\sum_{n=p}^{q}\alpha_n\sqrt{I-P_s}\,g_n\|_0^2 \xrightarrow[q,p\to\infty]{} 0\,, \tag{19}$$

by Lemma 2.2. This proves the convergence of $\mathcal{D}^+ f$. Now, taking $p = 1$ and letting $q \to \infty$ in (18) and using (4) we see that $\|\mathcal{D}^+ f\|_0^2 \leq 2\|g\|_0^2 = 2\|f\|_{-1}^2$.

The results about $\mathcal{D}^- f$, may be proved similarly. The proof for $\mathcal{R}$ follows from (17).

It remains to prove (15). Let $n \geq 1$. We have

$$\begin{aligned}(I-P_s)\alpha_n g_n(W_1) &= \frac{1}{2}\alpha_n(g_n(W_1) - Pg_n(W_0))\\ &\quad + \frac{1}{2}\alpha_n(g_n(W_0) - P^*g_n(W_1)) + \frac{1}{2}\alpha_n(I-P)g_n(W_1)) - \frac{1}{2}\alpha_n(I-P)g_n(W_0))\,.\end{aligned}$$

Summing from $n = 1$ to p, and letting $p \to \infty$, (15) follows. □

Write

$$M_n^+ := \sum_{k=0}^{n-1} \mathcal{D}^+ f \circ \theta^k \qquad \text{and} \qquad M_n^- := \sum_{k=0}^{n-1} \mathcal{D}^- f \circ \theta^k \,. \tag{20}$$

Theorem 3.2. *Let $f \in \mathcal{H}_{-1}$. We have*

$$\| \max_{1\le k\le n} |S_k| \|_{L^2(\mathbb{P}_m)} \le 2\sqrt{2}\sqrt{n}\|f\|_{-1}\,; \tag{21}$$

$$\limsup_{n\to+\infty} \frac{|S_n|}{\sqrt{2n\log\log n}} \le \sqrt{2}\|f\|_{-1} \qquad \mathbb{P}_m\text{-}a.s. \tag{22}$$

Moreover, for every $1 \le p < 2$ there exists $C_p > 0$ such that

$$\| \sup_{n\ge 3} \frac{|S_n|}{\sqrt{n\log\log n}} \|_p \le C_p \|f\|_{-1}\,. \tag{23}$$

Remark. The proof makes use of the corresponding results for martingale or reverse martingale differences. Those results hold in the context of Hilbert-valued (reverse) martingales, replacing the absolute value with the Hilbertian norm. The construction of $\mathcal{H}_{-1}$ as well as the forward–backward decomposition hold in that context. Hence, Theorem 3.2 also holds in that context.

Proof. We have

$$\begin{aligned} \| \max_{1\le k\le n} |S_k| \|_{L^2(\mathbb{P}_m)} \le \| \max_{1\le k\le n} |M_k^+| \|_{L^2(\mathbb{P}_m)} + \| \max_{1\le k\le n} |M_k^-| \|_{L^2(\mathbb{P}_m)} \\ + \|\mathcal{R}f\|_{L^2(\mathbb{P}_m)} + \| \max_{0\le k\le n-1} |\mathcal{R}\circ\theta^k| \|_{L^2(\mathbb{P}_m)}\,. \end{aligned}$$

Notice that for every $n \ge 1$, $\Big(M_n^- - M_{n-k}^-\Big)_{0\le k\le n-1}$ is a martingale. Hence, using Doob maximal inequality twice, we obtain

$$\| \max_{1\le k\le n} |M_k^+| \|_{L^2(\mathbb{P}_m)} \le 2\|M_n^+\|_{L^2(\mathbb{P}_m)} = 2\sqrt{n}\|\mathcal{D}^+ f\|_0 \le 2\sqrt{2}\|f\|_{-1}\,;$$

$$\| \max_{1\le k\le n} |M_k^{(-)}| \|_{L^2(\mathbb{P}_m)} \le 3\|M_n^-\|_{L^2(\mathbb{P}_m)} = 3\sqrt{n}\|\mathcal{D}^- f\|_0 \le 3\sqrt{2}\|f\|_{-1}\,.$$

On the other hand

$$\| \max_{0\le k\le n-1} |\mathcal{R}f\circ\theta^k| \|^2_{L^2(\mathbb{P}_m)} \le \sum_{k=0}^{n-1} \|\mathcal{R}f\circ\theta^k\|^2_{L^2(\mathbb{P}_m)} \le \frac{n}{2}\|f\|^2_{-1}\,.$$

Combining those estimates, we derive (21).

Let us prove (22). We now apply the LIL for sums of stationary and ergodic differences of martingales or of reverse martingales. For the latter, we refer to [18], see also [52] where a bounded LIL is obtained for differences of reverse martingales. We have

$$\limsup_{n\to+\infty} \frac{|M_n^-|}{\sqrt{2n\log\log n}} = \|\mathcal{D}^- f\|_{L^2(\mathbb{P}_m)} \le (\sqrt{2}/2)\|f\|_{-1} \qquad \mathbb{P}_m\text{-a.s.}\,,$$

$$\limsup_{n\to+\infty} \frac{|M_n^+|}{\sqrt{2n\log\log n}} = \|\mathcal{D}^+ f\|_{L^2(\mathbb{P}_m)} \le (\sqrt{2}/2)\|f\|_{-1} \qquad \mathbb{P}_m\text{-a.s.}\,,$$

and $\mathcal{R}f\circ\theta^n/\sqrt{n} \to 0$ $\mathbb{P}_m$-a.s., by the Borel–Cantelli lemma. This finishes the proof of (22).

It remains to prove (23). Firstly, (23) holds with M_n^+ (or M_n^-) in place of S_n, by Theorem 2.3 of [16]. To finish the proof it suffices to prove that $\| \sup_{n\geq 1} \frac{|\mathcal{R}f\circ\theta^n|}{n^{1/2}}\|_p \leq C\|\mathcal{R}f\|_0$. But we have

$$\| \sup_{n\geq 1} \frac{|\mathcal{R}f\circ\theta^n|}{n^{1/2}}\|_p^p \leq \|\mathcal{R}f\|_0^p + \sum_{n\geq 0} \frac{\mathbb{E}(|\mathcal{R}f\circ\theta^n|^p \mathbf{1}_{\{|\mathcal{R}f\circ\theta^n|\geq n^{1/2}\|\mathcal{R}f\|_0\}})}{n^{p/2}}$$
$$\leq \|\mathcal{R}f\|_0^p + \mathbb{E}\Big(|\mathcal{R}f|^p \sum_{1\leq n\leq |\mathcal{R}f|^2/\|\mathcal{R}f\|_0^2} \frac{1}{n^{p/2}}\Big) \leq C\|\mathcal{R}f\|_0^p,$$

and the result follows. □

We easily derive the following strengthening of Lemma 2.5.

Corollary 3.3. *Let P be a Markov operator on $L^2(m)$. For every $n\geq 1$, we have $\|P^n\|_{-1}\leq 1+4\sqrt{n}$.*

Proof. Let $f\in\mathcal{H}_{-1}$. By (21) of Theorem 3.2,

$$\|f+\ldots+P^{n-1}f\|_0 = \|\mathbb{E}(S_n|\mathcal{F}_0)\|_{L^2(\mathbb{P}_m)} \leq \|S_n\|_{L^2(\mathbb{P}_m)} \leq 2\sqrt{2}\sqrt{n}\|f\|_{-1}\,.$$

The end of the proof is like the proof of Lemma 2.5, with $\sqrt{n}$ instead of n in the denominator. □

4. Martingale approximation in $\overline{(I-P)L_0^2(\mathbb{S},m)}^{\mathcal{H}_{-1}}$

In this section, we construct a martingale approximation M_n of S_nf, for $f\in\mathbf{B} := \overline{(I-P)L_0^2(\mathbb{S},m)}^{\mathcal{H}_{-1}}$, and show that the weak invariance principle and the functional LIL may be derived from the respective properties of M_n thanks to Theorem 3.2.

Recall that by Lemma 2.3(ii), $(I-P)L^2(\mathbb{S},m)\subset\mathcal{H}_{-1}$, and by Lemma 2.4,

$$\mathbf{B} := \overline{(I-P)L^2(\mathbb{S},m)}^{\mathcal{H}_{-1}} = \overline{(I-P)\mathcal{H}_{-1}}^{\mathcal{H}_{-1}}\,.$$

For every $f\in L^2(\mathbb{S},m)$, define

$$\mathcal{D}((I-P)f) = f(W_1) - Pf(W_0)\,.$$

Note that $\mathcal{D}$ is well-defined since, if $(I-P)f_1 = (I-P)f_2$, then $f_1(W_1)-Pf_1(W_0) = f_2(W_1)-Pf_2(W_0)$ $\mathbb{P}_m$-a.s. (see (24) below).

Now, writing $(I-P)f = \sqrt{I-P_s}g$ (which is possible by item (ii) of Lemma 2.3), we have

$$\|f(W_1)-Pf(W_0)\|_{L^2(\mathbb{P}_m)}^2 = \|f\|_0^2 - \|Pf\|_0^2 \leq 2\langle f,(I-P)f\rangle_0; \tag{24}$$

and

$$\langle f,(I-P)f\rangle_0 = \langle f,\sqrt{I-P_s}g\rangle_0$$
$$\leq \|\sqrt{I-P_s}f\|_0\|g\|_0 = \langle f,(I-P)f\rangle_0^{1/2}\|g\|_0.$$

Hence $\langle f,(I-P)f\rangle_0 \leq \|g\|_0^2 = \|(I-P)f\|_{-1}^2$ and

$$\|\mathcal{D}((I-Pf))\|_{L^2(\mathbb{P})} \leq 2\|(I-P)f\|_{-1}\,.$$

Hence $\mathcal{D}$ defines a linear operator from $(I-P)L^2(\mathbb{S},m) \subset \mathcal{H}_{-1}$ to $L^2(\Omega,\mathcal{F}_1,\mathbb{P}_m) \ominus L^2(\Omega,\mathcal{F}_0,\mathbb{P}_m)$ that may be extended continuously to $\mathbf{B}$. Then

$$\|\mathcal{D}f\|_{L^2(\mathbb{P}_m)} \le 2\|f\|_{-1} \qquad \forall f \in \mathbf{B}\,. \tag{25}$$

Writing $M_n = M_n(f) = \mathcal{D}f + (\mathcal{D}f)\circ\theta + \cdots (\mathcal{D}f)\circ\theta^{n-1}$, for $f \in \mathbf{B}$, (M_n) is a martingale.

We obtain

Theorem 4.1. *Let $f \in \mathbf{B}$. We have*

$$\|\sup_{1\le k\le n} |S_k - M_k|\|_{L^2(\mathbb{P}_m)} = o(\sqrt{n})\,; \tag{26}$$

$$|S_n - M_n| = o(\sqrt{n\log\log n}) \qquad \mathbb{P}_m\text{-}a.s. \tag{27}$$

In particular, (S_n) satisfies the Donsker invariance principle (WIP) and Strassen's invariance principle (ASIP).

As we shall see the proof is based on Theorem 3.2 and Banach principles. In particular, (26) and (27) hold for f taking values in a Hilbert space, and the invariance principles hold as well in that setting (for Strassen's invariance principle, this follows from [16]). The WIP was obtained in [45] and [51] in the real-valued case and in [34] in the Hilbert-valued case. The ASIP was obtained in [52] for symmetric Markov operators, in the real-valued case.

Proof. We make use of Banach principles. To prove (26), we need a very slight modification of the usual Banach principle (see Theorem 7.1 p. 63 of [37]), while to prove (27), we need an almost sure version of it (see Theorem 7.2, p. 64 of [37]).

Let us prove (26). Define the maximal operators $\mathcal{M}_n$ on $\mathcal{H}_{-1}$ by

$$\mathcal{M}_n f = \|\sup_{1\le k\le n} |S_k(f) - M_k(f)|\|_2/\sqrt{n}.$$

By maximal operator we mean that $\mathcal{M}_n$ takes values in $\mathbb{R}^+$ or in a set of non-negative measurable functions, and that it is positively homogeneous and subadditive. By Theorem 3.2, those maximal operators are uniformly bounded on $\mathcal{H}_{-1}$, hence, by (a version of) the Banach principle, the set $E = \{f \in \mathcal{H}_{-1} : \lim_n \mathcal{M}_n f = 0\}$ is closed in $\mathcal{H}_{-1}$, for $\|\cdot\|_{-1}$. But one easily sees that $(I-P)L^2(\mathbb{S},m) \subset E$, which is the classical Gordin–Lifshitz result [30] (note that $S_n((I-P)g) - M_n((I-P)g) = Pg(W_0) - Pg(W_n)$).

Let us prove (27). Define maximal operators $\mathcal{M}_n$ on $\mathbf{B}$ by $\mathcal{M}_n f = |S_n(f) - M_n(f)|/\sqrt{n\log\log n}$. By Theorem 3.2, for every $f \in \mathcal{H}_{-1}$,

$$\sup_{n\ge 1} \mathcal{M}_n f < \infty \qquad \mathbb{P}_m\text{-a.s.}$$

Hence, by the Banach principle the set of $f \in \mathbf{B}$ for which (27) holds, is closed in $\mathcal{H}_{-1}$. But the result is clear on $(I-P)L^2(\mathbb{S},m)$. $\square$

5. The almost sure central limit theorem

We now give an almost sure central limit theorem for $f \in \mathbf{B}$. As previously, the proof will make use of martingales and reverse martingales. The almost sure central limit theorem in the setting of martingales may be found in Lifshits [38] and for reverse martingales in Chazottes–Gouëzel [10].

Our proof is longer than that of Theorem 4.1 by lack of maximal operators. So in some sense we redo the proof of a Banach principle here. Note that (27) is not enough to deduce the almost sure central limit theorem for S_n from the one for M_n (the rate $o(1/\sqrt{n})$ would be needed in (27)).

Theorem 5.1. *Let $f \in \mathbf{B}$. Then, for $\mathbb{P}_m$-a.e. $\omega \in \Omega$, the sequence of distributions $(\frac{1}{\log n}\sum_{k=1}^n \frac{\delta_{S_k(\omega)/\sqrt{k}}}{k})$ converges weakly to the normal distribution $\mathcal{N}(0, \|\mathcal{D}f\|_{L^2(\mathbb{P}_m)})$.*

Proof. For $\gamma \geq 0$, denote by ν_γ the probability measure associated with $\mathcal{N}(0,\gamma)$, where ν_0 stands for the Dirac measure at 0. It suffices to prove that for every $x, y \in \mathbb{Q}$, $x < y$, $(\frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{[x,y]}(S_k/\sqrt{k})}{k})$ converges $\mathbb{P}_m$-a.s. to $\nu_\sigma([x,y])$ where $\sigma = \|\mathcal{D}f\|_{L^2(\mathbb{P}_m)}$. Let $0 < \varepsilon < (y-x)/2$. We have

$$\frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{[x,y]}(S_k/\sqrt{k})}{k} \leq \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{[x-\varepsilon,y+\varepsilon]}(M_k/\sqrt{k})}{k} + \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon,+\infty[}(|S_k - M_k|/\sqrt{k})}{k},$$

and

$$\frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{[x,y]}(S_k/\sqrt{k})}{k} \geq \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{[x+\varepsilon,y-\varepsilon]}(M_k/\sqrt{k})}{k} - \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon,+\infty[}(|S_k - M_k|/\sqrt{k})}{k}.$$

Applying the almost sure central limit theorem for martingales (see e.g. [38]), we obtain that

$$\begin{aligned}&\limsup_n \left|\frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{[x,y]}(S_k/\sqrt{k})}{k} - \nu_\sigma([x,y])\right| \\ &\quad \leq \nu([y-\varepsilon,y+\varepsilon]) + \nu([x-\varepsilon,x+\varepsilon]) + \limsup_n \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon,+\infty[}(|S_k - M_k|/\sqrt{k})}{k}.\end{aligned}$$

Let $\eta > 0$ and $g \in (I-P)L^2(\mathbb{S}, m)$ be such that $\|f-g\|_{-1} < \eta$. We have

$$\begin{aligned}\frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon,+\infty[}(|S_k - M_k|/\sqrt{k})}{k} &\leq \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon/3,+\infty[}(|S_k(g) - M_k(g)|/\sqrt{k})}{k} \\ &+ \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon/3,+\infty[}(|S_k(f-g)|/\sqrt{k})}{k} + \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon/3,+\infty[}(|M_k(f-g)|/\sqrt{k})}{k}.\end{aligned}$$

We have $(S_n(g) - M_n(g))/\sqrt{n} = (Pg(W_0) - Pg(W_n))/\sqrt{n} \to 0$ $\mathbb{P}_m$-a.s., by the Borel–Cantelli lemma. Hence,

$$\limsup_{n\to\infty} \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon/3,+\infty[}(|S_k(g) - M_k(g)|/\sqrt{k})}{k} = 0 \qquad \mathbb{P}_m\text{-a.s.} \tag{28}$$

Using Proposition 3.1 and the notation (20), we have

$$S_n(f-g) = M_n^-(f-g) + M_n^+(f-g) + \mathcal{R}(f-g)\circ\theta^{n-1} - \mathcal{R}(f-g)\circ\theta^{-1}. \tag{29}$$

Using the inequality $\max(\|\mathcal{D}^-(f-g)\|_{L^2(\mathbb{P}_m)}, \|\mathcal{D}^+(f-g)\|_{L^2(\mathbb{P}_m)}) \leq \|f-g\|_{-1}$ and the almost sure central limit theorem for martingales and sums of stationary ergodic reverse martingale differences (see [10]), we obtain

$$\limsup_n \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon/3,+\infty[}(|S_k(f-g)|/\sqrt{k})}{k} \leq 4\nu(]\varepsilon/(9\eta), +\infty[) \qquad \mathbb{P}_m\text{-a.s.}$$

Using $\|\mathcal{D}(f-g)\|_{L^2(\mathbb{P}_m)} \le 2\|f-g\|_{-1}$ and applying the almost sure central limit theorem for $M_n(f-g)$, we obtain

$$\limsup_n \frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{]\varepsilon/3,+\infty[}(|M_k(f-g)|/\sqrt{k})}{k} \le 2\nu(]\varepsilon/(6\eta),+\infty[) \qquad \mathbb{P}_m\text{-a.s.}$$

Combining the above estimates and letting $\eta \to 0$ (along rational numbers), we obtain

$$\begin{aligned}&\limsup_n \left|\frac{1}{\log n}\sum_{k=1}^n \frac{\mathbf{1}_{[x,y]}(S_k/\sqrt{k})}{k} - \nu([x,y])\right| \\ &\le \nu([y-\varepsilon,y+\varepsilon]) + \nu([x-\varepsilon,x+\varepsilon]) \qquad \mathbb{P}_m\text{-a.s.},\end{aligned}$$

which converges to 0 as $\varepsilon \to 0$ (along rational numbers) and proves the result. □

Remark. When P is normal, or when P satisfies the sector condition, the theorem applies to $f \in \mathcal{H}_{-1} = \mathbf{B}$ (by Lemma 2.4 or Proposition 2.7, respectively).

6. A characterization of the sector condition

In this section we give a characterization of the sector condition (13), in terms of the numerical range of P on the complex $L^2(m)$. Recall that by Proposition 2.7, the sector condition guarantees that $\mathbf{B} = \mathcal{H}_{-1}$.

This characterization will be useful in the next section; it allows us to prove that the sector condition is stable by convex combinations.

We begin with some equivalent formulation of (13).

Lemma 6.1. *The following are equivalent for P on the real $L^2(m)$:*

(i) *P satisfies the sector condition: there exist $K > 0$ such that*

$$|\langle (I-P)g,h\rangle_0|^2 \le K\langle (I-P)g,g\rangle_0\langle (I-P)h,h\rangle_0 \quad \forall g,h \in L^2_0(\mathbb{S},m)\,. \tag{30}$$

(ii) *There exists $L > 0$ such that*

$$|\langle (P-P^*)g,h\rangle_0|^2 \le L\langle (I-P)g,g\rangle_0\langle (I-P)h,h\rangle_0 \quad \forall g,h \in L^2_0(\mathbb{S},m)\,.$$

(iii) *There exists $C > 0$ such that*

$$|\langle (P-P^*)g,h\rangle_0| \le C\big(\langle (I-P)g,g\rangle_0 + \langle (I-P)h,h\rangle_0\big) \quad \forall g,h \in L^2_0(\mathbb{S},m)\,. \tag{31}$$

Proof. By (7), on the real $L^2(m)$, (30) holds for P if and only if it holds for P^*, and then (30) holds for P_s (note that necessarily $K \ge 1$). Since $\frac{1}{2}(P-P^*) = (I-P_s) - (I-P)$, we easily obtain that (i) implies (ii).

Assume (ii). Since $I - P_s = \big(\sqrt{I-P_s}\big)^2$, symmetry of P_s and the Cauchy–Schwarz inequality yield

$$|\langle (I-P_s)g,h\rangle_0| \le \|\sqrt{I-P_s}g\|_0 \cdot \|\sqrt{I-P_s}h\|_0 = \Big(\langle (I-P_s)g,g\rangle_0\langle (I-P_s)h,h\rangle_0\Big)^{1/2}.$$

This together with (ii) yields (i), since $I - P = I - P_s + \frac{1}{2}(P^* - P)$.

(ii) implies (iii) since

$$\begin{aligned}|\langle (P-P^*)g,h\rangle_0| &\le \sqrt{L}\sqrt{\langle (I-P)g,g\rangle_0}\sqrt{\langle (I-P)h,h\rangle_0}\\ &\le \frac{1}{2}\sqrt{L}\Big(\langle (I-P)g,g\rangle_0+\langle (I-P)h,h\rangle_0\Big).\end{aligned}$$

Assume that (iii) holds. For $\lambda\in\mathbb{R}$ we replace g in (31) by λg, and obtain

$$\lambda\langle (P-P^*)g,h\rangle_0 \le |\langle (P-P^*)(\lambda g),h\rangle_0| \le \lambda^2 C\langle (I-P)g,g\rangle_0 + C\langle (I-P)h,h\rangle_0\,.$$

We thus have a non-negative quadratic real polynomial, so its discriminant is non-positive, which yields (ii) with $L=4C^2$. □

We shall now characterize the sector condition in terms of the numerical range. Recall that the numerical range of an operator T on a *complex* Hilbert space $\mathbf{H}$ is defined by

$$\Theta(T) := \{\langle Tf,f\rangle :\ \|f\|=1\}\,.$$

The numerical range of an operator is a convex set (Hausdorff–Toeplitz Theorem) whose closure contains the spectrum of the operator. When the operator is normal (i.e. $T^*T=TT^*$), the closure of the numerical range is exactly the convex hull of the spectrum, see e.g. [6].

For $C\ge 1$, define the regions

$$\begin{aligned}\Gamma_C &:= \{z\in\mathbb{C},\ |z|\le 1,\ |1-z|\le C\,\mathrm{Re}(1-z)\}\\ \tilde\Gamma_C &:= \{z\in\mathbb{C},\ |z|\le 1,\ |\,\mathrm{Im}\,z|\le C\,\mathrm{Re}(1-z)\}\,.\end{aligned}$$

For $|z|\le 1$ let $\alpha_z\in(-\pi,\pi)$ be the angle between the line segment $[1,z]$ and the x-axis. Then $\Gamma_C=\{|z|\le 1:\ \cos\alpha_z\ge 1/C\}$ and $\tilde\Gamma_C=\{|z|\le 1:\ |\tan\alpha_z|\le C\}$. By simple trigonometry, for every $C\ge 1$,

$$\Gamma_C\subset\tilde\Gamma_C\le\Gamma_{(C^2+1)^{1/2}}\,. \tag{32}$$

Lemma 6.2. *Let T be an operator on a real Hilbert space $\mathbf{H}$. The numerical range of the operator T acting on the complexified Hilbert space $\mathbf{H}^{\mathbb{C}}$ is contained in $\tilde\Gamma_C$ for some $C\ge 1$, if and only if for every $g,h\in\mathbf{H}$,*

$$|\langle (T-T^*)g,h\rangle_{\mathbf{H}}| \le C\big(\langle g-Tg,g\rangle_{\mathbf{H}}+\langle h-Th,h\rangle_{\mathbf{H}}\big)\,.$$

Proof. First, note that $\Theta(T)\subset\tilde\Gamma_C$ means that, for every $f\in\mathbf{H}^{\mathbb{C}}$ with $\|f\|_{\mathbf{H}^{\mathbb{C}}}=1$,

$$|\,\mathrm{Im}(\langle Tf,f\rangle_{\mathbf{H}^{\mathbb{C}}})|\le C(1-\mathrm{Re}(\langle Tf,f\rangle_{\mathbf{H}^{\mathbb{C}}})=C(\|f\|^2_{\mathbf{H}^{\mathbb{C}}}-\mathrm{Re}(\langle Tf,f\rangle_{\mathbf{H}^{\mathbb{C}}})\,.$$

Let $f\in\mathbf{H}^{\mathbb{C}}$ and write $f:=g+ih$ with $g,h\in\mathbf{H}$. Then, $\|f\|^2_{\mathbf{H}^{\mathbb{C}}}=\|g\|^2_{\mathbf{H}}+\|h\|^2_{\mathbf{H}}$ and

$$\mathrm{Re}\langle Tf,f\rangle_{\mathbf{H}^{\mathbb{C}}}=\langle Tg,g\rangle_{\mathbf{H}}+\langle Th,h\rangle_{\mathbf{H}};\qquad \mathrm{Im}\langle Tf,f\rangle_{\mathbf{H}^{\mathbb{C}}}=\langle (T^*-T)g,h\rangle.$$

Since for $\|f\|_{\mathbf{H}^{\mathbb{C}}}=1$ we have

$$1-\mathrm{Re}\langle Tf,f\rangle=\|g\|^2+\|h\|^2-\langle Tg,g\rangle-\langle Th,h\rangle=\langle g-Tg,g\rangle+\langle h-Th,h\rangle,$$

the result follows. □

Proposition 6.3. *Let P be a Markov operator. Then P satisfies the sector condition if and only if the numerical range of P on the complex $L^2(m)$ is contained in some $\tilde{\Gamma}_C$.*

Proof. By Lemma 6.1, the sector condition (30) is equivalent to (31), which is equivalent to the numerical range being contained in some $\tilde{\Gamma}_C$ by the previous lemma. □

Corollary 6.4. *Let $P_1, \dots, P_d$ be Markov operators on $L^2(m)$ which satisfy the sector condition. Then any convex combination $P = \sum_{1 \le i \le d} a_i P_i$ also satisfies the sector condition.*

Proof. Let $\Theta(P_i) \subset \tilde{\Gamma}_{C_i}$ and put $C = \max\{C_1, \dots, C_d\}$. Then $\Theta(P) \subset \tilde{\Gamma}_C$. □

As we shall see in the next proposition as well as in the next section, it is interesting to consider other types of regions (smaller than some Γ_C), known as Stolz regions.

Definition. A *Stolz region* is a subset of the closed unit disk which is the convex hull of the point 1 and a disk centered at the origin with radius strictly less than 1.

Remark. A closed subset A of Γ_C (of $\tilde{\Gamma}_C$) is included in some Stolz region if (and only if) $A \cap \{z : |z| = 1\} \subset \{1\}$. In particular, the numerical range of a Markov operator is in a Stolz region if and only if it is included in some Γ_C and admits only 1 as unimodular complex number.

Stolz regions may be parametrized in the following way. Let $\alpha \in [0, \pi/2[$ be the angle between the x-axis and the tangent line to the disk from the point 1. Then the radius of the disk is $\sin\alpha$. Thus, the Stolz region, denoted by Λ_α, is the convex hull of the disk of radius $\sin\alpha$ centered at 0 and the point 1. The boundary of Λ_α consists of an arc of the circle $\{|z| = \sin\alpha\}$ and the two linear segments $[1, z_\alpha]$ and $[1, \bar{z}_\alpha]$, where $z_\alpha = 1 - \cos\alpha e^{i\alpha}$.

The following property is well-known, but we have no reference for a proof, so we give one below.

Lemma 6.5. *Let $S = \Lambda_\alpha$ be a Stolz region. Then there exists $C \ge 1$ such that $|1 - z| \le C(1 - |z|)$ for every $z \in S$.*

Proof. Let $r < 1$ be the radius of the disk in S. Since $-r$ is the point with $|z| \le r$ farthest from 1, we have

$$|1 - z| \le 1 + r = \frac{1 + r}{1 - r}(1 - r) \le \frac{1 + r}{1 - r}(1 - |z|) \quad \forall |z| \le r.$$

The set of z such that $|1-z| \le \frac{1+r}{1-r}(1-|z|)$ is convex and contains 1 and the disk of radius r, so contains S. □

It follows from the lemma that any Stolz region is contained in some Γ_C. On the other hand, a closed subregion of Γ_C, whose intersection with the unit circle is (at most) $\{1\}$, is contained in some Stolz region.

Example 1. Let S be a Markov operator on $(\mathbb{S}, \mathcal{S})$ with invariant probability m which is ergodic, and let $\{p_k : k \in \mathbb{Z}\}$ be a probability distribution on $\mathbb{Z}$. Define $P := p_0 I + \sum_{k \ge 1}(p_{-k} S^{*k} + p_k S^k)$, which is a Markov operator with m invariant. If $\{k \in \mathbb{Z} : p_k \ne 0\}$, the support of $\{p_k\}$, is not contained in $d\mathbb{Z}$ for any $d > 1$, then P is ergodic. In general P is not normal. Let U be the operator induced by the two-sided Markov shift of S, and define $Q = \sum_{k \in \mathbb{Z}} p_k U^k$. Then Q is a normal operator, and the property of the two-sided shift yields that the projection E on $\mathbb{S}$ satisfies $EU^n = S^n$ and $EU^{*n} = S^{*n}$. A sufficient condition for $\sigma(Q)$ to be contained in Stolz region was given by Bellow, Jones and Rosenblatt [5]: $\sum_{k \in \mathbb{Z}} k^2 p_k < \infty$ and $\sum_{k \in \mathbb{Z}} k p_k = 0$. In that case, by normality, also $\Theta(Q)$ is contained in the same Stolz region. Since $EQ = P$, we obtain that

$\Theta(P)$ is contained in a Stolz region. When S is induced by an *invertible* ergodic transformation preserving m (the case treated in [5]), we can take $U = S$ (and then P is normal).

Next, we give a sufficient condition for the sector condition.

Proposition 6.6. *Let T be an operator on a complex Hilbert space $\mathcal{H}$. If for some $a > 0$ we have*

$$\|f\|^2 - \|Tf\|^2 \geq a|\langle (I-T)f, f\rangle| \quad \forall f \in \mathcal{H}, \tag{33}$$

then T is a contraction, the numerical range $\Theta(T)$ is included in a Stolz region, and $\sup_n n\|T^n - T^{n+1}\| < \infty$.

Proof. The inequality (33) yields $\|Tf\|^2 \leq \|f\|^2$, so T is a contraction. The (easily checked) identity

$$\|f\|^2 - \|Tf\|^2 + \|(I-T)f\|^2 = 2\,\mathrm{Re}\langle (I-T)f, f\rangle$$

and (33) yield $|\langle (I-T)f, f\rangle| \leq 2a^{-1}\,\mathrm{Re}\langle (I-T)f, f\rangle$ for every f. Fix $\|f\| = 1$ and put $z = \langle Tf, f\rangle$. Then $|1-z| \leq 2a^{-1}\,\mathrm{Re}(1-z)$, which shows that $a \leq 2$. Hence $\Theta(T) \subset \Gamma_{2/a}$. To show that $\Theta(T)$ is contained in a Stolz region, we show that $\overline{\Theta(T)}$ intersects the unit circle at most at 1. Let $|\zeta| = 1$ and $\{f_n\}$ with $\|f_n\| = 1$ such that $z_n := \langle Tf_n, f_n\rangle \to \zeta$. Then $\|Tf_n\| \to 1$, and by (33) $z_n \to 1$, so $\zeta = 1$. Hence $\Theta(T)$ is contained in a Stolz region, which implies (see [11, Proposition 2.3]) that $\sup_n n\|T^n - T^{n+1}\| < \infty$. □

Remark. Dungey [25] proved that (33) implies $\sup_n n\|T^n - T^{n+1}\| < \infty$, using a semi-group "domination", without referring to the numerical range. The beginning of our proof follows the beginning of his proof.

Corollary 6.7. *Let P be a Markov operator. If, for some $a > 0$, P on the complex $L^2(m)$ satisfies*

$$\|f\|^2 - \|Pf\|^2 \geq a|\langle (I-P)f, f\rangle| \quad \forall f \in L^2(m) \tag{34}$$

then P satisfies the sector condition.

Theorem 6.8. *Let T be a normal contraction of a complex Hilbert space. Then the following are equivalent:*

(i) $\sup_n n\|T^n(I-T)\| < \infty$.
(ii) *$\sigma(T)$ is contained in a Stolz region.*

Proof. Bellow, Jones and Rosenblatt [5, p. 11] proved that (ii) implies (i), while Nagy and Zemánek [44] proved that (for any power-bounded operator on a complex Banach space) (i) implies (ii). See [11] for additional information and references. □

7. On the relationship between $\mathcal{H}_{-1}$ and $\sqrt{I-P}L^2(m)$

Thanks to the previous section, we know conditions under which $\mathbf{B} = \mathcal{H}_{-1}$, hence, under which Theorem 4.1 and Theorem 5.1 hold for every $f \in \mathcal{H}_{-1}$. Therefore it remains to find conditions (suitable in view of applications) guaranteeing that $f \in \mathcal{H}_{-1}$. In particular we shall investigate the relationship between $\mathcal{H}_{-1}$ and $\sqrt{I-P}L^2(m)$. This study is motivated by the fact that $\sqrt{I-P}L^2(m)$ appears naturally in proving the CLT for a normal Markov operator P (as proved by Gordin and Lifshitz [31,7]), and by a recent result of Cohen–Cuny–Lin [11] (see Proposition 7.4 below) which characterizes the space $\sqrt{I-P}L^2(m)$ when the numerical range of P (not necessarily normal) is in a Stolz region.

Throughout this section P is a Markov operator with invariant probability m, which is ergodic (the only fixed points in $L^2(m)$ are constants). The operator P acts as a contraction on the real and on the complex $L^2(m)$ spaces.

Recall that when P is normal, we have $\mathbf{B} = \mathcal{H}_{-1}$ by Lemma 2.4. The (annealed) central limit theorem of Gordin and Lifshitz [31] (proved in [7, Sections IV.7–IV.9]; see also [22]) says that if P is normal and f is real in $\sqrt{I-P}L^2(m)$, then $\frac{1}{\sqrt{n}}S_n$ converges in distribution, in $(\Omega, \mathbb{P}_m)$, to a normal distribution. For P symmetric, $\mathcal{H}_{-1} = \sqrt{I-P}L^2(m) = \sqrt{I-P}L^2_0(m)$ by the definition; however, an example in [23, p. 15] shows that for P normal it is possible that

$$\mathbf{B} = \mathcal{H}_{-1} := \sqrt{I-P_s}L^2_0(m) = (I-P)L^2(m) \subsetneq \sqrt{I-P}L^2(m).$$

In such a case (see Proposition 9.6 below for other examples), our limit theorems apply only to coboundaries, so have no novelty. We therefore want to study when $\sqrt{I-P}L^2(m) \subset \mathcal{H}_{-1}$.

In the complex $L^2(m)$, when P is normal, a spectral characterization of $f \in \sqrt{I-P}L^2(m)$ [21] is

$$\int_{\sigma(P)} \frac{1}{|1-z|}\rho_f(dz) < \infty, \tag{35}$$

where ρ_f is the spectral measure of f – a finite measure on the closed unit disk $\overline{\mathbb{D}}$ supported on the spectrum $\sigma(P)$, such that $\langle P^n f, f\rangle_0 = \int_{\sigma(P)} z^n \rho_f(dz)$ for every $n \geq 0$.

In order to be able to use the spectral theorem, we will work in this section with the complex $L^2(m)$; since P is positive, the real $L^2(m)$ is an invariant subspace, and the consequences of our study will apply to real functions. Note that the series expansion [21] of the solution g of $\sqrt{I-P}g = f$ in the complex $L^2_0(m)$ shows that g is real if f is real.

Proposition 7.1. *Let P be normal. Then*

$$\mathcal{H}_{-1} = \{f \in L^2(\mathbb{S}, m) : \int_{\sigma(P)} \frac{1}{1-\operatorname{Re} z} d\rho_f(z) < \infty\}. \tag{36}$$

Consequently, $\{f \in L^2(m) : \sum_{n\geq 0} \|P^n f\|_0^2 < \infty\} \subset \mathbf{B} = \mathcal{H}_{-1} \subset \sqrt{I-P}L^2(m)$. If moreover, the spectrum of P is included in some region Γ_C, $C \geq 1$, we actually have $\mathbf{B} = \mathcal{H}_{-1} = \sqrt{I-P}L^2(m)$.

Proof. The equality $\mathbf{B} = \mathcal{H}_{-1}$ is in Lemma 2.4. By the functional calculus, $P_s = \operatorname{Re} P$ corresponds to $\operatorname{Re} z = (z+\bar{z})/2$. Thus, $f \in \mathcal{H}_{-1}$ if and only if $f \in \sqrt{1-\operatorname{Re} P}L^2(m)$, which yields (36).

Now $0 \leq 1 - \operatorname{Re} z \leq |1-z|$ for $|z| \leq 1$ yields that $\int_{\sigma(P)} \frac{1}{|1-z|}\rho_f(dz) < \infty$ when $f \in \mathcal{H}_{-1}$, by (36).

It follows from (36) that $f \in L^2(\mathbb{S}, m)$ is in $\mathcal{H}_{-1}$ if $\int_D \frac{\rho_f(dz)}{1-|z|} < \infty$, which is equivalent to

$$\sum_n \|P^n f\|_0^2 < \infty\,. \tag{37}$$

In particular, (37) implies $f \in \sqrt{I-P}L^2(m)$.

When $\sigma(P)$ is included in Γ_C, the equality $\mathcal{H}_{-1} = \sqrt{I-P}L^2(m)$ follows from (36), (35) and the definition of Γ_C. $\square$

Corollary 7.2. *Let P be a normal Markov operator. If P satisfies the sector condition, then $\mathbf{B} = \mathcal{H}_{-1} = \sqrt{I-P}L^2(m)$.*

In general, even for P normal which is mixing, $f \in \sqrt{I-P}L^2$ need not satisfy (37), as shown by the example of [23] with $\mathcal{H}_{-1} = (I-P)L^2(m) \subsetneq \sqrt{I-P}L^2(m)$.

Lemma 7.3. *Let Q be a normal contraction on a real Hilbert space $\mathbf{H}$, such that its spectrum on the complexification $\mathbf{H}^C$ is in some region Γ_C. Then*

$$\|\sqrt{I-Q}g\| \le \sqrt{C}\|\sqrt{I-Q_s}g\| \le \sqrt{C}\|\sqrt{I-Q}g\| \quad \textit{for every} \quad g \in \mathbf{H}.$$

Proof. By the spectral theorem and the definition of Γ_C we have

$$\|\sqrt{I-Q}g\|_{\mathbf{H}}^2 = \int\limits_{\Gamma_C} |1-z|\rho_g(dz) \le C \int\limits_{\Gamma_C} (1-\operatorname{Re} z)\rho_g(dz) = C\|\sqrt{I-Q_s}g\|_{\mathbf{H}}^2.$$

Similarly, normality always implies $\|\sqrt{I-Q_s}g\|^2 \le \|\sqrt{I-Q}g\|^2$. $\square$

Definition. A contraction P on the real $L^2(m)$ *admits a real normal dilation* if there exists a *real* real Hilbert space $\mathbf{H}$ containing $L^2(m)$ as a subspace and a normal contraction Q on $\mathbf{H}$, such that $P^n = EQ^n$, for every $n \ge 1$, where E is the orthogonal projection from $\mathbf{H}$ onto $L^2(m)$. Of course, the two-sided Markov shift of P Markovian with invariant probability is a normal dilation, so every such Markovian P has a positive normal dilation; our interest will be in normal dilations with particular properties.

It will be important for our study that P have a *real* normal dilation. Of course, Q on $\mathbf{H}$ extends to a normal contraction on $\mathbf{H}^C$, the complexification of H, still denoted by Q, which is a normal dilation of P on the *complex* $L^2(m)$.

Note that the numerical range of the dilation Q (on $\mathbf{H}^C$) contains the numerical range of P (on the complex $L^2(m)$).

Example 1 in the previous section is of an ergodic Markov operator P which has a real normal dilation with spectrum in a Stolz region.

Proposition 7.4. *Let P be an ergodic Markov operator on $L^2(m)$ admitting a real normal dilation Q on $\mathbf{H}$ whose spectrum (on $\mathbf{H}^C$) is included in a region Γ_C, $C \ge 1$. Then the following are equivalent for $f \in L^2(m)$:*

(i) $f \in \sqrt{I-P}L^2(m)$;
(ii) $f \in \sqrt{I-P_s}L^2(m) = \mathcal{H}_{-1}$;
(iii) $\sup_{n\ge1} \mathbb{E}(S_n^2(f))/n < \infty$;
(iv) $\sup_{\lambda\to1}(1-\lambda)\mathbb{E}\Big(\big(\sum_{n\ge0}\lambda^n f(W_n)\big)^2\Big) < \infty$;
(v) $\liminf_{\lambda\to1}\sum_{n\ge0}\lambda^n\langle f, P^n f\rangle_0 < \infty$;
(vi) $\liminf_{\lambda\to1}\sum_{n\ge0}\lambda^n\langle f, Q^n f\rangle_{\mathbf{H}} < \infty$;
(vii) $f \in \sqrt{I-Q_s}\mathbf{H}$;
(viii) $f \in \sqrt{I-Q}\mathbf{H}$;

If moreover, the spectrum of Q is included in a Stolz region, then the above conditions are all equivalent to

(ix) $\sum_{n\ge0}\|P^n f\|_0^2 < \infty$.
(x) $\sum_{n\ge0}|\langle f, P^n f\rangle_0| < \infty$.

Remarks. we will prove in Proposition 9.2 that (ix) $\Rightarrow$ (i) for *any* Markov operator. The set of equivalences (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (v), for P satisfying the sector condition, was proved in [34].

Proof. We first prove (i) $\Rightarrow$ (ii). Let $f \in \sqrt{I-P}L^2(m)$. Then there exists $h \in L_0^2(m)$ (unique by ergodicity) such that $f = \sqrt{I-P}h$. Since $\langle P^n h, g\rangle_0 = \langle Q^n h, g\rangle_{\mathbf{H}}$ for $g \in L^2(m)$, using Lemma 7.3 we obtain

$$\begin{aligned}|\langle f, g\rangle_0|^2 &= |\langle \sqrt{I-P}h, g\rangle_0|^2 = |\langle \sqrt{I-Q}h, g\rangle_{\mathbf{H}}|^2 \le \|h\|_0^2 \|\sqrt{I-Q^*}g\|_{\mathbf{H}}^2 \\ &\le C^2\|h\|_0^2\|\sqrt{I-Q_s}g\|_{\mathbf{H}}^2 = C\|h\|_0^2\langle (I-Q_s)g, g\rangle_{\mathbf{H}} = C\|h\|_0^2\langle (I-P_s)g, g\rangle_0 \\ &= C\|h\|_0^2\langle (I-P)g, g\rangle_0,\end{aligned}$$

with the last equality by (7) since g is real valued. By Lemma 2.3(i), $f \in \mathcal{H}_{-1}$.

(ii) $\Rightarrow$ (iii) follows from Theorem 3.2, equation (21).

(iii) $\Rightarrow$ (iv) follows from comparison between Cesàro and Abel means.

Let us prove that (iv) $\Rightarrow$ (v). Let $0 \le \lambda < 1$. We have

$$\begin{aligned}\mathbb{E}\Big(\big(\sum_{n\ge0}\lambda^n f(W_n)\big)^2\Big) &= \sum_{n\ge0}\lambda^{2n}\mathbb{E}(f(W_n)^2) + 2\sum_{0\le n<m}\lambda^{n+m}\mathbb{E}(f(W_n)f(W_m)) \\ &= \frac{\|f\|_0^2}{1-\lambda^2} + \frac{2}{1-\lambda^2}\sum_{m\ge1}\lambda^m\langle f, P^m f\rangle_0\end{aligned}$$

(v) $\Rightarrow$ (vi) since the series have the same terms, by the definition of a dilation.

Let us prove that (vi) $\Rightarrow$ (vii). Denote $Q_s = \frac{1}{2}(Q+Q^*)$ and $\mathbf{H}_{-1} := \sqrt{I-Q_s}\mathbf{H}$. Since Q acts on a real Hilbert space $\mathbf{H}$, we can apply the proof of Lemma 2.3 to Q – all that we needed was that the inner product is real – and obtain the characterization of $\mathbf{H}_{-1}$. Also the sector condition can be defined for Q on $\mathbf{H}$, and its characterization in Proposition 6.3 is valid, with the same proof.

Since Q on $\mathbf{H}^C$ is normal, $\Theta(Q) \subset \Gamma_C$, and by Proposition 6.3 Q on $\mathbf{H}$ satisfies the sector condition (30); hence for fixed $0 < \lambda < 1$ and every $g \in \mathbf{H}$ we have

$$|\langle (I-Q)\sum_{n\ge0}\lambda^n Q^n f, g\rangle_{\mathbf{H}}|^2 \le K|\langle (I-Q)\sum_{n\ge0}\lambda^n Q^n f, \sum_{n\ge0}\lambda^n Q^n f\rangle_{\mathbf{H}}|\,|\langle (I-Q)g, g\rangle_{\mathbf{H}}|\,.$$

For λ fixed, the characterization of the $\mathbf{H}_{-1}$-norm in (6) and the identity (12) with P replaced by Q yield

$$\begin{aligned}\|\langle (I-Q)\sum_{n\ge0}\lambda^n Q^n f\|_{\mathbf{H}_{-1}}^2 &\le K|\langle (I-Q)\sum_{n\ge0}\lambda^n Q^n f, \sum_{n\ge0}\lambda^n Q^n f\rangle_{\mathbf{H}}| \\ &= \frac{K}{\lambda}\langle f - A_\lambda f, \sum_{n\ge0}\lambda^n Q^n f\rangle_{\mathbf{H}} \le \frac{K}{\lambda}\langle f, \sum_{n\ge0}\lambda^n Q^n f\rangle_{\mathbf{H}}\,.\end{aligned}$$

By the Banach–Saks Theorem, using (vi), there exists $\lambda_m \to 1$ such that the Cesáro averages of $((I - Q)\sum_{n\ge0}\lambda_m^n Q^n f)_{m\ge1}$ converge in $\mathbf{H}_{-1}$, say to $h \in \mathbf{H}_{-1}$. By (5), the convergence holds also in $\mathbf{H}$; but in that case, by (12), the limit in $\mathbf{H}$ is known to be f, so that, $f = h \in \mathbf{H}_{-1}$.

The implication (vii) $\Rightarrow$ (viii) follows from Proposition 7.1.

By Derriennic–Lin [21], for any contraction T on a reflexive Banach space X, $f \in \sqrt{I-T}X$ if and only if $\sum_{n\ge0}\beta_n T^n f$ converges in X (see (3) for the definition of $\{\beta_n\}$). Hence, (viii) $\Rightarrow$ (i) follows from the definition of a dilation.

We now assume that $\sigma(Q)$ is contained in a Stolz region. Then $\Theta(Q)$, the numerical range of Q, is in a Stolz region, so $\Theta(P) \subset \Theta(Q)$ is in a Stolz region, and by [11, Proposition 6.4] (ix) $\Leftrightarrow$ (i).

(x) $\Rightarrow$ (v) is immediate.

We now prove (viii) $\Rightarrow$ (x). By the spectral theorem (see [21]), $f \in \sqrt{I-Q}\mathbf{H}$ is equivalent to $\int_{\sigma(Q)}\frac{1}{|1-z|}\rho_f(dz) < \infty$. Since $\sigma(Q)$ is contained in a Stolz region, by Lemma 6.5 there exists $C > 0$

such that $|1-z| \le C(1-|z|)$ for $z \in \sigma(Q)$, so $\int_{\sigma(Q)} \frac{1}{1-|z|}\rho_f(dz) < \infty$. For any positive integer j we have

$$\sum_{n=0}^{j} |\langle f, Q^n f\rangle_{\mathbf{H}}| = \sum_{n=0}^{j} |\int_{\sigma(Q)} z^n \rho_f(dz)| \le \sum_{n=0}^{j} \int_{\sigma(Q)} |z|^n \rho_f(dz) \le \int_{\sigma(Q)} \frac{1}{1-|z|}\rho_f(dz).$$

Since $\langle f, P^n f\rangle_0 = \langle f, Q^n f\rangle_{\mathbf{H}}$, letting $j \to \infty$ we obtain

$$\sum_{n=0}^{\infty} |\langle f, P^n f\rangle_0| = \sum_{n=0}^{\infty} |\langle f, Q^n f\rangle_{\mathbf{H}}| \le \int_{\sigma(Q)} \frac{1}{1-|z|}\rho_f(dz) < \infty,$$

so (x) holds. $\square$

Remarks. 1. Clearly (x) implies

(x′) $\sum_{n\ge 0}\langle f, P^n f\rangle$ converges.

On the other hand, (x′) ⇒ (v) follows from Abel's Theorem for power series. Thus, when $\sigma(Q)$ is included in a Stolz region, (x′) ⇔ (x).

2. In general, even if the operator is normal, (x) does not imply (ix), nor (i). At the end of [22] there is an example of T unitary on $L^2(m)$ (induced by a two-sided shift) and $f \in L^2(m)$ such that $\sum_{n\ge 1} |\langle T^n f, f\rangle| < \infty$; since $\|T^n f\| = \|f\|$ for $n \ge 1$, condition (ix) cannot hold. In that example $f \notin \sqrt{I-T}L^2(m)$.

Corollary 7.5. *Let P be an ergodic Markov operator with invariant probability m admitting a real normal dilation Q on $\mathbf{H}$ such that $\sigma(Q)$ is included in a region Γ_C, $C \ge 1$. Then*

$$\mathbf{B} = \mathcal{H}_{-1} = \sqrt{I-P}L^2(m).$$

Proof. Since Q is normal, the assumption implies that also the numerical range $\Theta(Q)$ is included in $\Gamma(C)$, hence also $\Theta(P) \subset \Gamma_C$. Then, by Proposition 6.3, P satisfies the sector condition. Hence $\mathbf{B} = \mathcal{H}_{-1}$ by Proposition 2.7. By the equivalence of (i) and (ii) of Proposition 7.4, $\mathcal{H}_{-1} = \sqrt{I-P}L^2(m)$. $\square$

In order to be able to use Proposition 7.4 we shall now characterize the existence of a normal dilation whose spectrum is included in a Stolz region.

Definition 2. For $\epsilon \in (0,1)$ and $0 \le t \le 1$ we have the expansion $(1-t)^\epsilon = 1 - \sum_{n=1}^{\infty} a_n^{(\epsilon)} t^n$, with $a_n^{(\epsilon)} > 0$ and $\sum_{n\ge 1} a_n^{(\epsilon)} = 1$. Then, for any power-bounded operator T on a (real or complex) Banach space, we define $(I-T)^\varepsilon := I - \sum_{n=1}^{\infty} a_n^{(\epsilon)} T^n$. In particular the series defines $(1-z)^\epsilon$ for $z \in \overline{\mathbb{D}}$.

Lemma 7.6. *Let $S = \Lambda_\alpha$ be a Stolz region. Then there exists $\epsilon \in (0,1)$ such that* $\sup_{z\in S} |1-(1-z)^{1+\epsilon}| \le 1$.

Proof. Fix ϵ and put $a_n = a_n^{(\epsilon)}$. We have

$$1-(1-z)^{1+\epsilon} = 1-(1-z)(1-z)^\epsilon = z + \sum_{n=1}^{\infty} a_n z^n (1-z). \tag{38}$$

Thus for $\delta \in (1,2)$, the function $1-(1-z)^\delta$ is holomorphic in the open unit disk $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$.

We show that there exists $\gamma > 1$ (close enough to 1), which depends on α, such that for any $\delta \in (1,\gamma)$ and any $z \in \Lambda_\alpha$,

$$|1-(1-z)^\delta| \le 1\,. \tag{39}$$

Since the function $z \mapsto 1-(1-z)^\delta$ is holomorphic in the interior of Λ_α and continuous on its closure, by the maximum modulus principle, we just have to prove (39) on the boundary of Λ_α (which consists of an arc and two line segments).

Let $\epsilon \in (0,1)$ and put $\delta = 1+\epsilon$. Since $a_n := a_n^{(\epsilon)}$ is decreasing to zero [21], for $|z| \le r := \sin\alpha$ the expansion (38) yields

$$\begin{aligned}|1-(1-z)^\delta| &= |z + a_1 z - \sum_{n=2}^{\infty}(a_{n-1}-a_n)z^n| \le r + a_1 r + \sum_{n=2}^{\infty}(a_{n-1}-a_n)r^n \\ &\le r + a_1 r + \sum_{n=2}^{\infty}(a_{n-1}-a_n)r = (1+2a_1)r.\end{aligned}$$

But $a_1^{(\epsilon)} = \epsilon$ [21], so $\max_{|z|\le r}|1-(1-z)^\delta| \le (1+2\epsilon)r$, which tends to $r<1$ as $\epsilon \to 0^+$, so for $\epsilon > 0$ small enough (fixed) (39) holds for $|z| \le \sin\alpha$.

The points on the line segment $[1, z_\alpha]$ are of the form $z = 1 - re^{i\alpha}$, with $0 \le r \le \cos\alpha$. For such z we have

$$\begin{aligned}|1-(1-z)^\delta|^2 &= |1-r^\delta \mathrm{e}^{i\alpha\delta}|^2 = 1 + r^{2\delta} - 2r^\delta\cos(\alpha\delta) \\ &\le 1 + r^\delta\big((\cos\alpha)^\delta - 2\cos(\alpha\delta)\,\big) := 1 + r^\delta\psi_\alpha(\delta)\,.\end{aligned}$$

Since $\psi_\alpha(\delta)$ is clearly continuous on $[1,2]$ and $\psi_\alpha(1) = -\cos\alpha < 0$, there exists $\gamma > 1$, such that for every $\delta \in (1,\gamma)$ we have $\psi_\alpha(\delta) < 0$, so (39) holds also on the line segments of the boundary of Λ_α. $\square$

Remark. The lemma is stated by Dungey [26, p. 1737], and his proof uses some semigroup theory.

Theorem 7.7. *Let T be a contraction of a real Hilbert space $\mathbf{H}$. Then T admits a normal dilation Q with $\sigma(Q)$ in a Stolz region if and only if there exists a contraction S on $\mathbf{H}$ such that $T = I-(I-S)^\alpha$ for some $\alpha \in (0,1)$.*

Proof. Assume that $T = I-(I-S)^\alpha$ for some $\alpha \in (0,1)$, with S a contraction on $\mathbf{H}$. Let U be the unitary dilation of S on a larger *real* Hilbert space $\mathbf{K}$ (see for instance [48, Appendix, Section 4]), and define $Q = I-(I-U)^\alpha$. Then Q is a normal dilation of T, and by Dungey [26, Theorem 1.1] Q satisfies $\sup_n n\|Q^n - Q^{n+1}\| < \infty$. By Theorem 6.8 $\sigma(Q)$ on the complexification $\mathbf{K}^{\mathbb{C}}$ is contained in a Stolz region.

Assume now that Q on $\mathbf{K}$ is a normal dilation of T, with $\sigma(Q)$ (on $\mathbf{K}^{\mathbb{C}}$) in a Stolz region. By Lemma 7.6 there is $\epsilon \in (0,1)$ such that $\sup_{z\in\sigma(Q)}|1-(1-z)^{1+\epsilon}| \le 1$. Then $V := I-(I-Q)^{1+\epsilon}$ is a (normal) contraction, by the spectral theorem. Projecting back to the original space we obtain that $S := I-(I-T)^{1+\epsilon}$ is a contraction on $\mathbf{H}$. The functional calculus for the normal operator Q yields $I-Q = (I-V)^{1/(1+\epsilon)}$, so $I-T = (I-S)^{1/(1+\epsilon)}$ yields the assertion. $\square$

Remark. Theorem 7.7 provides an effective way to build Markov operators admitting a *real* normal dilation whose spectrum is included in some Stolz region: take any Markov operator R on $L^2(m)$ and define $P := I-(I-R)^\alpha$ for some $0<\alpha<1$. Then, P is a Markov operator with the desired property, when in the proof of Theorem 7.7 U is the two-sided shift of R.

Proposition 7.8. *Let $P_1, \dots, P_d$ be d commuting contractions of a real space $\mathbf{H}$. Assume that each P_i has a normal dilation with spectrum included in a Stolz region. Then, the product $P_1 \cdots P_d$ admits a normal dilation with spectrum included in a Stolz region, and so does any convex combination $\alpha_1 P_1 + \ldots + \alpha_d P_d$.*

Proof. It is enough to prove the result when $d = 2$ and then proceed by induction.

Let P_1, P_2 be commuting contractions admitting each a normal dilation whose spectrum is included in a Stolz region. By the proof of Theorem 7.7, there exists $\varepsilon_1, \varepsilon_2 > 0$, such that $S_1 := I - (I - P_1)^{1+\varepsilon_1}$ and $S_2 := I - (I - P_2)^{1+\varepsilon_2}$ are contractions. Clearly S_1 and S_2 commute. Hence, by Andô's dilation theorem [3] (see e.g. Theorem 6.1 Chapter I of [49]; the proof is valid also for real spaces), there exist two commuting unitary operators U_1 and U_2 on a larger *real* Hilbert space $\mathbf{K}$ containing $\mathbf{H}$, such that, if E is the orthogonal projection from $\mathbf{K}$ onto $\mathbf{H}$, we have

$$S_1^i S_2^j = E U_1^i U_2^j \qquad \forall i, j \geq 0 .$$

Define $Q_1 := I - (I - U_1)^{1/(1+\varepsilon_1)}$ and $Q_2 := I - (I - U_2)^{1/(1+\varepsilon_2)}$. Then, $Q_1 Q_2$ and $\alpha_1 Q_1 + \alpha_2 Q_2$ are normal, and on the complexification $\mathbf{K}^{\mathbb{C}}$ they have their numerical ranges in Stolz regions. By construction they are respectively dilations of $P_1 P_2$ and $\alpha_1 P_1 + \alpha_2 P_2$, hence the result. $\square$

Remark. We needed to proceed by induction because Andô's theorem does not extend to more than two contractions, by a result of Parrott [47].

8. Limit theorems for functions in $\sqrt{I - P}L^2(m)$ and examples

In this section we combine the results of the previous sections to obtain conditions for our limit theorems to hold for $f \in \sqrt{I - P}L^2(m)$ (where P is a Markov operator as in the previous sections), and provide some examples.

Theorem 8.1. *Let P be a Markov operator with invariant probability m which is ergodic. Assume that P admits a real normal dilation Q with spectrum $\sigma(Q)$ contained in a domain Γ_C. Let $f \in L^2(m)$. If $f \in \sqrt{I - P}L^2(m)$, in particular (see Proposition 9.2 below) if $\sum_{n \geq 1} \|P^n f\|_0^2 < \infty$, then $f \in \mathbf{B}$ and the conclusions of Theorems 4.1 and 5.1 hold.*

For normal operators, we have the following (see Proposition 7.1).

Theorem 8.2. *Let P be a normal Markov operator with invariant probability m which is ergodic. Let $f \in L^2(m)$. If $\int_{\sigma(P)} \frac{1}{1 - \operatorname{Re} z} \rho_f(dz) < \infty$, in particular if $\sum_{n \geq 1} \|P^n f\|_0^2 < \infty$, then $f \in \mathbf{B}$ and the conclusions of Theorems 4.1 and 5.1 hold.*

Corollary 8.3. *Let P be a normal Markov operator with invariant probability m which is ergodic. If P satisfies the sector condition, then Theorem 8.1 applies to P.*

Proof. By (32) and Proposition 6.3, $\sigma(P)$ is contained in some Γ_C. $\square$

Remarks. 1. Under the assumptions on P in Theorem 8.1, we also have the annealed CLT for every $f \in \sqrt{I - P}L^2(m)$, by applying (the discrete version of) Olla's CLT for $f \in \mathbf{B}$ [45, p. 80]. In general, the annealed CLT holds for every $f \in L^2(m)$ satisfying the Maxwell–Woodroofe condition [41]

$$\sum_{n \geq 1} \frac{\| \sum_{k=1}^{n} P^k f \|}{n^{3/2}} < \infty \tag{40}$$

with no additional assumptions on P ergodic. If f satisfies (40), then $f \in \sqrt{I-P}L^2(m)$ [23]. However, even for P symmetric the converse need not be true – see examples in [17]. Note that the CLT may fail for some P and $f \in \sqrt{I-P}L^2(m)$ [50], while the theorem of Gordin and Lifshitz [31] for P normal applies to all functions $f \in \sqrt{I-P}L^2(m)$, even those which do not satisfy (40).

For $\alpha > 1/2$, the rate in the mean ergodic theorem [21, Corollary 2.15] yields that $(I-P)^\alpha L^2(m)$ is contained in the set MW of $f \in L^2(m)$ satisfying (40). If $(I-P)L^2(m)$ is not closed (P is not uniformly ergodic), then there exist functions in $(I-P)^\alpha L^2(m)$ which are not coboundaries [21], so $(I-P)L^2(m)$ is strictly contained in MW; thus in the example of [23] cited above $\mathcal{H}_{-1} = (I-P)L^2(m) \subsetneq MW$, while under the conditions of Theorem 8.1 $MW \subset \mathcal{H}_{-1}$ with no equality in the example of [17]. Hence condition (40) and the condition $f \in \mathcal{H}_{-1}$ are not comparable, and thus Olla's CLT gives some new information for some non-normal P – see the first two examples below.

2. Cuny [15] obtained the law of iterated logarithm (LIL) for any $f \in L^2(m)$ satisfying the Maxwell–Woodroofe condition (40), with no additional assumptions on P. His result supersedes the previous results of [43,53,14], who assumed on f various conditions which imply (40). For P symmetric, Wu [52] obtained the functional LIL for $f \in \sqrt{I-P}L^2(m)$. Theorem 8.1 and Corollary 8.3 apply also to some non-symmetric P and functions which do not satisfy (40). This situation is somewhat similar to the annealed CLT – see the previous remark.

3. As far as we know, the almost sure central limit theorem under projective conditions has been considered only in Merlevède–Peligrad–Peligrad [42]. They worked under a strengthening of the Maxwell–Woodroofe condition.

Example 2. Let S be a Markov operator with m an invariant probability which is ergodic, and let $P = \sum_{n=1}^\infty a_n^{(\alpha)} S^n$, where $\alpha \in (0,1)$ and $(1-t)^\alpha = 1 - \sum_{n=1}^\infty a_n^{(\alpha)} t^n$ for $|t| \le 1$. Then m is P-invariant and also P is ergodic. We have $P = I - (I-S)^\alpha$, and by the remark to Theorem 7.7 P has a real normal dilation Q with $\sigma(Q)$ in a Stolz region. Hence Theorem 8.1 applies. If S is not normal, neither is P.

Example 3. Let $P_1, \dots, P_d$ be commuting ergodic normal Markov operators on $L^2(m)$ which satisfy the sector condition and let $P = \sum_{i=1}^d a_i P_i$ be a convex combination. If $Pf = f$, then uniform convexity yields that $P_i f = f$ for every i, so f is constant; hence P is ergodic. By Corollary 6.4 P satisfies the sector condition. By a Theorem of Fuglede [29] (see also [32]), each P_i commutes also with P_j^*, so P is normal. Hence Corollary 8.3 applies.

Example 4. Let $P_1, \dots, P_d$ be commuting ergodic Markov operators on $L^2(m)$ admitting real normal dilations $N_1, \dots, N_d$ with $\sigma(N_1), \dots, \sigma(N_d)$ contained in Stolz regions. Define $P = P_1 \cdots P_d$ and $R = a_1P_1 + \cdots + a_dP_d$. By Proposition 7.8, P and R are a Markov operators admitting real normal dilations, M and N respectively, with $\sigma(M)$ and $\sigma(N)$ in Stolz regions.

Hence Theorem 8.1 applies to P and to R, provided we show that those operators are ergodic. The fact that R is ergodic follows from uniform convexity. Let us prove that P is ergodic. It suffices to consider the case $d = 2$.

By Theorem 6.8, we have $\sup_n n\|N_1^n - N_1^{n+1}\| < \infty$ and $\sup_n n\|N_2^n - N_2^{n+1}\| < \infty$. Projecting back to the original space, it follows that $\sup_n n\|P_1^n - P_1^{n+1}\| < \infty$ and $\sup_n n\|P_2^n - P_2^{n+1}\| < \infty$. Then by commutativity

$$\sup_n n\|P^n - P^{n+1}\| \le \sup_n n\|P_1^n(P_2^n - P_2^{n+1})\| + \sup_n n\|P_2^{n+1}(P_1^n - P_1^{n+1})\| < \infty.$$

Hence also P has spectrum in a Stolz region, by Theorem 6.8. Denote $Ig = \int g\,dm$ for $g \in L^2(m)$. By ergodicity and the above estimates, $P_1^n g \to Ig$ and $P_2^n g \to Ig$ in norm for every $g \in L^2(m)$. Since constants are invariant, we have

$$\|P^n g - Ig\| = \|P_1^n(P_2^n g - Ig)\| \leq \|P_2^n g - Ig\| \to 0,$$

which shows that P is ergodic. Hence Theorem 8.1 applies to P.

Note that if $P_1, \dots, P_d$ are normal, then each P_i commutes with each P_j^* by [29], so P is normal.

Example 5. Let G be a locally compact Abelian (LCA) σ-compact group G with dual group $\hat{G}$, and let $\{\theta(t) : t \in G\}$ be a continuous probability preserving ergodic action of G in $(\mathbb{S}, \mathcal{S}, m)$. For a probability μ on G define the μ-average of the action by $Pf = \int_G f \circ \theta(t) d\mu(t)$. Then P is a Markov operator preserving m. We assume that the smallest closed group containing the support of μ is G, so P is ergodic (the Choquet–Deny theorem). Since G is commutative, the transformations $\{\theta(t)\}$ induce commuting unitary operators. Hence P is normal. If μ has a *bounded angular ratio*, which means that $|\hat{\mu}(\gamma)| < 1$ for $0 \neq \gamma \in \hat{G}$ and

$$\sup_{0 \neq \gamma} \frac{|1 - \hat{\mu}(\gamma)|}{1 - |\hat{\mu}(\gamma)|} = C < \infty,$$

then $\sigma(P)$ is contained in a Stolz region [11, Proposition 5.3] (for $G = \mathbb{Z}$ see [5]). Hence Theorem 8.1 applies.

9. On sufficient conditions for limit theorems for Markov chains

Let P be an ergodic normal Markov operator on $L^2(m)$. We know that the CLT holds for $f \in \sqrt{I-P}L^2(m)$ [31], and by Lemma 2.4 and Proposition 7.1 $\mathbf{B} = \mathcal{H}_{-1} \subset \sqrt{I-P}L^2(m)$, so the CLT holds for $f \in \mathcal{H}_{-1}$. In Theorem 8.2 we have shown that the WIP as well as the ASIP hold for $f \in \mathcal{H}_{-1}$. A natural question is whether for normal Markov operators those results still hold for $f \in \sqrt{I-P}L^2(m)$. We answer below in the negative.

Proposition 9.1. *Let $0 < \varepsilon < 1$. There exists a normal Markov operator on some $L^2(m)$ and $f \in \sqrt{I-P}L^2(m)$, such that $\max_{1 \le k \le n} |S_k| / \sqrt{n \log^\varepsilon n} \xrightarrow[n\to\infty]{} +\infty$. In particular, the WIP and the ASIP cannot hold (so $f \notin \mathcal{H}_{-1} = \mathbf{B}$).*

Proof. Let (X, Σ, m, θ) be an invertible and ergodic dynamical system. We define a unitary (hence normal) operator P on $L^2(m)$ by setting $Pf := f \circ \theta$ for every $f \in L^2(m)$. Let $(W_n)_{n\in\mathbb{N}}$ be the corresponding canonical Markov chain. It is well-known that for any $f \in L^2(m)$, since P is given by a transformation, the law of the process $(f(W_n)_{n\in\mathbb{N}}$ under $\mathbb{P}_m$ is the same as the law of the process $(f(\circ\theta^n)_{n\in\mathbb{N}}$ under m. It follows from the proof of Theorem 2.5 of [13] that for any $0 < \varepsilon < 1$, that the exists $f \in L^2(m)$, such that

$$\frac{\max_{1\le k\le n} |f + \ldots + f \circ \theta^{n-1}|}{\sqrt{n \log^\varepsilon n}} \xrightarrow[n\to\infty]{} +\infty \qquad m\text{-a.s.},$$

and the first assertion follows. The fact that the ASIP does not hold is obvious. To see that the WIP does not hold, notice that otherwise we should have $\sup_{n\ge 1} \mathbb{P}(\max_{1\le k\le n} |f + \ldots + f \circ \theta^{n-1}| \geq A\sqrt{n}) \xrightarrow[A\to\infty]{} 0$. □

The next proposition was needed in Theorem 8.1.

Proposition 9.2. *If $f \in L^2(m)$ satisfies $\sum_n \|P^n f\|_0^2 < \infty$, then $f \in \sqrt{I-P}L^2(m)$.*

Proof. Recall that P admits a unitary dilation, that is, there exists a Hilbert space $\mathcal{H}$ containing $L^2(m)$ and a unitary operator U on $\mathcal{H}$ such that for every $n \in \mathbb{N}$, and every $f \in L^2(m)$, $EU^n f = P^n f$ where E is the orthogonal projection onto $L^2(m)$.

We first make two observations.

Lemma 9.3. *For every $n \in \mathbb{N}$ and every $\ell \geq 1$, the spaces $U^{-n-\ell}P^{n+\ell}L^2(m)$ and $(U^{-n}P^n - U^{-n-1}P^{n+1})L^2(m)$ are orthogonal (in $\mathcal{H}$).*

Proof. Let $f, g \in L^2(m)$. Let $n \in \mathbb{N}$ and $\ell \geq 1$. We have

$$\begin{aligned}
&\langle (U^{-n}P^n - U^{-n-1}P^{n+1})f, U^{-n-\ell}P^{n+\ell}g\rangle_{\mathcal{H}} \\
&= \langle U^{\ell}P^n f, P^{n+\ell}g\rangle_{\mathcal{H}} - \langle U^{\ell-1}P^{n+1}f, P^{n+\ell}g\rangle_{\mathcal{H}} \\
&= \langle P^{n+\ell}f, P^{n+\ell}g\rangle_{\mathcal{H}} - \langle P^{n+\ell}f, P^{n+\ell}g\rangle_{\mathcal{H}} = 0 . \qquad \square
\end{aligned}$$

Lemma 9.4. *Let $f \in L^2(m)$. The following conditions are equivalent.*

(i) $\sum_{n\in\mathbb{N}} \|P^n f\|_0^2 < \infty$;
(ii) $\|P^n f\|_{\mathcal{H}} \to 0$ *and* $\sum_{n\in\mathbb{N}} n\|U^{-n}P^n f - U^{-n-1}P^{n+1}f\|_{\mathcal{H}}^2 < \infty$.

Proof. Notice that if (i) holds, then $\|P^n f\| \to 0$. Hence, in any case, we may assume that $\|P^n f\| \to 0$. Hence, for every $n \in \mathbb{N}$, we have (with convergence in $\mathcal{H}$)

$$P^n f = \sum_{k\geq 0}(U^{-k}P^{n+k}f - U^{-k-1}P^{n+k+1}f) . \tag{41}$$

By the above lemma the terms of that series lie in orthogonal spaces. Hence,

$$\begin{aligned}
\|P^n f\|_{\mathcal{H}}^2 &= \sum_{k\geq 0} \|U^{-k}P^{n+k}f - U^{-k-1}P^{n+k+1}\|_{\mathcal{H}}^2 \\
&= \sum_{k\geq n} \|U^{-k}P^{k}f - U^{-k-1}P^{k+1}\|_{\mathcal{H}}^2 ,
\end{aligned}$$

where we used that U^{-1} is an isometry (and a change of variable) for the last identity. The desired result, then follows by Fubini. $\square$

Continuation of the proof of Proposition 9.2. Let $f \in L^2(m)$ be such that $\sum_n \|P^n f\|_0^2 < \infty$. By lemma [21], we just have to prove that the series $\sum_n \beta_n P^n$ converges in $L^2(m)$, which is equivalent to convergence of $\sum_n P^n f/\sqrt{n}$ [21]. We shall check Cauchy's criterion.

Let $q > p \geq 1$ be integers. Denote $V_{p,q}f := \sum_{n=p}^{q} \frac{P^n f}{\sqrt{n}}$. Using (41) and that U^{-1} is an isometry, we see that

$$\begin{aligned}
\|V_{p,q}f\|_{\mathcal{H}}^2 &= \sum_{k\geq 0} \|U^{-k}P^{k}V_{p,q}f - U^{-k-1}P^{k+1}f\|_{\mathcal{H}}^2 \\
&\leq \sum_{k\geq 0} \Big(\sum_{n\geq p} \frac{\|U^{-k}P^{n+k}f - U^{-k-1}P^{n+k+1}f\|_{\mathcal{H}}}{\sqrt{n}} \Big) \\
&= \sum_{k\geq 0} \Big(\sum_{n\geq p} \frac{\|U^{-n-k}P^{n+k}f - U^{-n-k-1}P^{n+k+1}f\|_{\mathcal{H}}}{\sqrt{n}} \Big)^2 .
\end{aligned}$$

For every $m \in \mathbb{N}$, denote $u_m := \|U^{-m}P^m f - U^{-m-1}P^{m+1}f\|_{\mathcal{H}}$. By Lebesgue dominated theorem for the counting measure on $\mathbb{N}$, we just have to prove that

$$\sum_{n\ge p} \frac{u_{n+k}}{\sqrt{n}} \xrightarrow[p\to+\infty]{} 0$$

and that

$$\sum_{k\ge 0} \Big(\sum_{n\ge 1} \frac{u_{n+k}}{\sqrt{n}}\Big)^2 < \infty\,. \tag{42}$$

Let $k \ge 0$. By Cauchy–Schwarz we have

$$\Big(\sum_{n\ge p} \frac{u_{n+k}}{\sqrt{n}}\Big)^2 \le \Big(\sum_{n\ge p} \frac{1}{n\sqrt{n+k}}\Big) \sum_{n\ge p} \sqrt{n+k}\,u_{n+k}^2 \le \frac{C}{\sqrt{p}} \sum_{n\ge p+k} \sqrt{n}\,u_n^2 \xrightarrow[p\to+\infty]{} 0\,, \tag{43}$$

by assumption.

Let us prove (42). We first notice that, by Cauchy–Schwarz,

$$\Big(\sum_{n=1}^{k+1} \frac{u_{n+k}}{\sqrt{n}}\Big)^2 \le \Big(\sum_{j=1}^{k+1} \frac{1}{j^{1/2}}\Big) \sum_{n=1}^{k+1} \frac{u_{n+k}^2}{n^{1/2}} \le C(k+1)^{1/2} \sum_{n=1}^{k+1} \frac{u_{n+k}^2}{n^{1/2}}\,.$$

Now, interverting the order of summation twice, we see that

$$\begin{aligned}
\sum_{k\ge 0} (k+1)^{1/2} \sum_{n=1}^{k+1} \frac{u_{n+k}^2}{n^{1/2}} &\le \sum_{k\ge 0} \sum_{n=1}^{k+1} \frac{(n+k+1)^{1/2} u_{n+k}^2}{n^{1/2}} \\
&= \sum_{n\ge 1} \frac{1}{n^{1/2}} \sum_{k\ge n-1} (n+k+1)^{1/2} u_{n+k}^2 = \sum_{n\ge 1} \frac{1}{n^{1/2}} \sum_{k\ge 2n-1} (k+1)^{1/2} u_k^2 \\
&= \sum_{k\ge 1} (k+1)^{1/2} u_k^2 \sum_{1\le n\le (k+1)/2} \frac{1}{n^{1/2}} \le C \sum_{k\ge 1} (k+1) u_k^2 < \infty\,,
\end{aligned}$$

by assumption. Hence

$$\sum_{k\ge 0} \Big(\sum_{n=1}^{k+1} \frac{u_{n+k}}{\sqrt{n}}\Big)^2 < \infty\,. \tag{44}$$

On the other hand, by (43) with $p = k+2$, we see that

$$\sum_{k\ge 0} \Big(\sum_{n\ge k+2} \frac{u_{n+k}}{\sqrt{n}}\Big)^2 \le \sum_{k\ge 0} \frac{C}{\sqrt{k+1}} \sum_{n\ge 2k+2} \sqrt{n}\,u_n^2 \le C' \sum_{n\ge 2} n u_n^2 < \infty\,. \tag{45}$$

Combining (44) and (45) with the identity $(a+b)^2 \le 2(a^2+b^2)$, we see that (42) holds. □

By Proposition 9.2, the condition $\sum_n \|P^n f\|_0^2 < \infty$ is always stronger than the condition $f \in \sqrt{I-P}L^2(m)$. For normal operators, or for operators admitting a normal dilation with spectrum in a sector, the condition $\sum_n \|P^n f\|_0^2 < \infty$ implies that $f \in \mathcal{H}_{-1}$ (see Proposition 7.1 or Proposition 7.4), so for such operators convergence of $\sum_n \|P^n f\|_0^2$ is sufficient for the WIP (and the ASIP).

One may wonder whether this is also true for other operators.

Proposition 9.5. *(See Dedecker [20].)*

(i) *There exists a Markov operator P on some $L^2(m)$ and $f \in L^2(m)$ with $\sum_{n\geq 1}(\log n)\|P^n f\|_0^2 < \infty$, such that $\frac{1}{\sqrt{n}}\|S_n(f)\|_{L^2(\mathbb{P}_m)} \xrightarrow[n\to\infty]{} +\infty$. In particular, by Theorem 3.2, $f \notin \mathcal{H}_{-1}$.*

(ii) *There exists a Markov operator P on some $L^2(m)$ and $f \in L^2(m)$ with $\sum_{n\geq 1}(\log n)\|P^n f\|_0^2 < \infty$, such that $\frac{1}{\sqrt{n}}\|S_n(f)\|_{L^2(\mathbb{P}_m)} \xrightarrow[n\to\infty]{} 1$, but the CLT fails. In particular, f does not satisfy the Maxwell–Woodroofe condition (40).*

Proof. Part (i) of Proposition 9.5 follows from Remark 3 of [20], and part (ii) follows from Corollary 2.4(ii) there. The paper [20] is written in the setting of dynamical systems but, as observed for instance in [19], it has an equivalent formulation in terms of Markov chains. In that case the Markov operator is a co-isometry (P^* is an isometry). □

Remarks. 1. Volný [50] constructed an example of a (non-normal, of course) Markov operator P with $f \in \sqrt{I-P}L^2(m)$ for which the CLT does not hold. In his example the asymptotic variance $\lim_n \|S_n(f)\|_{L^2(\mathbb{P}_m)}/\sqrt{n}$ does not exist. In view of Proposition 9.2, in Dedecker's example given by Proposition 9.5(ii), f is in $\sqrt{I-P}L^2(m)$ and has a finite asymptotic variance, but the CLT still does not hold.

2. The function f in Proposition 9.5(ii) is not in $\mathbf{B}$, since by Olla [45, p. 80] the CLT holds for functions in $\mathbf{B}$. By the next proposition, $f \notin \mathcal{H}_{-1}$.

As we shall see now, for co-isometries the symmetrization method does not bring any novelty.

Proposition 9.6. *Let P be a Markov operator on $L^2(m)$ which is a co-isometry. Then, $\mathcal{H}_{-1} = (I-P)L^2(m) = \mathbf{B}$.*

Proof. We already know that $(I-P)L^2(m) \subset \mathbf{B} \subset \mathcal{H}_{-1}$. Let $f \in \mathcal{H}_{-1}$. We have to prove that $f \in (I-P)L^2(m)$. By a result of Browder [8], it suffices to prove that

$$\sup_{n\geq 1}\|f + Pf + \ldots + P^n f\|_0 < \infty\,.$$

Define $V_n := I + P + \ldots + P^n$ and $f_n := V_n f$. Let $g \in L^2(m)$ be such that $f = \sqrt{I-P_s}g$. Using (7), we obtain

$$\begin{aligned}\|f_n\|_0^4 &= \langle V_n\sqrt{I-P_s}g, f_n\rangle_0^2 \leq \|g\|_0^2\,\|\sqrt{I-P_s}V_n^* f_n\|_0^2 \\ &= \|g\|_0^2\,\langle (I-P)^* V_n^* f_n, V_n^* f_n\rangle_0 = \|g\|_0^2\,\langle f_n - (P^*)^{n+1}f_n, V_n^* f_n\rangle_0 \\ &= \|g\|_0^2\,\langle f_n, (I-P^{n+1})V_n^* f_n\rangle_0\,.\end{aligned}$$

By the assumption $PP^* = I$ we have $P^{n+1}V_n^* = V_{n+1} - I$, hence (on the real L^2)

$$|\langle f_n, (I-P^{n+1})V_n^* f_n\rangle_0| = |\langle f_n, V_n^* f_n\rangle_0 - \langle f_n, (V_{n+1}-I)f_n\rangle_0|$$

$$|\langle f_n, V_n^* f_n\rangle_0 - \langle f_n, (V_{n+1}^*-I)f_n\rangle_0| \leq 2\|f_n\|_0^2\,.$$

Hence, $\|f_n\|_0^4 \leq 2\|g\|_0^2\,\|f_n\|_0^2$, and the result follows. □

10. Poincaré's inequality and uniform ergodicity

In this section we use the properties of $\mathcal{H}_{-1}$ to show that P (ergodic with invariant probability m) is uniformly ergodic if and only if P satisfies *Poincaré's inequality*: there exists $C > 1$ such that

$$\|f\|_0^2 \leq C\langle f, (I-P)f\rangle_0 \qquad \forall \text{ real } \; f \in L_0^2(m). \tag{46}$$

Remark. Originally, Poincaré's inequality was with $I - P$ replaced by $-\Delta$, where Δ is the Laplacian (and the functions f taken in the domain of Δ), as a tool for estimating the eigenvalues of Δ. For the history of the inequality we refer to Allaire [2]. The book [4] treats the inequality with Δ replaced by a general symmetric infinitesimal generator A of a (necessarily reversible) Markov semi-group.

Lemma 10.1. *Let P be an ergodic Markov operator with invariant probability m. Then the following are equivalent.*

(i) *P satisfies Poincaré's inequality (46);*
(ii) *$\mathcal{H}_{-1} = L_0^2(m)$.*

Moreover, if any of the above holds, then P is power-bounded on $\mathcal{H}_{-1}$, with $\sup_n \|P^n\|_{-1} \leq 2\sqrt{C}$. In particular $\mathbf{B} = \mathcal{H}_{-1} = L_0^2(m)$.

Proof. Assume (i). Let $f, h \in L_0^2(m)$. By (46) we have

$$\langle f, h\rangle^2 \leq \|f\|_0^2 \, \|h\|_0^2 \leq C\|f\|_0^2 \langle h, (I-P)h\rangle \, .$$

Hence by Lemma 2.3(i), $f \in \mathcal{H}_{-1}$ and $\|f\|_{-1} \leq \sqrt{C}\|f\|_0$. Hence, (ii) holds and the norms $\|\cdot\|_0$ and $\|\cdot\|_{-1}$ are equivalent. In particular P is power bounded on $\mathcal{H}_{-1}$. By (5), $\|P^n\|_{-1} \leq 2\sqrt{C}$. For the equality $\mathbf{B} = \mathcal{H}_{-1}$ see end of the proof of Lemma 2.4.

Assume that $\mathcal{H}_{-1} = L_0^2(m)$. By Lemma 2.2 it follows that $\sqrt{I - P_s}$ is a bijection of $L_0^2(m)$. By the open mapping theorem it admits a continuous (with respect to $\|\cdot\|_0$) inverse, say G_s. Then for $f \in L_0^2(m)$ we have

$$\|f\|_0^2 = \|G_s\sqrt{I-P_s}f\|_0^2 \leq \|G_s\|^2 \, \langle f, (I-P_s)f\rangle_0 = \|G_s\|^2 \, \langle f, (I-P)f\rangle_0 \, ,$$

i.e. P satisfies Poincaré's inequality. □

For $p \in [1, \infty)$, the Markov operator P is called *L^p-uniformly ergodic* if $\frac{1}{n}\sum_{k=1}^n P^k$ converges in the operator norm on L^p. If P is L^2-uniformly ergodic, then it is L^p-uniformly ergodic for any $1 < p < \infty$ (by the Riesz–Thorin theorem; see also [24, Corollary 3.5]).

Theorem 10.2. *Let P be an ergodic Markov operator with invariant probability m. Then P is L^2-uniformly ergodic if and only if it satisfies Poincaré's inequality.*

Proof. Assume first the P satisfies Poincaré's inequality. By Lemma 10.1(ii), $\mathcal{H}_{-1} = L_0^2(m)$. Then the inequality (21) in Theorem 3.2 yields that

$$\sup_{n\geq 1} \frac{1}{n}\|\sum_{k=1}^n f(W_k)\|_{L^2(\mathbb{P}_m)} < \infty \qquad \text{for every } \; f \in L_0^2(m).$$

Thus condition (iii) of [24, Theorem 2.2] holds, so P is uniformly ergodic by that theorem.

Assume now that P is uniformly ergodic. Then $I - P$ is invertible on $L_0^2(m)$ (since $\|\frac{1}{n}\sum_{k=1}^n P^k\|_{L_0^2(m)} \to 0$), and $(I-P)L^2(m) = (I-P)L_0^2(m) = L_0^2(m)$. Using Lemma 2.3(ii) we obtain

$$\mathcal{H}_{-1} \subset L_0^2(m) = (I-P)L_0^2(m) \subset \mathcal{H}_{-1},$$

which yields $\mathcal{H}_{-1} = L_0^2(m)$. By Lemma 10.1 P satisfies Poincaré's inequality. □

Remark. For P *symmetric* (and ergodic), it has been known for some time that Poincaré's inequality (46) is equivalent to P on the complex $L^2(m)$ having a spectral gap at 1, i.e. $\sigma(P) \subset [-1,\theta] \cup \{1\}$ for some $0 < \theta < 1$. Since we have not found a reference for the proof, we give one which is independent of our previous results. Let L_0^2 be the subspace of the complex $L^2(m)$ of functions with integral zero, and let T be the restriction of P to L_0^2. Since P is symmetric $\langle Pf, f\rangle_0$ is real on the complex $L^2(m)$, and it is easy to check that (46) implies the same inequality also for every *complex* valued $f \in L^2(m)$ with $\int f\,dm = 0$. Then (46) means

$$\sup\{\langle Tf, f\rangle_0 : f \in L_0^2,\ \|f|_0 = 1\} \le 1 - \frac{1}{C}.$$

Since T is a symmetric contraction on L_0^2, this is equivalent, see [48, Section 107], to $\sigma(T) \subset [-1, 1 - \frac{1}{C}]$ which, by ergodicity, is equivalent to P having a spectral gap. In particular $I - T$ is invertible on L_0^2, so the ergodic decomposition yields the uniform ergodicity of P.

Corollary 10.3. *Let P be an ergodic Markov operator with invariant probability m. Then P is L^2-uniformly ergodic if and only if P_s is L^2-uniformly ergodic.*

Proof. By (7), Poincaré's inequality is the same for P and for P_s. □

Remark. In continuous time, we have a C_0-semi-group $\{P_t\}_{t\ge 0}$ of Markov operators with infinitesimal generator A, and the quadratic form $\langle f, (I-P)f\rangle_0$ in (46) is replaced by the Dirichlet quadratic form $\langle f, -Af\rangle_0$, defined for real f in the domain of A; Ligget [39, Theorem 2.3] proved that Poincaré's inequality is equivalent in that case to L^2-operator norm convergence of P_t, as $t \to \infty$, with exponential rate (without averaging). In the discrete case, uniform ergodicity still allows for periods, even when P is symmetric (take any finite state reversible ergodic Markov chain which is 2-periodic; see [24, Proposition 3.3]).

Proposition 10.4. *Let P be an ergodic Markov operator with invariant probability m. Then the following are equivalent.*

(i) *P satisfies Poincaré's inequality;*
(ii) *P is L^2-uniformly ergodic;*
(iii) *There exists $d \ge 1$ such that $\{P^{dn}\}_{n\ge 1}$ converges in L^2 operator norm.*

In particular, with the same $d \ge 1$ as in (iii)*, there exist $C > 0$ and $0 < \rho < 1$, such that $\|P^{dn}f - E_d(f)\|_0 \le C\rho^n\|f\|_0$, for every $f \in L^2(m)$, where E_d is the orthogonal projection on the functions invariant by P^d.*

Proof. The equivalence of (i) and (ii) follows from Theorem 10.2. The exponential convergence follows easily from (iii).

Let us prove that (ii) is equivalent to (iii).

If P is L^2-uniformly ergodic, the averages converge also uniformly on the complex $L^2(m)$. As shown in the proof of [24, Proposition 3.3], all unimodular spectral points of P are simple poles (hence isolated in the spectrum), so there are only a finite number of them. Since P is a positive operator, they are roots of unity, and there is $d \geq 1$ such that $\sigma(P^d) \cap \{\lambda \in \mathbb{C} : |\lambda| = 1\} = \{1\}$, by [40, Proposition 1]. Hence the spectral radius of P^d restricted to $\overline{(I - P^d)L^2}$ is less than one, which implies that $\{P^{dn}\}_{n\geq 1}$ converges in operator norm.

The converse follows easily from the operator norm convergence of $\frac{1}{n}\sum_{k=1}^{n} P^{dk+j}$, as $n \to \infty$, for each $0 \leq j \leq d-1$. $\square$

The smallest d for which (iii) of Proposition 10.4 holds is called the *period* of P; when $d = 1$ we say that P is *aperiodic*.

Remark. As mentioned above, the analogue of this proposition in continuous-time, the equivalence of Poincaré's inequality and the exponential L^2-operator norm convergence of P_t, is well-known [39]. We know of no reference for our discrete time result.

Lemma 10.5. *Let P be an ergodic Markov operator with invariant probability m. Assume that P satisfies*

$$\langle Pf, f\rangle_0 > -1/2 \qquad \textit{for every real } f \in L^2(m) \textit{ with } \|f\|_0 = 1\,. \tag{47}$$

Then P (on the complex $L^2(m)$) has no unimodular eigenvalue other than 1. In particular, if P is in addition uniformly ergodic, it is aperiodic.

Proof. Since P is an ergodic Markov operator, it is well known that the set of unimodular eigenvalues of P is a group (see e.g. [1] or [12]). Assume that P on the complex L^2 has a unimodular eigenvalue $e^{i\theta}$ with eigenfunction $f \neq 0$. Equality in the Cauchy–Schwarz inequality implies that $P^*f = e^{-1\theta}f$. Hence $P_s f = \cos\theta \cdot f$.

If $\mathrm{e}^{i\theta}$ is not a root of unity, it generates a dense subgroup of the circle, contradicting (47).

If $\mathrm{e}^{i\theta}$ is a d-th root of unity with (minimal) $d \geq 2$, then for $k \in \{0, \ldots, d-1\}$, the numbers $\mathrm{e}^{2ik\pi/d}$ are eigenvalues. If $d = 2s$ for some $s \geq 1$, taking $k = s$, we see that -1 is an eigenvalue of P, contradicting (47). If $d = 2s+1$ for some $s \geq 1$, taking $k = s$, we see that $\cos(2k\pi/d) = -\cos(\pi/(2s+1)) \leq -\cos(\pi/3) = -1/2$, contradicting (47). $\square$

Our next task is to estimate the exponent ρ in Proposition 10.4(iii) when P is L^2-uniformly ergodic and aperiodic.

When P is symmetric positive definite, i.e. satisfies (47), the proof of uniform ergodicity in the remark preceding Corollary 10.3 yields that C_0, the smallest C in Poincaré's inequality (46), satisfies $\rho = 1 - \frac{1}{C_0}$.

Next, observe that if we have Poincaré's inequality for the *multiplicative symmetrization*, i.e.

$$\|f\|_0^2 \leq K\langle f, I - P^*Pf\rangle_0 \qquad \text{for every real} \quad f \in L_0^2(m),$$

then $\|Pf\|_0 \leq \sqrt{(K-1)/K}\|f\|_0$, so P is aperiodic with $\rho \leq \sqrt{(K-1)/K}$.

The next result is inspired by the work of Fill [28] on Markov chains with *finite* state space.

Proposition 10.6. *Let P be an ergodic Markov operator with invariant probability m, which satisfies Poincaré's inequality (46) with constant $C > 1$, and assume that for some $\gamma > (C-1)/C$ we have*

$$\langle Pf, f\rangle_0 \geq \gamma\langle Pf, Pf\rangle_0 \qquad \forall f \in L^2(m). \tag{48}$$

Then, for every $f \in L_0^2(m)$

$$\|P^n f\|_0^2 \leq \left(\frac{C-1}{\gamma C}\right)^n \|f\|_0^2 \qquad \forall n \geq 1\,.$$

Proof. Let $f \in L^2_0(m)$. Using (46) and (48) we have

$$\|f\|_0^2 \le C\|f\|_0^2 - \gamma C\|Pf\|_0^2 .$$

Hence,

$$\|Pf\|_0^2 \le \frac{C-1}{\gamma C}\|f\|_0^2 ,$$

and the result follows by induction. □

Remarks. 1. Since $P1 = 1$, necessarily $\gamma \le 1$.

2. Let C_0 be the smallest constant for which (46) is satisfied, and put $\rho_s = (C_0-1)/C_0$. If (48) is satisfied, then P_s is positive-definite, so it is aperiodic and ρ_s yields its rate of convergence. Minimality of C_0 yields that $\gamma > \rho_s$, so the proposition yields the estimate $\rho \le \rho_s/\gamma$.

3. The hypothesis (48) may be rewritten as $\langle P_s f, f\rangle_0 \ge \gamma\langle P^*Pf, f\rangle_0$.

Example 6. Let Q be an ergodic L^2-uniformly ergodic Markov operator with invariant probability m, and define $P = \frac{1}{2}(I+Q)$. By Theorem 10.2 Q satisfies Poincaré's inequality with a constant C, and it is easily checked that P satisfies it with a constant $2C$; hence P is L^2-uniformly ergodic.

$$4\|Pf\|_0^2 = \|f\|_0^2 + \|Qf\|_0^2 + 2\langle Qf, f\rangle_0 \le 2\big(\|f\|_0^2 + \langle Qf, f\rangle_0\big) = 4\langle Pf, f\rangle_0 \tag{49}$$

shows that P satisfies (48) with $\gamma = 1$, so by the above $\rho \le \rho_s$.

Note that if we define $P = \alpha I + (1-\alpha)Q$ with $\alpha \in (\frac{1}{2}, 1)$, we can reduce it to $P = \frac{1}{2}(I+Q')$ for some other Markovian contraction Q', so P satisfies (48) with $\gamma = 1$.

A special case is when P is a finite-dimensional ergodic Markov matrix with $p_{i,i} \ge \frac{1}{2}$ for every i (called *strongly aperiodic* in [28]). $Q = 2P - I$ is then Markovian since it is positive and $Q1 = 1$.

Remarks. 1. Let P satisfy (48) with $\gamma = 1$, and define $Q = 2P - I$. Then the computation in (49) shows that necessarily $\|Q\| \le 1$. If P satisfies Poincaré's inequality, so does Q. Since $I - Q = 2(I - P)$, we have $(I-Q)L^2(m)$ closed, so Q is L^2-uniformly ergodic, and $Qf = f$ implies f constant by ergodicity of P.

2. If Q is as in Example 6, but there is some f with $\langle Qf, f\rangle_0 < 0$ (i.e. Q_s is not positive definite), then for $\alpha > 0$ small enough $P^{(\alpha)} := \alpha I + (1-\alpha)Q$ does not satisfy (48), nor even (47), since $\lim_{\alpha\to 0^+}\langle P^{(\alpha)}f, f\rangle_0 < 0$. When Q on $\{0,1\}$ is given by $q_{i,j} = 1 - \delta_{i,j}$, then $P^{(\alpha)}$ does not satisfy (47) for any $\alpha < \frac{1}{2}$.

3. The operator $P = \alpha I + (1-\alpha)Q$ with Q as in Example 6 and $\alpha \in (0,1)$ is always L^2-uniformly ergodic aperiodic. Since $I - P = (1-\alpha)(I-Q)$, we have $(I-P)L^2$ closed, and invertible on L^2_0. By Foguel–Weiss $\|P^n(I-P)\| \to 0$, so $\|P^n\|_{L^2_0} \to 0$, which yields aperiodicity.

4. A Markov operator P has the representation $P = \alpha I + (1-\alpha)Q$ with Q Markovian and $\alpha \in (0,1)$ if and only if $Pf \ge \alpha f$ for every $f \ge 0$. Indeed, if $Pf \ge \alpha f$ for every $f \ge 0$, we define $Q := (P - \alpha I)/(1-\alpha)$. Then $Qf \ge 0$ for $f \ge 0$, $Q1 = 1$, and $Q(x, A) := Q1_A(x)$ is a transition probability with $mQ = m$. Hence Q is a contraction of $L^2(m)$. The converse is obvious.

Acknowledgments

Part of this work was carried out during the second author's visits to Laboratoire MAS of Ecole Centrale de Paris.

The authors are grateful to P. Cattiaux, D. Chafaï and L. Miclo for their comments on Section 10 and for pointing out references [4] and [28].

References

[1] J. Aaronson, M. Lin, B. Weiss, Mixing properties of Markov operators and ergodic transformations, and ergodicity of Cartesian products, Israel J. Math. 33 (3–4) (1979) 198–224.
[2] G. Allaire, A la recherche de l'inégalité perdue, Matapli 98 (2012) 52–64.
[3] T. Andô, On a pair of commutative contractions, Acta Sci. Math. (Szeged) 24 (1963) 88–90.
[4] D. Bakry, I. Gentil, M. Ledoux, Analysis and Geometry of Markov Diffusion Operators, Springer Cham, Berlin–Heidelberg, 2014.
[5] A. Bellow, R. Jones, J. Rosenblatt, Almost everywhere convergence of powers, in: Almost Everywhere Convergence, Academic Press, Boston–San Diego, 1989, pp. 99–120.
[6] S.K. Berberian, The numerical range of a normal operator, Duke Math. J. 31 (1964) 479–483.
[7] A.N. Borodin, I.A. Ibragimov, Limit Theorems for Functionals of Random Walks, Proc. Steklov Inst. Math., vol. 195, 1994 (in Russian); English translation: Amer. Math. Soc., 1995.
[8] F.E. Browder, On the iteration of transformations in noncompact minimal dynamical systems, Proc. Amer. Math. Soc. 9 (1958) 773–780.
[9] P. Cattiaux, D. Chafaï, A. Guillin, Central limit theorems for additive functionals of ergodic Markov diffusion processes, ALEA Lat. Am. J. Probab. Math. Stat. 9 (2) (2012) 337–382.
[10] J.-R. Chazottes, S. Gouëzel, On almost-sure versions of classical limit theorems for dynamical systems, Probab. Theory Related Fields 138 (1–2) (2007) 195–234.
[11] G. Cohen, C. Cuny, M. Lin, Almost everywhere convergence of powers of some positive L_p contractions, J. Math. Anal. Appl. 420 (2014) 1129–1153.
[12] C. Cuny, Weak mixing of random walks on groups, J. Theoret. Probab. 16 (2003) 923–933.
[13] C. Cuny, Some optimal pointwise ergodic theorems with rate, C. R. Math. Acad. Sci. Paris 347 (2009) 953–958.
[14] C. Cuny, Pointwise ergodic theorems with rate with applications to limit theorems for stationary processes, Stoch. Dyn. 11 (2011) 135–155.
[15] C. Cuny, Limit theorems under the Maxwell–Woodroofe condition in Banach spaces, arXiv:1403.0772 (29 pp.).
[16] C. Cuny, ASIP for martingales in 2-smooth Banach spaces. Applications to stationary processes, Bernoulli 21 (1) (2015) 374–400.
[17] C. Cuny, M. Lin, Pointwise ergodic theorems with rate and application to the CLT for Markov chains, Ann. Inst. Henri Poincaré Probab. Stat. 45 (3) (2009) 710–733.
[18] C. Cuny, F. Merlevède, Strong invariance principles with rate for "reverse" martingales and applications, J. Theoret. Probab. 28 (2015) 137–183.
[19] C. Cuny, D. Volný, A quenched invariance principle for stationary processes, ALEA Lat. Am. J. Probab. Math. Stat. 10 (1) (2013) 107–115.
[20] J. Dedecker, On the optimality of McLeish's conditions for the central limit theorem, C. R. Math. Acad. Sci. Paris 353 (2015) 557–561.
[21] Y. Derriennic, M. Lin, Fractional Poisson equations and ergodic theorems for fractional coboundaries, Israel J. Math. 123 (2001) 93–130.
[22] Y. Derriennic, M. Lin, The central limit theorem for Markov chains with normal transition operators started at a point, Probab. Theory Related Fields 119 (4) (2001) 508–528.
[23] Y. Derriennic, M. Lin, The central limit theorem for random walks on orbits of probability preserving transformations, in: Topics in Harmonic Analysis and Ergodic Theory, in: Contemp. Math., vol. 444, Amer. Math. Soc., Providence, RI, 2007, pp. 31–51.
[24] Y. Derriennic, M. Lin, Variance bounding Markov chains, L_2-uniform ergodicity and the CLT, Stoch. Dyn. 11 (2011) 81–94.
[25] N. Dungey, A class of contractions in Hilbert space and applications, Bull. Pol. Acad. Sci. Math. 55 (2007) 347–355.
[26] N. Dungey, Subordinated discrete semigroups of operators, Trans. Amer. Math. Soc. 363 (2011) 1721–1741.
[27] R. Emilion, Mean-bounded operators and mean ergodic theorems, J. Funct. Anal. 61 (1) (1985) 1–14.
[28] J. Fill, Eigenvalue bounds on convergence to stationarity for non-reversible Markov chains, with an application to the exclusion process, Ann. Appl. Probab. 1 (1991) 62–87.
[29] B. Fuglede, A commutativity theorem for normal operators, Proc. Natl. Acad. Sci. USA 36 (1950) 35–40.
[30] M.I. Gordin, B.A. Lifshitz, Central limit theorem for stationary Markov processes, Dokl. Akad. Nauk SSSR 239 (1978) 766–767 (in Russian); English translation: Sov. Math., Dokl. 19 (1978) 392–394.
[31] M.I. Gordin, B.A. Lifshitz, A remark about a Markov process with normal transition operator, in: Third Vilnius Conf. Probab. Stat., vol. 1, Akad. Nauk Litov., Vilnius, 1981, pp. 147–148 (in Russian).
[32] P. Halmos, Commutativity and spectral properties of normal operators, Acta Sci. Math. (Szeged) 12 (1950) 153–156.
[33] E. Hille, R.S. Phillips, Functional Analysis and Semi-groups, rev. ed., Amer. Math. Soc. Colloq. Publ., vol. 31, American Mathematical Society, Providence, RI, 1957.
[34] Y. Jiang, L. Wu, Hilbertian invariance principle for empirical process associated with a Markov process, Chin. Ann. Math. Ser. B 24 (2003) 1–16.
[35] C. Kipnis, 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.
[36] T. Komorowski, C. Landim, S. Olla, Fluctuations in Markov Processes. Time Symmetry and Martingale Approximation, Grundlehren Math. Wiss. (Fundam. Princ. Math. Sci.), vol. 345, Springer, Heidelberg, 2012.
[37] U. Krengel, Ergodic Theorems, De Gruyter, Berlin, 1985.
[38] M.A. Lifshits, Almost sure limit theorem for martingales, in: Limit Theorems in Probability and Statistics, vol. II, Balatonlelle, 1999, János Bolyai Math. Soc., Budapest, 2002, pp. 367–390.

[39] T. Ligget, Exponential L_2 convergence of attractive reversible nearest particle systems, Ann. Probab. 17 (1989) 403–432.
[40] M. Lin, The uniform zero-two law for positive operators in Banach lattices, Studia Math. 131 (1998) 149–153.
[41] M. Maxwell, M. Woodroofe, Central limit theorem for additive functionals of Markov chains, Ann. Probab. 28 (2000) 713–724.
[42] F. Merlevède, C. Peligrad, M. Peligrad, Almost sure invariance principles via martingale approximations, Stochastic Process. Appl. 122 (2012) 170–190.
[43] Y. Miao, G. Yang, The law of the iterated logarithm for additive functionals of Markov chains, Statist. Probab. Lett. 78 (2008) 265–270.
[44] B. Nagy, J. Zemánek, A resolvent condition implying power boundedness, Studia Math. 134 (1999) 143–151.
[45] S. Olla, Central limit theorems for tagged particles and for diffusions in random environment, in: Milieux aléatoires, in: Panor. Synthèses, vol. 12, Soc. Math. France, Paris, 2001, pp. 75–100.
[46] H. Osada, T. Saitoh, An invariance principle for non-symmetric Markov processes and reflecting diffusions in random domains, Probab. Theory Related Fields 101 (1995) 45–63.
[47] S. Parrott, Unitary dilations for commuting contractions, Pacific J. Math. 34 (1970) 481–490.
[48] F. Riesz, B. Sz. Nagy, Leçons d'Analyse Fonctionnelle, 3rd ed., Akadémiai Kiadó, Budapest, 1955.
[49] B. Sz.-Nagy, C. Foiaş, Harmonic Analysis of Operators on Hilbert Space, Translated from French and revised, North-Holland Publishing Co., American Elsevier Publishing Co., Inc., Akadémiai Kiadó, Amsterdam–London, New York, Budapest, 1970.
[50] D. Volný, Martingale approximation and optimality of some conditions for the central limit theorem, J. Theoret. Probab. 23 (2010) 888–903.
[51] L. Wu, Forward–backward martingale decomposition and compactness results for additive functionals of stationary Markov processes, Ann. Inst. Henri Poincaré 35 (1999) 121–141.
[52] L. Wu, Functional law of iterated logarithm for additive functionals of reversible Markov processes, Acta Math. Appl. Sin. 16 (2) (2000) 149–161.
[53] O. Zhao, M. Woodroofe, Law of the iterated logarithm for stationary processes, Ann. Probab. 36 (2008) 127–142.