A BERRY-ESSEEN THEOREM AND EDGEWORTH EXPANSIONS FOR UNIFORMLY ELLIPTIC INHOMOGENEOUS MARKOV CHAINS

DMITRY DOLGOPYAT AND YEOR HAFOUTA

ABSTRACT. We prove a Berry-Esseen theorem and Edgeworth expansions for partial sums of the form $S_N=\sum_{n=1}^{N}f_n(X_n,X_{n+1})$, where $\{X_n\}$ is a uniformly elliptic inhomogeneous Markov chain and $\{f_n\}$ is a sequence of uniformly bounded functions. The Berry-Esseen theorem holds without additional assumptions, while expansions of order 1 hold when $\{f_n\}$ is irreducible, which is an optimal condition. For higher order expansions, we then focus on two situations. The first is when the essential supremum of f_n is of order $O(n^{-\beta})$ for some $\beta\in(0,1/2)$. In this case it turns out that expansions of any order $r<\frac{1}{1-2\beta}$ hold, and this condition is optimal. The second case is uniformly elliptic chains on a compact Riemannian manifold. When f_n are uniformly Lipschitz continuous we show that S_N admits expansions of all orders. When f_n are uniformly Hölder continuous with some exponent $\alpha\in(0,1)$, we show that S_N admits expansions of all orders $r<\frac{1+\alpha}{1-\alpha}$. For Hölder continues functions with $\alpha<1$ our results are new also for uniformly elliptic homogeneous Markov chains and a single functional $f=f_n$. In fact, we show that the condition $r<\frac{1+\alpha}{1-\alpha}$ is optimal even in the homogeneous case.

1. Introduction

Let $Y_1,Y_2,Y_3,...$ be a uniformly bounded sequences of independent random variables. Set $\bar S_N=\sum_{n=1}^{N}(Y_n-\mathbb{E}(Y_n))$, $V_N=\mathrm{Var}(S_N)$ and $\sigma_N=\sqrt{V_N}$. The classical central limit theorem (CLT) states that if $\sigma_N\to\infty$ then, as $N\to\infty$, the distribution of $\hat S_N=\bar S_N/\sigma_N$ converges to the standard normal distribution. A related classical result is the Berry-Esseen theorem [27] which is a quantification of the CLT stating that there is an absolute constant $C_0>0$ so that for every $N\geq 1$

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(\hat S_N\leq t)-\Phi(t)\right|\leq C_0\sigma_N^{-3}\sum_{j=1}^{N}\mathbb{E}\big[\big|Y_j-\mathbb{E}[Y_j]\big|^3\big] \tag{1.1}$$

where Φ is the standard normal distribution function (we refer to [6] for similar result obtained simultaneously). In [28], Esseen proved, in particular, that the optimal constant C_0 in the RHS of (1.1) is greater than 0.4. Since then there were many efforts to provide close to tight upper bounds on C_0, and currently the smallest possible known

Date: November 9, 2021.
DD was partially supported by the NSF and the BSF.

choice for C_0 is $C_0 = 0.56$, see [52] and references therein. In particular, when Y_n are uniformly bounded then with $\|Y\|_\infty = \sup_n \|Y_n\|_\infty$ we have

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(\hat{S}_N \le t) - \Phi(t)\right| \le C_0 \|Y\|_\infty \sigma_N^{-1}. \tag{1.2}$$

It turns out that the rate of σ_N^{-1} in (1.2) is optimal, see below. By now the optimal convergence rate in the CLT was obtained for wide classes of stationary Markov chains [49, 50, 42] and other weakly dependent random processes including chaotic dynamical systems [56, 37, 42, 35, 45, 46], uniformly bounded stationary sufficiently fast ϕ-mixing sequences [54], U-statistics [10, 34] and locally dependent random variables [2, 4, 12] (the last three papers use Stein's method).

The rate σ_N^{-1} is optimal for two reasons. First, for the lattice random variables the distribution function $t \mapsto \mathbb{P}(\hat{S}_N \le t)$ has jumps of order σ_N^{-1}. Secondly even if the distributions of the summands have smooth densities the rate of convergence is still $O\left(\sigma_N^{-1}\right)$ if the third moment of the sum is different from Gaussian. To address the moment obstacle one could introduce appropriate corrections[1]. Namely, fix $r \ge 1$. We say that the *Edgeworth expansions of order* r hold if there are polynomials $P_{1,N}, ..., P_{r,N}$ with degrees not depending on N and coefficients uniformly bounded in N so that

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(\hat{S}_N \le t) - \Phi(t) - \sum_{j=1}^{r} \sigma_N^{-j} P_{j,N}(t)\phi(t)\right| = o(\sigma_N^{-r}) \tag{1.3}$$

where $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ is the standard normal density function. These expansions provide a more accurate approximations of the distribution function of $\hat{S}_N$ in comparison with the Berry-Esseen theorem.

For independent random variables it was proven by Esseen in [27], that the expansion of order 1 holds iff the distribution of S_N is non-lattice. The conditions for higher order expansions are not yet completely understood. Sufficient conditions for the Edgeworth expansions of an arbitrary order were first obtained in [14] under the assumption that the characteristic function of the sum $\mathbb{E}(e^{itS_N})$ decays exponentially in N uniformly for large t. Later the same expansions were obtained in [27, 29, 7, 9, 1] under weaker decay conditions[2], where the second paper considered non identically distributed variables and the fourth and fifth considered random iid vectors. Later Edgeworth expansions were proven for several classes of weakly dependent random variables including stationary Markov chains ([49, 50, 30]), chaotic dynamical systems ([13, 30, 31]) and certain classes of local statistics ([8, 41, 5, 11]). In particular, Hervé-Pène proved in [43] that for several classes of stationary processes the first order Edgeworth expansion holds if the system is irreducible, in the sense that S_N can not be represented as $S'_N + H_N$ where S'_N is lattice valued and H_N is bounded. We also mention that in [3, 55] so called weak expansions, i.e. expansions of the form $\mathbb{E}\left(\phi\left(S_N/\sigma_N\right)\right)$ where ϕ is a smooth test function were studied.

[1] In the case the arithmeticity obstacle is present, that is, the distribution is lattice, one can consider asymptotic expansions of $\mathbb{P}(S_N = k)$ see [29, 33, 44, 19] and references wherein.

[2] The decay conditions used in the above papers are optimal, since one can provide examples where the decay is slightly weaker and there are oscillatory corrections to Edgeworth expansion, see [18, 19].

Both Berry–Esseen Theorem and Edgeworth expansions require a detailed control of the characteristic function. For dependent variables, the most powerful method for analyzing the characteristic function is the spectral approach developed by Nagaev [49, 50] (see [36, 42] for the detailed exposition of the spectral method). Since the spectral method relies on perturbation theory for the spectrum of linear operators, extending it to a non stationary setting turned out to be a non trivial task. Recently a significant progress on this problem was achieved by using a contraction properties of the projective metric which allows to prove spectral gap type estimates for the non-stationary compositions of linear operators ([48, 58, 25, 26]). In particular, complex sequential Ruelle-Perron-Frobenius Theorem, proven in [40] provides a powerful tool for proving the Central Limit Theorem and its extensions in the non stationary case. This theorem allows to obtain both Berry–Esseen theorem ([40, 39]) and Edgeworth expansions ([38, 24]) in the non stationary setting for both Markov chains and dynamical systems.

However, the results of [40, 39, 38, 24] are in a certain sense perturbative. Namely, those papers study either a small perturbation of a fixed stationary system, or they deal with random systems assuming that a system comes to a small neighborhood of a fixed system with a positive frequency. One difficulty in studying the non-stationary case is that there could be large cancellations of the consecutive terms, so that the variance of the sum, can be much smaller then the sum of the variances of the summands. Recently [20] developed a structure theory for Markov chains which allows to find, for each additive functional, a representative in the same homology class (the homologous functionals satisfy the same limit theorems) with the smallest L^2 distance from either zero or from a given lattice in $\mathbb{R}$. This structure theory was used in [20] to prove the local limit theorem for non-stationary Markov chains in both diffusive and large deviations regimes.

In the present paper we combine the methods of [40] and [20] to obtain several optimal results concerning the convergence rate in the CLT for bounded additive functional of uniformly elliptic non-stationary Markov chains. Our results include

- Berry–Esseen bound, which holds without any additional assumptions;
- first order Edgeworth expansion in the irreducible case, extending theorems of Esseen and of Hervé-Pène;
- higher order expansions for the chains with either decaying L^∞ norm or with bounded Hölder norm.

We emphasize that our assumptions concern only regularity of the observables. No additional assumptions dealing with either the growth of variance or with the decay of characteristic function away from zero are made.

The structure of the paper is the following. Section 2 contains the precise statements of our results. The necessary background from [40, 20] is given in Section 3. In Section 4 we discuss the Edgeworth expansions. In general, those expansions follow from the asymptotics of the characteristic function around 0, together with decay of the characteristic functions over appropriate domains. In Section 4 we will show that the desired expansions around the origin hold under certain logarithmic growth conditions.

We demonstrate that under the above growth conditions the asymptotics of the characteristic function near zero always comes from the Edgeworth polynomials (regardless of whether the Edgeworth expansions hold or not). Those polynomials are defined canonically, and we show that under our logarithmic growth conditions the polynomials have bounded coefficients. The main step in our proofs is a verification of the latter growth conditions for the uniformly elliptic Markov chains considered in this paper. This is accomplished in Section 5. Using the sequential complex Perron-Frobenius Theorem from [40], the required estimates are obtained by studying the behavior around the origin of a resulting sequential complex pressure functions. For independent variables the n-th pressure function coincides with the logarithm of the characteristic function of the n-th summand, and our arguments essentially reduce to the ones in [27, 29]. In comparison with [40], where the Markov chains in random environment were studied, the main difficulty is that the variance does not grow linearly fast in the number of summands N. The Berry–Essen theorem is a direct consequence of the detailed asymptotics of the characteristic function near zero established in Section 5. The first order expansion also follows by combining the same estimates with the results of [20].

In order to achieve the desired rate of decay away from 0, an additional structure is needed. Thus we consider two special classes of additive functionals. The first is when the essential supremum of the n-th summand converges to 0 as $n \to \infty$. We show in Section 6 that if $\|f_n\|_\infty = O\left(n^{-\beta}\right)$ for some $\beta \in (0, 1/2)$ then the partial sums admit expansions of any order $r < \frac{1}{1-2\beta}$, and that this condition is optimal. The second type of additive functionals we consider are Hölder continuous functions. If $\{X_n\}$ is a Markov chain evolving on a compact Riemannian manifold with uniformly bounded and bounded away from 0 densities and $S_N = \sum_{n=1}^{N} f_n(X_n, X_{n+1})$, then we show in Section 7 that when f_n's are uniformly bounded Lipschitz functions then S_N admits Edgeworth expansions of all orders, while when f_n's are uniformly bounded Hölder continuous functions with exponent $\alpha \in (0, 1)$, then S_N admits expansions of every order $r < \frac{1+\alpha}{1-\alpha}$, and that the latter condition is optimal. In fact, we will show that the condition $r > \frac{1+\alpha}{1-\alpha}$ is optimal even in the stationary case when $\{X_n\}$ is homogeneous Markov chain and $f_n = f$ does not depend on n.

2. Main results

2.1. **A Berry-Esseen theorem and expansions of order 1.** Let $(\mathcal{X}_i, \mathcal{F}_i)$, $i \geq 1$ be a sequence of measurable spaces. For each i, let $R_i(x, dy)$, $x \in \mathcal{X}_i$ be a measurable family of (transition) probability measures on $\mathcal{X}_{i+1}$. Let μ_1 be any probability measure on $\mathcal{X}_1$, and let X_1 be an $\mathcal{X}_1$-valued random variable with distribution μ_1. Let $\{X_j\}$ be the Markov started from X_1 with the transition probabilities

$$\mathbb{P}(X_{j+1} \in A | X_j = x) = R_j(x, A),$$

where $x \in \mathcal{X}_j$ and $A \subset \mathcal{X}_{j+1}$ is a measurable set. Each R_j also gives rise to a transition operator given by

$$R_j g(x) = \mathbb{E}[g(X_{j+1})|X_j = x] = \int g(y) R_j(x, dy)$$

which maps an integrable function g on $\mathcal{X}_{j+1}$ to an integrbale function on $\mathcal{X}_j$ (the integrability is with respect to the laws of X_{j+1} and X_j, respectively). We assume here that there are probability measures $\mathfrak{m}_j$, $j > 1$ on $\mathcal{X}_j$ and families of transition probabilities $p_j(x, y)$ so that

$$R_j g(x) = \int g(y) p_j(x, y) d\mathfrak{m}_{j+1}(y).$$

Moreover, there exists $\varepsilon_0 > 0$ so that for any j we have

$$\sup_{x,y} p_j(x, y) \leq 1/\varepsilon_0, \tag{2.1}$$

and the transition probabilities of the second step[3] transition operators $R_j \circ R_{j+1}$ of X_{j+2} given X_j are bounded from below by ε_0 (this is the uniform ellipticity condition):

$$\inf_{j \geq 1} \inf_{x,z} \int p_j(x, y) p_{j+1}(y, z) d\mathfrak{m}_{j+1}(y) \geq \varepsilon_0. \tag{2.2}$$

Next, for a uniformly bounded sequence of measurable functions $f_n : \mathcal{X}_n \times \mathcal{X}_{n+1} \to \mathbb{R}$ we set $Y_n = f_n(X_n, X_{n+1})$ and

$$S_N = \sum_{n=1}^{N} (Y_n - \mathbb{E}(Y_n)). \tag{2.3}$$

Set $V_N = \mathrm{Var}(S_N)$ and $\sigma_N = \sqrt{V_N}$. Then by [20, Theorem 2.2] we have $\lim_{N \to \infty} V_N = \infty$ if and only if one can not decompose Y_n as

$$Y_n = \mathbb{E}(Y_n) + a_{n+1}(X_{n+1}, X_{n+2}) - a_n(X_n, X_{n+1}) + g_n(X_n, X_{n+1})$$

where a_n are uniformly bounded functions and $\sum_n g_n(X_n, X_{n+1})$ converges almost surely.

The CLT in the case $V_N \to \infty$ is due to [15], see [59] for a modern proof. Our first result here is a version of the Berry-Esseen theorem. Denote

$$\hat{S}_N = (S_N - \mathbb{E}[S_N]) / \sigma_N. \tag{2.4}$$

[3]The assumptions that we have uniform lower bound on the two step density and that the summands f_n introduced below depend only on two variables are taken form [20]. In fact, the arguments of [20] also work in the case we have uniform ellipticity after an arbitrary fixed number of steps and f_n depend on finitely many variables around x_n require only minor modifications (but lead to a significant complication of the notation). On the other hand there are some new effects in the case f depends on two variables which could not be seen in the case (considered in [15]) where f_n depend on a single variable. In this paper we keep the convention from [20] and assume two step ellipticity and two step dependence for additive functionals.

1. **Theorem.** *Suppose that* $\lim_{N\to\infty} V_N = \infty$. *Then there is a constant* $C > 0$ *which depends only on* $\sup_n \|Y_n\|_{L^\infty}$ *and* ε_0 *so that for any* $N \geq 1$,

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(\hat{S}_N \leq t) - \Phi(t)\right| \leq C\sigma_N^{-1} \tag{2.5}$$

where Φ *is the standard normal distribution function.*

Next we introduce some terminology from [20]. We say that a sequence Z_N of random variables is *center tight* if there are constants c_N such that $\{Z_N - c_N\}$ is tight. Two additive functionals f_n and $\tilde{f}_n$ are *homologous* if $\sum_{n=1}^{N}(f_n(X_n, X_{n+1}) - \tilde{f}_n(X_n, X_{n+1}))$ is center tight. We say that $\{f_n\}$ is *reducible* if it is homologous to an additive functional taking values in $h\mathbb{Z}$ for some $h > 0$. If $\{f_n\}$ is not reducible, it is called *irreducible.*

2. **Theorem.** *If* V_N *diverges and* $\{f_n\}$ *is irreducible then* S_N *satisfies the Edgeworth expansion of order 1, where*

$$P_{1,N}(t) = \frac{\mathbb{E}[(S_n - \mathbb{E}[S_n])^3]}{6V_N}(t^3 - 3t).$$

Next, we say that f_n *stably*[4] *obeys Edgeworth expansion of order* r if any additive functional homologous to f_n satisfies Edgeworth expansions of order r.

3. **Corollary.** f_n *stably obeys Edgeworth expansion of order 1 iff it is irreducible.*

Proof. If f_n is irreducible then any homologous additive functional $\tilde{f}_n$ is also irreducible, so by Theorem 2, $\tilde{f}_n$ obeys Edgeworth expansion of order 1.

If f_n is reducible then its homology class contains an $h\mathbb{Z}$ valued functional $\tilde{f}_n$, for some $h > 0$. By the LLT of [20, Section 5], $\tilde{S}_N$ has jumps of order $1/\sqrt{V_N}$, so $\tilde{S}_N$ does not obey expansion of order 1. □

2.2. High order expansions.

2.2.1. *Summands with small essential supremum.* We obtain the following extension of the Edgeworth expansions for function f_n which converge to 0 as $n \to \infty$.

4. **Theorem.** *Suppose that* $\lim_{N\to\infty} V_N = \infty$, *and that there are* $C > 0$ *and* $\beta \in (0, 1/2)$ *so that for all* $n \in \mathbb{N}$ *we have* $\|f_n\|_\infty \leq \frac{C}{n^\beta}$. *Let* $r \geq 1$ *be an integer satisfying*

$$r < \frac{1}{1 - 2\beta}. \tag{2.6}$$

Then S_N *admits an Edgeworth expansion of order* r. *In particular, if* $\|f_n\|_\infty = O(n^{-1/2})$ *then* S_N *admits Edgeworth expansions of all orders.*

[4]The notion of stable Edgeworth expansion is motivated by the notion of stable local limit theorem studied in [53, 57]. We note that [19] obtains conditions for the stability of Edgeworth expansions for the sums of independent integer valued random variables (in the integer case one studies the expansions for $\mathbb{P}(S_N = k_N)$).

The following result shows that the conditions of Theorem 4 are optimal.

5. **Theorem.** *For every* $\beta \in (0, \frac{1}{2})$ *there exists a sequence of centered independent random variables* X_n *so that* $C_1 n^{-\beta} \leq \|X_n\|_{L^\infty} \leq C_2 n^{-\beta}$ *for some* $C_1, C_2 > 0$ *and all* n *large enough,* $V(S_N)$ *is of order* $N^{1-2\beta}$ *but* $S_N = \sum_{n=1}^{N} X_n$ *fails to satisfy Edgeworth expansions of any order* s *such that* $s > \frac{1}{1-2\beta}$.

Taking $\beta \in (0, 1/4)$ we have $\frac{1}{1-2\beta} < 2$, and we get from Theorem 5 that S_N might not admit Edgeworth expansions of order larger than 1 if $\|f_n\|_\infty \asymp n^{-\beta}$.

2.2.2. *Markov chains on compact Riemannian manifolds.* Let us assume that $\{X_n\}$ is a Markov chain on a compact Riemannian manifold M with transition densities $p_n(x, y)$ bounded and bounded away from 0, uniformly in n. Let $\alpha \in (0, 1]$ and let $f_n : M \times M \to \mathbb{R}$ be observables satisfying $\|f_n\|_\alpha := \max(\sup |f_n|, v_\alpha(f_n)) \leq 1$, where $v_\alpha(f_n)$ is the Hölder constant of f_n corresponding to the exponent α. Consider the sum

$$S_N = \sum_{n=1}^{N} f_n(X_n, X_{n+1}).$$

6. **Theorem.** *Suppose that* $V_N = V(S_N) \to \infty$.

(i)If $\alpha = 1$ *then* S_N *satisfies the Edgeworth expansion of all orders.*

(ii) If $\alpha < 1$ *then* S_N *satisfies the Edgeworth expansion of any order* $r < \frac{1+\alpha}{1-\alpha}$.

For smooth functions, expansions of all orders were obtained in [30] for stationary Markov chains and functions $f_n = f$ which do not depend on n. Here we have to overcome the difficulty that the variance of $f_n(X_n, X_{n+1})$ might be small, and hence the proof differs from the one in [30] even for smooth functions, so it is also new in the stationary case. The proof of Theorem 6 follows the approach of [16]. We note that similar estimates are used in [16, 17] to prove polynomial bounds for the decay of correlations for hyperbolic suspension flows with Hölder roof functions. However, the bound of [16, 17] are not explicit whereas here we get an explicit (and optimal, see below) control on the possible location of resonances.

We see that as $\alpha \to 1$, the largest order of the expansions ensured by Theorem 6(ii) diverges to ∞. The following theorem shows that the conditions of Theorem 6(ii) are optimal.

7. **Theorem.** *Let* $\{x_n\}$ *be iid random variables uniformly distributed on* $[-1, 1]$. *For every* $0 < \alpha < 1$ *there exists an increasing odd function* $f : [-1, 1] \to [-1, 1]$ *which is Hölder continuous with exponent* α *and is onto* $[-1, 1]$, *so that* $S_n = \sum_{j=1}^{n} f(x_j)$ *does not admit Edgeworth expansion of any order* $r > \frac{1+\alpha}{1-\alpha}$.

Theorem 7 show that the conditions of Theorem 6(ii) are optimal even in the stationary case. The idea in the proof of Theorem 7 is to first approximate α by numbers of the form $\alpha_{q,p} = \ln(p)/\ln(p+q)$, for some $p, q \geq 2$ so that $q|(p-1)$. Then, the restriction of the function f to $[0,1]$ will be the, so called, Cantor function (see [32]) corresponding to a certain Cantor set with Hausdorff dimension $\alpha_{q,p}$.

2.3. **The canonical form of the Edgeworth polynomials.** We note that in the non-stationary setting, (1.3) does not define the Edgeworth polynomials uniquely since we could always modify the coefficients by terms of order $o(\sigma_N^{-r})$. However, it turns out that one could make a canonical choice which a simple computation of its coefficient in a quite general setting including additive functionals of uniformly elliptic Markov chains considered here.

Given a nonconstant random variable S with finite moments of all orders, let $a_j(S)$ denote the normalized cumulant

$$a_j(S)=\frac{1}{V(S)i^j}\frac{d^j}{dt^j}\Big|_{t=0}\ln\left[\mathbb{E}\left(e^{it(S-\mathbb{E}(S))}\right)\right].$$

8. **Theorem.** *There exist polynomials $\mathfrak{P}_j(z;a_3,a_4,\dots,a_{3j})$ such that for each integer $r\geq 1$ there is a positive constant $\delta_r=\delta_r(\varepsilon_0,K)$, $K=\sup_n\|f_n(X_n,X_{n+1})\|_{L^\infty}$, such that if S_N and $\hat{S}_N$ are given by (2.3) and (2.4), respectively, then denoting*

$$P_{j,N}(z)=\mathfrak{P}_j(z,a_3(S_N),\dots,a_{3j}(S_N)),\tag{2.7}$$

$\mathcal{E}_{r,N}(z)=\Phi(z)+\phi(z)\sum_{j=1}^{r}\sigma_N^{-j}P_{j,N}(z)$ and letting $\widehat{\mathcal{E}}_{r,N}$ denote the Fourier transform of $\mathcal{E}_{r,N}(z)$ we have

$$\int_{-\delta_r\sigma_N}^{\delta_r\sigma_N}\left|\frac{\mathbb{E}\left(e^{it\hat{S}_N}\right)-\widehat{\mathcal{E}_{r,N}}(t)}{t}\right|dt=O\left(\sigma_N^{-(1+r)}\right).\tag{2.8}$$

We note that our proofs of Theorems 2, 4 and 6 provide the Edgeworth expansions with the above polynomials $P_{j,n}$.

The polynomials $\mathfrak{P}_j$ are given in Definition 15. In §5.4 we show that for additive functionals of the Markov chains considered in this paper the Edgeworth polynomials have bounded coefficients. This is done by verifying Assumption 17 which ensures the boundness for an abstract sequence of random variables.

We note that (2.8) holds without any additional assumptions. However, to ensure that the term $\mathcal{E}_{r,N}(z)$ provides a good approximation to $\mathbb{P}(\hat{S}_N\leq z)$ we need to control the LHS of (2.8) on longer intervals of size $B\sigma_N^r$ for an arbitrary B. In the case $r=1$ the contribution of $[-B\sigma_N,B\sigma_N]\setminus[-\delta_1\sigma_N,\delta_1\sigma_N]$ is analyzed in [20]. The case $r>1$ is addressed in Sections 6 and 7 where we control the characteristic function of $\hat{S}_N$ under the assumptions of Theorems 4 and 6, respectively.

3. Background

3.1. **A sequential RPF theorem.** For all $j\in\mathbb{N}$ and $z\in\mathbb{C}$, let $R_z^{(j)}$ the operator given by

$$R_z^{(j)}g(x)=\mathbb{E}[g(X_{j+1})e^{zf_j(X_j,X_{j+1})}|X_j=x]=R_j(e^{zf_j(x,\cdot)}g)(x)$$

where $g : \mathcal{X}_{j+1} \to \mathbb{R}$ is a bounded function. Denote by B_j the space of bounded functions on $\mathcal{X}_j$, equipped with the supremum norm $\|\cdot\|_\infty$. For every integer $j \geq 1$, $n \in \mathbb{N}$ and $z \in \mathbb{C}$ consider the n-th order iterates $R_z^{j,n} : B_{j+n} \to B_j$ given by

$$R_z^{j,n} = R_z^{(j)} \circ R_z^{(j+1)} \circ \cdots \circ R_z^{(j+n-1)}. \tag{3.1}$$

We have the following.

9. **Theorem.** *There exists a complex neighborhood U of 0 which depends only on $\|f\|_\infty := \sup_j \sup |f_j|$ and ε_0 (from the definition of the uniform ellipticity) so that for any $z \in U$ and an integer $j \geq 1$ there exists a triplet $\lambda_j(z)$, $h_j^{(z)}$ and $\nu_j^{(z)}$ consisting of a nonzero complex number $\lambda_j(z)$, a complex function $h_j^{(z)} \in B_j$ and a continuous linear functional $\nu_j^{(z)} \in B_j^*$ satisfying that $\nu_j^{(z)}(\mathbf{1}) = 1$, $\nu_j^{(z)}(h_j^{(z)}) = 1$ and*

$$R_z^{(j)} h_{j+1}^{(z)} = \lambda_j(z) h_j^{(z)}, \quad \text{and} \quad (R_z^{(j)})^* \nu_j^{(z)} = \lambda_j(z) \nu_{j+1}^{(z)}$$

where $(R_z^{(j)})^* : B_j^* \to B_{j+1}^*$ is the dual operator of $R_j^{(z)}$ and B_j^* is the dual space of the Banach space B_j. When $z = t \in \mathbb{R}$ then $h_j^{(t)}$ is strictly positive, $\nu_j^{(t)}$ is a probability measure and there exist constants $a, b > 0$, which depend only on $\|f\|_\infty$ and ε_0 so that $\lambda_j(t) \in [a, b]$ and $h_j^{(t)} \geq a$. When $t = 0$ we have $\lambda_j(0) = 1$ and $h_j^{(0)} = \mathbf{1}$.

Moreover, this triplet is analytic and uniformly bounded. Namely, the maps

$$\lambda_j(\cdot) : U \to \mathbb{C}, \; h_j^{(\cdot)} : U \to B_j \quad \text{and} \quad \nu_j^{(\cdot)} : U \to B_j^*$$

are analytic, where B_j^* is the dual space of B_j, and there exists a constant $C > 0$ so that

$$\max\left(\sup_{z \in U} |\lambda_j(z)|, \sup_{z \in U} \|h_j^{(z)}\|_\infty, \sup_{z \in U} \|\nu_j^{(z)}\|_\infty \right) \leq C \tag{3.2}$$

where $\|\nu\|_\infty$ is the operator norm of a linear functional $\nu : B_j \to \mathbb{C}$.

Furthermore, there exist constants $C > 0$ and $\delta \in (0,1)$ such that for all $n \geq 1$, $j \in \mathbb{N}$, $z \in U$ and $q \in B_{j+n}$,

$$\left\| \frac{R_z^{j,n} q}{\lambda_{j,n}(z)} - \big(\nu_{j+n}^{(z)}(q)\big) h_j^{(z)} \right\|_\infty \leq C \|q\|_\infty \cdot \delta^n \tag{3.3}$$

and

$$\left\| \frac{(R_z^{j,n})^* \mu}{\lambda_{j,n}(z)} - \big(\mu h_j^{(z)}\big) \nu_{j+n}^{(z)} \right\|_\infty \leq C \|\mu\|_\infty \cdot \delta^n \tag{3.4}$$

where $\lambda_{j,n}(z) = \prod_{k=0}^{n-1} \lambda_{j+k}(z)$.

The proof of Theorem 9 was given in [40, Ch.4&6] by a successive application of the complex projective contraction.We remark that the arguments in [40, Ch.4&6] formally require us to have a two sided sequence of operators, and in order to overcome this technical difficulty, for $j \leq 0$ we define $\mathcal{X}_j = \mathcal{X}_1$ and $R_z^{(j)} g(x) = \mathbb{E}[g(X_1)]$. This amount to taking independent copies $\{Z_j : j \leq 0\}$ of X_1, setting $X_j = Z_j$ for $j \leq 0$ and $f_j = 0$.

In fact, in [40, Ch.4&6] the setup included random operators $R_z^{(j)} = R_z^{\theta^j \omega}$, when $\omega \in \Omega$ and $(\Omega, \mathcal{F}, P, \theta)$ is some invertible measure preserving system, which is not necessarily ergodic. The main reason for considering random operators in [40], and not just a sequence of operators, is that the random Ruelle-Perron-Frobenius (RPF) theorem was needed in the proof of the local CLT from [40, Ch. 2], where random operators arise after a certain conditioning argument. The measurability of the resulting RPF triplets $\lambda_\omega(z), h_\omega^{(z)}, \nu_\omega^{(z)}$ as functions of ω played an important rule in that proof, which lead to consider a more general steup of random operators in [40, Ch. 4], for which there is meaning to such measurability. However, in our purely sequential setup such measurability issues do not arises, and thus we can just repeat the arguments from [40, Ch. 4] pertaining to a fixed ω and ignore the ones addressing measurability.

10. **Remark.** In the proof of the Berry-Esseen theorem and the Edgeworth expansions it will be convenient to assume that $a_n := \mathbb{E}[f_n(X_n, X_{n+1})] = 0$. This amount to replacing f_n with $f_n - a_n$, and hence to replacing R_z^j with $e^{-a_j z} R_z^j$ and replacing $\lambda_j(z)$ with $e^{-za_j} \lambda_j(z)$.

3.2. **The structure constants.** As it was mentioned in the introduction a new feature of our work is that we do not make any assumptions on how slow variance of S_N grows. In this section we recall a few results from [20] which provide some geometric control on the variance.

By a *random hexagon based at* n we mean a tuple

$$P_n = (\mathscr{X}_{n-2}, \mathscr{X}_{n-1}, \mathscr{X}_n; \mathscr{Y}_{n-1}, \mathscr{Y}_n, \mathscr{Y}_{n+1})$$

where $(\mathscr{X}_{n-2}, \mathscr{X}_{n-1})$ and $(\mathscr{Y}_n, \mathscr{Y}_{n+1})$ are independent, $(\mathscr{X}_{n-2}, \mathscr{X}_{n-1})$ and (X_{n-2}, X_{n-1}) are equality distributed, $(\mathscr{Y}_n, \mathscr{Y}_{n+1})$ and (X_n, X_{n+1}) are equality distributed and $\mathscr{X}_n$ and $\mathscr{Y}_{n-1}$ are conditionally independent given the previous choices and are sampled according to the *bridge distributions*

$$\mathbb{P}(\mathscr{X}_n \in E | \mathscr{X}_{n-1} = x_{n-1}, \mathscr{Y}_{n+1} = y_{n+1}) = \mathbb{P}(X_n \in E | X_{n-1} = x_{n-1}, X_{n+1} = y_{n+1})$$

and

$$\mathbb{P}(\mathscr{Y}_{n-1} \in E | \mathscr{X}_{n-2} = x_{n-2}, \mathscr{Y}_n = y_n) = \mathbb{P}(X_{n-1} \in E | X_{n-2} = x_{n-2}, X_n = y_n).$$

The balance $\Gamma(P_n)$ of the hexagon is given by

$$\Gamma(P_n) = f_{n-2}(\mathscr{X}_{n-2}, \mathscr{X}_{n-1}) + f_{n-1}(\mathscr{X}_{n-1}, \mathscr{X}_n) + f_n(\mathscr{X}_n, \mathscr{Y}_{n+1})$$
$$-f_{n-2}(\mathscr{X}_{n-2}, \mathscr{Y}_{n-1}) - f_{n-1}(\mathscr{Y}_{n-1}, \mathscr{Y}_n) - f_n(\mathscr{Y}_n, \mathscr{Y}_{n+1}).$$

Next, let

$$u_n^2 = \mathbb{E}[\Gamma(P_n)^2]. \tag{3.5}$$

11. **Theorem** ([20], Theorem 2.1)**.** *There exist positive constants C_1, C_2, C_3, C_4 so that for any $m \geq 0$ and $N \geq 3$,*

$$C_1 \sum_{n=m+3}^{m+N} u_n^2 - C_2 \leq V_N = \mathrm{Var}(S_N - S_m) \leq C_3 \sum_{n=m+3}^{m+N} u_n^2 + C_4. \tag{3.6}$$

It turns out that the hexagon process also allows to control the characteristic function of S_N. Denote

$$d_n(\xi)^2 = \mathbb{E}[|e^{i\xi\Gamma(P_n)} - 1|^2] = 4\mathbb{E}[\sin^2(\xi\Gamma(P_n)/2)], \quad D_N(\xi) = \sum_{n=1}^{N} d_n^2(\xi). \tag{3.7}$$

12. **Lemma** ([20], eq. (4.2.6)). *There are constants $C, c > 0$ so that for each N and $\xi \in \mathbb{R}$, the characteristic function $\Phi_N(\xi) = \mathbb{E}\left(e^{i\xi S_N}\right)$ satisfies*

$$|\Phi_N(\xi)| \leq Ce^{-cD_N(\xi)}. \tag{3.8}$$

3.3. **Mixing and moment estimates.** Next we discuss the mixing properties of $\{X_n\}$.

13. **Lemma** (Proposition 1.11 (2), [20]). *There exist $\delta \in (0,1)$ and $A > 0$ so that for all $n, k \in \mathbb{N}$ we have*

$$\big|\text{Cov}\big(f_n(X_n, X_{n+1}), f_{n+k}(X_{n+k}, X_{n+k+1})\big)\big| \leq A\delta^k.$$

Next, for each j and n consider the random variable $S_{j,n}$ given by

$$S_{j,n} = \sum_{k=j}^{j+n-1} f_k(X_k, X_{k+1}). \tag{3.9}$$

Then $S_{1,n} = S_n$.

14. **Lemma** (Lemma 2.16, [20]). *For every integer $p \geq 1$ there are constant $C_p, R_p > 0$ so that for all j and n,*

$$\left|\mathbb{E}\left[\left(S_{j,n} - \mathbb{E}(S_{j,n})\right)^p\right]\right| \leq R_p + C_p\Big(\text{Var}(S_{j,n})\Big)^{[p/2]}.$$

4. Edgewoth expansions under logarithmic growth assumptions

4.1. **The Edgewoth polynomials.** Let S be a random variable with finite moments of all orders. We recall that the k-th cumulant of S is given by

$$\Gamma_k(S) = \frac{1}{i^k}\frac{d^k}{dt^k}\big(\ln \mathbb{E}[e^{itS}]\big)\big|_{t=0}.$$

Note that $\Gamma_k(aS) = a^k\Gamma_k(S)$ for every $a \in \mathbb{R}$. Moreover, $\Gamma_1(S) = \mathbb{E}[S]$, $\Gamma_2(S) = \text{Var}(S)$ and for $k \geq 3$ by (1.34) in [51], we have

$$\Gamma_k(S) = \sum_{v=1}^{k} \frac{(-1)^{v-1}}{v} \sum_{k_1+\ldots+k_v=k} \frac{k!}{k_1!k_2!\cdots k_v!}\alpha_{k_1}\alpha_{k_2}\cdots\alpha_{k_v} \tag{4.1}$$

where $\alpha_m = \alpha_m(S) = \mathbb{E}[S^m]$ (this formula is a consequence of the Taylor expansion of the function $\ln(1+z)$).

The cumulants of order $k \geq 3$ measure the distance of the distribution of $\hat{S} = \big(S - \mathbb{E}[S]\big)/\sigma$, from the standard normal distribution, where $\sigma = \sqrt{\text{Var}(S)}$, assuming of course that $\sigma > 0$. We have $\Gamma_k(S) = 0$ for all $k \geq 3$ if and only if $\hat{S}$ is standard normal, and we refer to [51] for conditions on $\Gamma_k(\hat{S})$ which insure that the distribution function of $\hat{S}$ is close to the standard normal distribution function in the uniform metric.

We also refer to [3, 55] for expansions of expectations of smooth functions of $\hat{S}$ which involve growth properties of cumulants.

Next, let us assume that $\mathbb{E}[S] = 0$ and $\sigma^2 = \mathbb{E}[S^2] > 0$. Consider the function

$$\Lambda(t; S) = \ln \mathbb{E}[e^{itS/\sigma}] + t^2/2.$$

Then $\Lambda(0; S) = 0$, $\Lambda'_n(0; S) = \mathbb{E}[S] = 0$, $\Lambda''_n(0; S) = \mathbb{E}[S^2]/\sigma^2 - 1 = 0$, and for $k \geq 3$ we have

$$\Lambda^{(k)}(0) := \frac{d^k}{dt^k}\Lambda(t; S)\big|_{t=0} = i^k \Gamma_k(S)\sigma^{-k}.$$

Thus, the k-th Taylor polynomial of $\Lambda(t; S)$ is given by

$$\mathcal{P}_k(t; S) = \sum_{j=3}^{k} \frac{i^j \Gamma_j(S)}{j!\sigma^j} t^j = \sum_{j=3}^{k} i^j a_j(S)\sigma^{-(j-2)} t^j.$$

where[5] $a_j(S) = \frac{\Gamma_j(S)}{j!\sigma^2}$. Let us consider the formal power series

$$\Gamma(t; S) = \sum_{j \geq 3} \frac{i^j \Gamma_j(S)}{j!\sigma^j} t^j = \sum_{j \geq 3} i^j a_j(S)\sigma^{-(j-2)} t^j,$$

where $a_j(S)$ is viewed as a variable independent of σ. This leads to the following formal series

$$\exp(\Gamma(t; S)) = 1 + \sum_{j \geq 1} \frac{i^j \Gamma(t; S)^j}{j!} = 1 + \sum_{j \geq 1} \sigma^{-j} A_j(t; S)$$

where $A_j(t; S)$ is the polynomial given by

$$A_j(t; S) = \sum_{m=1}^{j} \frac{1}{m!} \sum_{k_1, \dots, k_m \in \mathcal{A}_{j,m}} \prod_{u=1}^{m} i^{k_i} a_{k_i}(S) t^{j+2m}$$

and $\mathcal{A}_{j,m}$ is the set of all m-tuples $(k_1, \dots, k_m)$ of integers such that

$$k_i \geq 3 \quad \text{and} \quad \sum_i k_i = 2m + j.$$

15. **Definition.** The j-th Edgewoth polynomial S is the unique polynomial $P_j(t; S)$ so that the Fourier transform of $\phi(t)P_j(t; S)$ is $e^{-t^2/2}A_j(t; S)$, where $\phi(t)$ is the standard normal density.

Notice that the polynomials $A_j(t; S)$ and $P_j(t; S)$ depend on S only through the first $3j$ moments. Note also that $A_j(0; S) = 0$ for all j.

[5] The reason we divide $\Gamma_j(S)$ by σ^2 is that under suitable restrictions on S, the quantities $|\Gamma_j(S)\sigma^{-2}|$ will be bounded by some constant not depending on S (see next section). This will be the case when $S = S_n$, for which the latter quantities will be bounded in n. Here S_n are the sums considered in Section 2.

16. **Remark.** In order to compute $A_j(t;S)$ for $j \le k$ it is enough to expand $e^{\mathcal{P}_{k+2}(t;S)}$ to a power series and represent it in the form $1 + \sum_{j\ge 1} \sigma_n^{-j} \tilde{A}_j(t;S)$. Indeed, it follows that $\tilde{A}_j(t;S) = A_j(t;S)$ for all $j \le k$ since

$$\Gamma(t,S) - \mathcal{P}_{k+2}(t;S) = \sigma^{-(k+1)} \sum_{j=k+3}^{\infty} i^j a_j(S) \sigma^{-(j-k-3)} t^j.$$

Thus, to compute $A_j(t;S)$, $j \le k$ we first write

$$e^{\mathcal{P}_{k+2}(t;S)} = 1 + \sum_{j=1}^{\infty} \frac{\mathcal{P}_{k+2}(t;S)^j}{j!}.$$

Now, since $\mathcal{P}_k(t;S)$ has a factor[6] σ^{-1}, we can compute $A_j(t;S)$, $j \le k$ by considering only the first k summands

$$1 + \sum_{j=1}^{k} \frac{\mathcal{P}_{k+2}(t;S)^j}{j!}.$$

After writing the above expression in the form $1+\sum_{j=1}^{\infty} \sigma^{-j} \bar{A}_{j,k}(t;S)$ (this is a finite sum) we have $A_j(t;S) = \bar{A}_{j,k}(t;S)$ for all $j \le k$.

In particular $\mathcal{P}_3(t;S) = \frac{i^3 a_3(S) t^3}{6\sigma} = \frac{A_1(t;S)}{\sigma}$, whence

$$P_1(t;S) = \frac{a_3(S)}{6}(t^3 - 3t) = \frac{\mathbb{E}[(S - \mathbb{E}[S])^3]}{6\sigma^2}(t^3 - 3t)$$

where we have used that the transform Fourier of $(t^3 - 3t)\phi(t)$ is $i^3 e^{-\frac{1}{2}\xi^2} \xi^3$.

4.2. **A Berry-Esseen theorem and Edgeworth expansions via decay of characteristic functions.** Let W_n be a sequence of centered random variables so that $\lim_{n\to\infty} \mathrm{Var}(W_n) = \infty$. Let us set

$$\Gamma_n(t) = \Gamma(t;W_n), \quad \Lambda_n(t) = \Lambda(t;W_n), \quad A_{j,n}(t) = A_j(t;W_n), \quad P_{j,n}(t) = P_j(t;W_n).$$

Let us also set $\sigma_n = \sqrt{\mathrm{Var}(W_n)}$. We will prove here Edgeworth expansions under the following logarithmic growth assumptions.

17. **Assumption.** For some $k \ge 3$, for all $j \le k$ there exist constants $C_j, \varepsilon_j > 0$ so that

$$\sup_{t \in [-\varepsilon_j \sigma_n, \varepsilon_j \sigma_n]} |\Lambda_n^{(j)}(t)| \le C_j \sigma_n^{-(j-2)}. \tag{4.2}$$

Note that under Assumption 17 the polynomials $A_{j,n}$ and $P_{j,n}$, $j \le k$ have bounded coefficients (for that it is enough to only consider $t = 0$). For $t = 0$ conditions of the form $|\Lambda_n^{(j)}(0)| = |\Gamma_j(W_n/\sigma_n)| \le (j!)^{1+\gamma} \sigma_n^{-(j-2)}, \gamma \ge 0$ appear in literature [21, 23, 51] in the context of moderate deviations and related results (see also references therein).

The relevance of Assumption 17 stems from the following facts.

[6]Recall that $a_j(S)$ are viewed as constants.

18. **Proposition.** *Let Assumption 17 hold with $k = 3$. Then there exists a constant $C > 0$ so that for every $n \geq 1$ we have*

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(W_n/\sigma_n \leq t) - \Phi(t)| \leq C\sigma_n^{-1}$$

where Φ is the standard normal distribution and density function.

19. **Proposition.** *Let $r \geq 1$ be an integer. Let Assumption 17 hold with $k = r + 3$. Suppose also that for every $B > 0$ and all $\delta > 0$*

$$\int_{\delta\leq|x|\leq B\sigma_n^{r-1}} |\mathbb{E}(e^{ixW_n})/x|dx = o(\sigma_n^{-r}). \tag{4.3}$$

Then

$$\sup_t \left|\mathbb{P}(W_n/\sigma_n \leq t) - \Phi(t) - \sum_{j=1}^{r} \sigma_n^{-j} P_{j,n}(t)\phi(t)\right| = o(\sigma_n^{-r}) \tag{4.4}$$

where Φ and ϕ are the standard normal distribution and density function, respectively.

4.3. **Auxillary estimates.** Here we present several technical estimates needed in the proofs of Propositions 18 and 19.

We need two lemmata.

20. **Lemma.** *Let $k \geq 3$ be an integer and let Assumption 17 hold with this k. Then there exist constants $\delta_k, B_k > 0$ so that for every real $t \in [-\sigma_n, \sigma_n]$,*

$$|\mathcal{P}_{k,n}(t)| \leq B_k\sigma_n^{-1}|t|^3 = B_k t^2|t/\sigma_n|.$$

Therefore, there is a constant $\delta_k > 0$ so that for every $t \in [-\delta_k\sigma_n, \delta_k\sigma_n]$,

$$|e^{\mathcal{P}_{k,n}(t)}| \leq e^{t^2/4}.$$

21. **Lemma.** *Let Assumption 17 hold with $k = 3$. Then there exist $\delta_0 > 0$ and $\alpha \in (0, 1/2)$ so that for every real t such that $|t/\sigma_n| \leq \delta_0$ we have $|e^{\Lambda_n(t)}| \leq e^{\alpha t^2}$.*

Proof of Lemmas 20 and 21. Let us first prove Lemma 20. By taking $t = 0$ in (4.2) and using that $\Gamma_j(aW) = a^jW$ we have $|\Gamma_j(W_n)| \leq C_j\sigma_n^2$. Thus, if $|t/\sigma_n| < 1$ then with $A_k = \max_{3\leq j\leq k} C_j$ and $B_k = kA_k$ we have

$$|\mathcal{P}_{k,n}(t)| \leq \sum_{j=3}^{k} \frac{|\Gamma_j(W_n)|}{j!\sigma_n^j}|t|^j \leq A_k t^2 \sum_{j=3}^{k} |t/\sigma_n|^{j-2}/j! \leq A_k t^2 \sum_{j=3}^{k} |t/\sigma_n| \leq B_k|t|^3\sigma_n^{-1}.$$

Hence, if $|t/\sigma_n| \leq \frac{1}{4B_k} := \delta_k$ then $|\mathcal{P}_{k,n}(t)| \leq t^2/4$ and so

$$\left|e^{\mathcal{P}_{k,n}(t)}\right| \leq e^{t^2/4}.$$

Finally, to prove Lemma 21, using that the second Taylor polynomial $\mathcal{P}_{2,n}(t)$ of Λ_n around the origin vanishes, we can write write $\Lambda_n(t) = \mathcal{P}_{2,n}(t) + \mathcal{R}_{2,n}(t) = \mathcal{R}_{2,n}(t)$, where $\mathcal{R}_{2,n}(t)$ is the Taylor remainder of order 2 around the origin. Then by the Lagrange form of the Taylor remainder we can write $\mathcal{R}_{2,n}(t) = \frac{t^3\Lambda_n'''(t_1)}{3!}$ for some t_1 such that $|t_1| \leq |t|$. Therefore, by Assumption 17 we have

$$\mathcal{R}_{2,n}(t) = O(t^3/\sigma_n) = t^2O(|t/\sigma_n|), \text{ if } |t| \leq \varepsilon_1.$$

Thus when $|t/\sigma_n|$ is small enough $|\Lambda_n(t)| = |\mathcal{R}_{2,n}(t)| < t^2/3$, and Lemma 21 follows with $\alpha = 1/3$. □

Combining Lemma 21 with Proposition 24 proven in Section 5 we recover the following result, which was proved in [20, Theorem 6.1] for the uniformly elliptic Markov chains considered in this paper.

22. **Corollary.** *Under assumption 17 with $k = 3$ there exist constants $c > 0$ and $\delta > 0$ so that for every natural n and $t \in [-\delta, \delta]$ we have*

$$|\mathbb{E}[e^{itW_n}]| \leq e^{-ct^2\sigma_n^2}.$$

The key step in estimating the rate of convergence for the CLT is the following.

23. **Proposition.** *Let $r \geq 0$ be an integer and let Assumption 17 hold with $k = r + 3$. Then there is a constant $\delta_r > 0$ such that*

$$\int_{-\delta_r\sigma_n}^{\delta_r\sigma_n} \left| \frac{\mathbb{E}[e^{itW_n/\sigma_n}] - e^{-t^2/2}(1 + Q_{r,n}(t))}{|t|} \right| dt = O(\sigma_n^{-r-1})$$

where for $r = 0$ we set $Q_{0,n}(t) = 0$ and for $r \geq 1$,

$$Q_{r,n}(t) = \sum_{j=1}^{r} \sigma_n^{-j} A_{j,n}(t)$$

Proof. Write

$$\mathbb{E}[e^{itW_n/\sigma_n}] = e^{-t^2/2}e^{\Lambda_n(t)} = e^{-t^2/2}e^{\mathcal{P}_{r+2,n}(t)+\mathcal{R}_{r+2,n}(t)} \tag{4.5}$$

where $\mathcal{R}_{r+2,n}(t)$ is the Taylor remainder of order $r + 2$ around 0. Using the Lagrange form of Taylor remainders together with Assumption 17 we get that

$$\mathcal{R}_{r+2,n}(t) = O(t^{r+3}\sigma_n^{-(r+1)}). \tag{4.6}$$

Next, by the mean value theorem and Lemmas 20 and 21 there are constants $\delta_r > 0$, $C_0 > 0$ and $b_0 \in (0, 1/2)$ so that if $|t/\sigma_n| \leq \delta_r$ then

$$\left|e^{\Lambda_n(t)} - e^{\mathcal{P}_{r+2,n}(t)}\right| \leq C_0 e^{b_0 t^2} |\mathcal{R}_{r+2,n}(t)|. \tag{4.7}$$

Moreover, by Lemma 20 and the Lagrange form of Taylor remainders,

$$\left| e^{\mathcal{P}_{r+2,n}(t)} - \left(1 + \sum_{j=1}^{r} \frac{\mathcal{P}_{j+2,n}(t)^j}{j!}\right) \right| \leq D_r e^{b_0 t^2} \sigma_n^{-(r+1)} |t|^{3(r+2)} \tag{4.8}$$

where $D_r > 0$ is some constant (when $r = 0$ then the left hand side vanishes since $\mathcal{P}_{2,n}(t) = 0$). Combining (4.5), (4.6), (4.7) and (4.8), for every real t so that $|t/\sigma_n| \leq \delta_r$ we have

$$\left| \mathbb{E}[e^{itW_n/\sigma_n}] - e^{-t^2/2} \left(1 + \sum_{j=1}^{r} \frac{\mathcal{P}_{j+2,n}(t)^j}{j!}\right) \right| \leq Ce^{-ct^2}\sigma_n^{-(r+1)} \max\left(|t|, |t|^{(r+3)(r+2)}\right)$$

where $c = 1/2 - b_0 > 0$. Next, by Remark 16, we have

$$\sum_{j=1}^{r} \frac{\mathcal{P}_{r+2,n}(t)^j}{j!} = Q_{r,n}(t) + \max(|t|, |t|^{r(r+2)}) O(\sigma_n^{-r-1})$$

where the term $\max(|t|, |t|^{r(r+2)}) O(\sigma_n^{-r-1})$ comes from the terms which include powers of σ_n^{-1} larger than r (when $r = 0$ both the left hand side and $Q_{r,n}(t)$ equal 0). We conclude that

$$\left| \mathbb{E}[e^{itW_n/\sigma_n}] - e^{-t^2/2}(1 + Q_{r,n}(t)) \right| \leq Ce^{-ct^2} \sigma_n^{-(r+1)} \max\left(|t|, |t|^{(r+3)(r+2)}\right).$$

Therefore,

$$\int_{-\delta_r\sigma_n}^{\delta_r\sigma_n} \left| \frac{\mathbb{E}[e^{itW_n/\sigma_n}] - e^{-t^2/2}(1 + Q_{r,n}(t))}{|t|} \right| dt$$

$$\leq C\sigma_n^{-(r+1)} \int_{-\infty}^{\infty} e^{-ct^2} \left(1 + |t|^{(r+3)(r+2)-1}\right) dt \leq C'\sigma_n^{-(r+1)}$$

completing the proof of the proposition. □

4.4. Proofs of Propositions 18 and 19.

Proof of Proposition 18. The first step in the proof is quite standard. We use generalized Esseen inequality [29, §XVI.3]. Let $F : \mathbb{R} \to \mathbb{R}$ be a probability distribution function and $G : \mathbb{R} \to \mathbb{R}$ be a differential function with bounded derivative so that $G(-\infty) = 0$. Let $f(t) = \int e^{itx} dF(x)$ and $g(t) = \int e^{itx} dG(x)$ be the corresponding Fourier transforms. Then for every $T > 0$ we have

$$\sup_{x \in \mathbb{R}} |F(x) - G(x)| \leq 2 \int_{-T}^{T} \left| \frac{f(t) - g(t)}{t} \right| dt + \frac{24\|G'\|_\infty}{\pi T}. \tag{4.9}$$

Taking F to be the distribution of W_n/σ_n, G to be the standard normal distribution and $T_n = \delta_1\sigma_n$ where δ_1 comes from Lemma 20 we conclude that Proposition 18 will follow if we prove that

$$\int_{-\delta_1\sigma_n}^{\delta_1\sigma_n} \left| \frac{\mathbb{E}[e^{itW_n/\sigma_n}] - e^{-t^2/2}}{t} \right| dt \leq C\sigma_n^{-1} \tag{4.10}$$

for some constant C. Finally, (4.10) follows from Proposition 23 with $r = 0$. □

Proof of Proposition 19. Relying on Proposition 23, the proof proceeds essentially in the same way as [27, 29]. We provide the details for readers' convenience.

Let $F = F_n$ be the distribution function of W_n/σ_n, and $G = G_{n,r}$ be the function whose Fourier transform is $e^{-t^2/2}(1 + Q_{n,r}(t))$, where $Q_{n,r}$ comes from Proposition 23. Then $G_{n,r}$ has the form

$$G_{n,r}(t) = \Phi(t) + \sum_{j=1}^{r} \sigma_n^{-j} P_{j,n}(t)\phi(t)$$

where $P_{j,n}$'s are the Edgeworth polynomials of W_n.

Let $\varepsilon > 0$ and $B = 1/\varepsilon$. Applying (4.9) with $F = F_n$, $G = G_n$ and $T = B\sigma_n^r$ we obtain

$$\sup_t \left| P(W_n/\sigma_n \le t) - \Phi(t) - \sum_{j=1}^{r} \sigma_n^{-j} P_{j,n}(t)\phi(t) \right| \le I_1 + I_2 + I_3 + O(\varepsilon)\sigma_n^{-r}$$

where for δ small enough

$$I_1 = \int_{-\delta\sigma_n}^{\delta\sigma_n} \left| \frac{\mathbb{E}[e^{itW_n/\sigma_n}] - e^{-t^2/2}(1 + Q_{r,n}(t))}{t} \right| dt$$

$$I_2 = \int_{\delta\sigma_n \le |t| \le B\sigma_n^r} \left| \frac{\mathbb{E}[e^{itW_n/\sigma_n}]}{t} \right| dt, \quad I_3 = \int_{|t| \ge \sigma_n \delta} e^{-t^2/2} \left| \frac{1 + Q_{r,n}(t)}{t} \right| dt.$$

By Proposition 23 we have $I_1 = o(\sigma_n^{-r})$, (4.3) gives that $I_2 = o(\sigma_n^{-r})$, while $I_3 = O(e^{-c\sigma_n^2})$ for some $c > 0$ since $Q_{r,n}$ is a polynomial with bounded coefficients and degree depending only on r. □

5. Application to uniformly elliptic inhomogeneous Markov chains

5.1. **Verification of Assumption 17.** In this section we consider uniformly bounded additive functional S_N of a Markov chain X_n which satisfies (2.1) and (2.2). We prove the following.

24. **Proposition.** *The sequence of random variables S_n verifies Assumption 17 for every k, namely, if $\Lambda_n(t) = \ln \mathbb{E}[e^{itS_n/\sigma_n}] + t^2/2$ then for every $k \ge 3$ there exist constants $\delta_k, C_k > 0$ so that for all n,*

$$\sup_{t \in [-\sigma_n\delta_k, \sigma_n\delta_k]} |\Lambda_n^{(k)}(t)| \le C_k \sigma_n^{-(k-2)}.$$

The proof of Proposition 24 is based on the construction of sequential pressure functions, as described in the following section.

25. **Remark.** In [51, Theorem 4.26] the authors show that if $S_n = \sum_{j=1}^{n} Y_j/\sigma_n$, and $\{Y_j\}$ is an exponentially fast ϕ-mixing uniformly bounded centered Markov chain, such that $\mathrm{Var}(Y_j)$ is bounded away from 0 then there is a constant C such that for all $m \in \mathbb{N}$ $|\Gamma_j(S_n/\sigma_n)| \le C^m m! \sigma_n^{-(m-2)}$. It follows that the function Λ_n is real analytic and, hence, Assumption 17 holds for every k. By [20, Proposition 1.22], the Markov chains $\{X_n\}$ considered in this paper are also exponentially fast ϕ-mixing, however, we consider functionals $Y_n = f_n(X_n, X_{n+1})$ whose variance can be small, and so Proposition 24 cannot be derived from [51, Theorem 4.26] despite the related setup.

5.2. **The sequential pressure function. Definition and basic properties.** Recall Theorem 9. For every $j \ge 1$, denote by μ_j the distribution of X_j (which is a probability measure on $\mathcal{X}_j$). Recall that $\lambda_j(z)$ is uniformly bounded in j and $\lambda_j(0) = 1$. Let $\Pi_j(z)$

denote the analytic branch of the logarithms of $\lambda_j(z)$, such that $\Pi_j(0) = 0$. We call $\Pi_j(z)$ the *sequential pressure functions.* Then

$$\sup_j \sup_{|z|\leq s_0} |\Pi_j(z)| \leq c_0 \tag{5.1}$$

where s_0 and c_0 are some positive constants. We note that all the derivatives of Π_j at $z = 0$ are real numbers, since the function $\lambda_j(z)$ is positive for real z's.

26. **Remark.** By Remark 10, upon replacing f_n with $f_n - \mathbb{E}[f_n(X_n, X_{n+1})]$, the resulting pressure function becomes $\Pi_j(z) - \mathbb{E}[f_n(X_n, X_{n+1})]z$. This has no affect on the value of the pressure function at $z = 0$ and on the derivatives of it of any order larger than 1. Thus, it will essentially make no difference in the following arguments if we have already centralized f_n or not.

Let j, n be positive integers. Set

$$\Gamma_{j,n}(z) = \ln \mathbb{E}[e^{zS_{j,n}}], \quad \Pi_{j,n}(z) = \sum_{s=j}^{j+n-1} \Pi_s(z)$$

where $S_{j,n}$ is defined in (3.9).

27. **Lemma.** *There is a constant $a > 0$ with the following property: for every integer $k \geq 0$ there exists $c_k > 0$ such that for each j, n for all complex z so that $|z| \leq a$ we have*

$$\left|\Gamma_{j,n}^{(k)}(z) - \Pi_{j,n}^{(k)}(z)\right| \leq c_k \tag{5.2}$$

where $g^{(k)}(z)$ denotes the k-th derivative of a function $g(z)$.

Note that for $k = 0, 1, 2$ and $z = 0$ we have $\Gamma'_{j,n}(0) = \mathbb{E}[S_{j,n}]$, $\Gamma''_{j,n}(0) = \text{Var}(S_{j,n})$ while for larger k's $\Gamma_{j,n}^{(k)}(0)$ is just the k-th cumulant of $S_{j,n}$. In particular,

$$\Pi'_{1,n}(0) = \mathbb{E}(S_n) + O(1) \quad \text{and} \quad \Pi''_{1,n}(0) = \sigma_n^2 + O(1).$$

Proof. Since $h_j(0) = \mathbf{1}$ and the norms $\|h_j^{(z)}\|_\infty$ are uniformly bounded in j around 0, it follows from the Cauchy integral formula that $\frac{\partial h_j}{\partial z}$ is uniformly bounded around the origin. Hence, if δ_0 is small enough then for any complex z with $|z| \leq \delta_0$ we have

$$\frac{1}{2} < \inf_j |\mu_1(h_1^{(z)})|. \tag{5.3}$$

Recall that $\mathbb{E}[e^{zS_{j,n}}] = \mu_j(R_z^{j,n}\mathbf{1})$. By (3.3), if $|z|$ is sufficiently small then for all j and n we have

$$\mathbb{E}[e^{zS_{j,n}}] = e^{\sum_{s=j}^{j+n-1} \Pi_s(z)}(\mu_j(h_j^{(z)}) + \delta_{j,n}(z)) \tag{5.4}$$

where $\delta_{j,n}$ is an analytic function so that $|\delta_{j,n}(z)| \leq C\delta^n$ for some $C > 0$ and $\delta \in (0, 1)$ which do not depend on j and n. In fact, since $h_j^{(0)} = \mathbf{1}$ we have $\delta_{j,n}(0) = 0$ and so Cauchy integral formula also implies $|\delta_{j,n}(z)| \leq C|z|\delta^n$. Using (5.3), we can take the logarithms of both sides of (5.4) and derive that when $|z|$ is sufficiently small, there is a constant c_0 so that

$$\left|\Gamma_{j,n}(z) - \Pi_{j,n}(z)\right| \leq c_0 \tag{5.5}$$

Applying the Cauchy integral formula once more we conclude that for each k there exists a constant $c_k > 0$ so that for every j and n we have

$$\left|\Gamma_{j,n}^{(k)}(z) - \Pi_{j,n}^{(k)}(z)\right| \leq c_k \tag{5.6}$$

and the lemma follows. $\square$

5.3. **The derivatives of the pressure function around the origin.** Here we prove several useful auxiliary estimates.

28. **Lemma.** *Let $k \geq 2$ be an integer, and let S be a real-valued random variable with finite first k moments. Let us define $\varphi(t) = \mathbb{E}\left(e^{itS}\right)$ and $\Lambda(t) = \ln \varphi(t)$. Then there exists a constant D_k which depends only on k so that with $r_0 = \dfrac{1}{2\sqrt{\mathbb{E}(S^2)}}$*

$$\sup_{t\in[-r_0,r_0]} |\Lambda^{(k)}(t)| \leq D_k \mathbb{E}[|S|^k].$$

Proof. We first recall that for the characteristic function $\varphi(t) = \mathbb{E}\left(e^{itS}\right)$ of a random variable S with finite first k moments and any real t we have

$$|\varphi(t) - 1| \leq |t|\mathbb{E}[|S|] \leq |t|\|S\|_{L^2}$$

and that for $j = 0, 1, 2, ..., k$ we have

$$|\varphi^{(j)}(t)| \leq \mathbb{E}[|S|^j]. \tag{5.7}$$

Next, let $\Lambda(t) = \ln \varphi(t)$ and $r_0 = \dfrac{1}{2\sqrt{\mathbb{E}(S^2)}}$. Then $|\varphi(t)| \geq \frac{1}{2}$ for all $t \in [-r_0, r_0]$. By Faá di Bruno's formula, for every $t \in [-r_0, r_0]$ we have

$$|\Lambda^{(k)}(t)| = \left|\sum_{m_1,...,m_k} \frac{k!}{\prod_{j=1}^k (m_j!(j!)^{m_j})} \cdot \frac{1}{\varphi(t)^{\sum_{j=1}^k m_j}} \prod_{j=1}^k \left((i)^j \mathbb{E}[S^j e^{itS}]\right)^{m_j}\right|$$

where $(m_1, ..., m_k)$ range over all the k-tuples of nonnegative integers such that $\sum_j jm_j = k$. Now the lemma follows from (5.7) and the Hölder inequality. $\square$

29. **Lemma.** *Fix some integer $k \geq 2$ and let $B_1 < B_2$ be constants. Then if B_1 is sufficiently large there are constants D and r_0 depending only on B_1, B_2 and k so that for every $t \in [-r_0, r_0]$ and each $j, n \in \mathbb{N}$ such that $B_1 \leq Var(S_{j,n}) \leq B_2$, we have*

$$|\Pi_{j,n}^{(k)}(it)| \leq D.$$

Proof. Let $\Lambda_{j,n}(t) = \ln \mathbb{E}[e^{itS_{j,n}}]$. Then, in the notation of Lemma 27, $\Lambda_{j,n}(t) = \Gamma_{j,n}(it)$. Applying Lemma 28 with $S = S_{j,n}$ and using (5.6) and Lemma 14 we obtain that for every $t \in [-r_0, r_0]$ we have

$$|\Pi_{j,n}^{(k)}(it)| \leq c_k + |\Lambda_{j,n}^{(k)}(t)| \leq c_k + D_k \mathbb{E}[|S_{j,n}|^k] \leq c_k + C\big(\mathrm{Var}(S_{j,n})\big)^{k/2} \leq c_k + CB_2^{k/2}$$

competing the proof of the lemma. $\square$

30. **Corollary.** *For every $k\geq 2$ there exist constants $\varepsilon_k>0$ and $C_k>0$ so that for each $n\in\mathbb{N}$ and $t\in[-\varepsilon_k,\varepsilon_k]$,*

$$|\Pi_{1,n}^{(k)}(it)|\leq C_k\sigma_n^2.$$

Hence, with $\tilde\Pi_n(t)=\Pi_{1,n}(it/\sigma_n)$ he have

$$\sup_{t\in[-\varepsilon_k\sigma_n,\varepsilon_k\sigma_n]}|\tilde\Pi_n^{(k)}(t)|\leq C_k\sigma_n^{-(k-2)}.$$

Proof. Fix some $k\geq 2$. Let B_1 and B_2 be large constants so that Lemma 29 holds. Let r_0 be the constant specified in Lemma 29. Let $I_1,I_2,...,I_{m_n}$ be disjoint intervals whose union cover $\{1,...,n\}$ so that

$$B_1\leq \mathrm{Var}(S_{I_l})\leq B_2$$

where for each l we set $S_{I_l}=\sum_{j\in I_l}f_j(X_j,X_{j+1})$. Note that it is indeed possible to find such intervals if B_1 and B_2/B_1 are sufficiently large because of Theorem 11. Indeed, with u_n^2 denoting the structural constants appearing there, there are constants $C_1,C_2>0$ so that for any $n\geq 3$ and j,

$$C_1^{-1}\sum_{m=j}^{j+n-1}u_m^2-C_2\leq \mathrm{Var}(S_{j,n})\leq C_1\sum_{m=j}^{j+n-1}u_m^2+C_2. \tag{5.8}$$

It is also clear that m_n/σ_n^2 is uniformly bounded away from 0 and ∞ (if n is large enough). Now, by Lemma 29 there are $\varepsilon_k>0$ and $A_k>0$ so that for each $1\leq l\leq m_n$ and $t\in[-\varepsilon_k,\varepsilon_k]$,

$$\left|\sum_{j\in I_l}\Pi_j^{(k)}(it)\right|\leq A_k.$$

Hence, $|\Pi_{1,n}(it)|\leq \sum_l\left|\sum_{j\in I_l}\Pi_j^{(k)}(it)\right|\leq A_km_n\leq C_k\sigma_n^2.$ □

5.4. **Verification of Assumption 17.**

Proof of Proposition 24. Since both sides of (5.4) with $j=1$ are analytic, $|\delta_{1,n}(z)|\leq C|z|\delta^n$ for some $\delta\in(0,1)$ and $C>0$. Moreover $\mu_1(h_1^{(0)})=1$. Hence, if $|z|$ is small enough then

$$\ln\mathbb{E}[e^{zS_n}]=\Pi_{1,n}(z)+G_n(z)$$

where $G_n(z)=\ln\big(\mu_1(h_1^{(z)})+\delta_{1,n}(z)\big)$, which is an analytic and uniformly bounded function around the origin (uniformly in n). Thus Proposition 24 follows from Corollary 30. □

31. **Corollary.** *Let $r\geq 1$. Suppose that for any $B>0$ and $\delta>0$ small enough,*

$$\int_{\delta\leq|x|\leq B\sigma_n^{r-1}}|\mathbb{E}(e^{ixS_n})/x|dx=o(\sigma_n^{-r}). \tag{5.9}$$

Then

$$(5.10)\qquad \sup_t \left| \mathbb{P}((S_n - \mathbb{E}[S_n])/\sigma_n \le t) - \Phi(t) - \sum_{j=1}^{r} \sigma_n^{-j} P_{j,n}(t)\phi(t) \right| = o(\sigma_n^{-r})$$

where Φ and ϕ are the standard normal distribution and density function, respectively, and $P_{j,n}(t) = P_j(t, \hat{S}_n)$ are the Edgeworth polynomials of $\bar{S}_n = S_n - \mathbb{E}[S_n]$.

Corollary 31 follows from Proposition 19 since S_n verifies Assumption 17.

5.5. **A Berry-Esseen theorem and Expansions of order** 1.

Proof of Theorems 1 and 2. First, Theorem 1 follows from Propositions 24 and 18.

Next, applying Theorem 3.5 and inequality (4.2.7) from [20] we see that if $\{f_n\}$ is irreducible then condition (5.9) with $r = 1$ is satisfied. This proves the result. □

6. High order expansions for summands with small essential supremum, proof of Theorem 4 and 5

6.1. **Existence of expansions.** Recall (3.7). In order to prove Theorem 4, we need the following:

32. **Lemma.** [20, eq. (3.3.7)] $\exists \delta > 0$ *s.t. if* $\|f_n\|_\infty |\xi| \le \delta$ *then* $d_n^2(\xi) \ge \dfrac{\xi^2 u_n^2}{2}$.

Proof of Theorem 4. Let us fix some $r < \frac{1}{1-2\beta}$, and take some $r < r_0 < \frac{1}{1-2\beta}$. We claim that there are constants $c, C > 0$ so that for all N large enough we have

$$|\Phi_N(\xi)| \le \exp\left(-c\xi^2 V_N\right) \text{ for } |\xi| \le C\sigma_N^{r_0-1}.$$

This is enough for the Edgeworth expansion of order r to hold by Corollary 31.

In order to prove the claim, let $N_0 = N_0(N)$ be the smallest positive integer such that $\sigma_N^{r_0-1}\|f_n\|_\infty \le \delta$ for all $n > N_0$ where δ is the number from Lemma 32. Then, since $\|f_n\| = O(n^{-\beta})$

$$(6.1)\qquad N_0 = O\left(\sigma_N^{\frac{r_0-1}{\beta}}\right) = O\left(V_N^{\frac{r_0-1}{2\beta}}\right).$$

Let us show now that $N_0 = o(N)$, which in particular yields that $N_0 < N/2$ if N is large enough. The assumption that $\|f_n\|_\infty = O(n^{-\beta})$ also implies that $u_n^2 = O(n^{-2\beta})$ and so by (3.6),

$$(6.2)\qquad V_N = O(N^{1-2\beta}).$$

Combining this with $r_0 < \frac{1}{1-2\beta}$ we see that $\sigma_N^{\frac{r_0-1}{\beta}} = O(N^\kappa)$, where

$$(6.3)\qquad \kappa = \frac{(r_0-1)}{2\beta}(1-2\beta) = 1 - \frac{1 - r_0(1-2\beta)}{2\beta} < 1.$$

Therefore, $N_0 = O(N^\kappa)$.

Next, let us write

$$\sum_{n=N_0+1}^{N} u_n^2 = \sum_{k=0}^{3} \sum_{\substack{N_0<n\leq N,\\ n \bmod 4=k}} u_n^2 := \sum_{k=0}^{3} U_{N_0,N,k}.$$

Let k_N be so that $U_{N_0,N,k_N} = \max\{U_{N_0,N,k} : 0 \leq k \leq 3\}$. Then by (3.6) there are constants $C, D > 0$ so that

$$V(S_N - S_{N_0}) \leq CU_{N_0,N,k_N} + D. \tag{6.4}$$

Combining (3.8), Lemma 32, and (6.4) we see that the characteristic function of S_N satisfies

$$|\Phi_N(\xi)| \leq \exp\left(-c\xi^2 V(S_N - S_{N_0})\right) \text{ for } |\xi| \leq C\sigma_N^{r_0-1} \tag{6.5}$$

where $C > 0$ is some constant which depends on β, r_0, and ε_0 but not on ξ or N. Note that by Lemma 13 we have

$$V_N = V_{N_0} + V(S_N - S_{N_0}) + 2\text{Cov}(S_{N_0}, S_N - S_{N_0}) = V_{N_0} + V(S_N - S_{N_0}) + O(1).$$

It follows that

$$V(S_N - S_{N_0}) = V_N - V_{N_0} + O(1).$$

On the other hand, by (6.2),

$$V_{N_0} \leq N_0^{1-2\beta} \leq C' V_N^{\kappa}$$

where κ is given by (6.3). Therefore $V(S_N - S_{N_0}) = V_N + O\left(V_N^{\kappa}\right)$. Combining this with (6.5) gives

$$|\Phi_N(\xi)| \leq \exp\left(-c\xi^2 (V_N + O(V_N^{\kappa}))\right) \text{ for } |\xi| \leq C\sigma_N^{r_0-1}$$

and the claim follows since $\kappa < 1$. □

6.2. **Optimality.**

Proof of Theorem 5. Fix some $0 < \beta < 1/2$, and take an integer $s > \frac{1}{1-2\beta}$. Then

$$s_\beta := (s-1)\left(\frac{1}{2} - \beta\right) > \beta.$$

Take $c \in (\beta, s_\beta)$. Set $q_n = 2^{[c \log_2 n]}$ and $p_n = [n^{-\beta} q_n]$. Let

$$a_n = \frac{p_n}{q_n}.$$

Since $c > \beta$ we have

$$n^{-\beta}(1 + o(1)) = n^{-\beta} - 2^{-[c \log_2 n]} \leq a_n \leq n^{-\beta}.$$

Let Y_n be an iid sequence so that $P(Y_n = \pm 1) = \frac{1}{2}$. Set

$$X_n = a_n Y_n = \frac{p_n}{q_n} Y_n.$$

Then, $\mathbb{E}[X_n] = 0$, $|X_n| = a_n \asymp n^{-\beta}$ and $V(X_n) = a_n^2 \asymp n^{-2\beta}$. Next, since q_n divides q_N if $n \leq N$ we have

$$q_N S_N = S_N 2^{[c \log_2 N]} \in \mathbb{Z}$$

and so the minimal jump of S_N is at least $\frac{1}{q_N}$. Therefore, if α_N is a possible value of S_N then

$$\mathbb{P}\big(S_N \in (\alpha_N, \alpha_N + \frac{1}{2} 2^{-[c \log_2 N]}]\big) = 0.$$

On the other hand, if S_N obeyed an expansion of order s then, choosing $\alpha_N = O(\sigma_N)$ and denoting $\varepsilon_N = 2^{-[c \log_2 N]} \sigma_N^{-1}$, we would get

$$0 = \mathbb{P}\big(S_N \in (\alpha_N, \alpha_N + \frac{1}{2} 2^{-[c \log_2 N]}]\big) = \mathbb{P}\big(S_N/\sigma_N \in (\alpha_N/\sigma_N, \alpha_N/\sigma_N + \varepsilon_N]\big)$$

$$= \mathbb{P}\big(S_N/\sigma_N \le \alpha_N/\sigma_N + \varepsilon_N\big) - \mathbb{P}\big(S_N/\sigma_N \le \alpha_N/\sigma_N\big)$$

$$= \Phi(\alpha_N/\sigma_N + \varepsilon_N) - \Phi(\alpha_N/\sigma_N)$$

$$+ \frac{1}{\sqrt{2\pi}} \sum_{j=1}^{s} \left(P_{j,N}(\alpha_N/\sigma_N + \varepsilon_N) e^{-\frac{1}{2}(\alpha_N/\sigma_N + \varepsilon_N)^2} - P_{j,N}(\alpha_N/\sigma_N) e^{-\frac{1}{2}\alpha_N^2 \sigma_N^{-2}} \right) \sigma_N^{-j}$$

$$+ o(\sigma_N^{-s}) \ge C\varepsilon_N + o(\sigma_N^{-s}) \ge C' 2^{-c \log_2 N} \sigma_N^{-1} + o(\sigma_N^{-s}).$$

Since σ_N^2 if of order $\displaystyle\sum_{n=1}^{N} n^{-2\beta} \asymp N^{1-2\beta}$ we must have

$$c > \frac{(s-1)(1-2\beta)}{2} = s_\beta$$

which contradicts that $c \in (\beta, s_\beta)$. Taking $s = s(\beta)$ to be the smallest integer such that $s > \frac{1}{1-2\beta}$ we see that the expansions of orders $r > \frac{1}{1-2\beta}$ do not hold. □

7. High order expansions for Hölder continuous functions on Riemannian manifolds.

7.1. **Distribution of Hölder functions.** The following estimate plays an important role in the proof of Theorem 6.

33. **Lemma.** *For every Riemanian manifold $\mathcal{X}$ there is a constant $\mathfrak{c}$ such that for each real-valued function φ on $\mathcal{X}$ with $\|\varphi\|_\alpha \le 1$ and each t, ε*

$$\nu(\varphi \in [t, t+\varepsilon]) \ge \mathfrak{c} \varepsilon^{1/\alpha} \min(\nu(\varphi \ge t + \varepsilon), \mu(\varphi \le t))$$

where ν is the normalized Riemannian volume on $\mathcal{X}$.

Proof. Since $\mathcal{X}$ is compact, it can be covered by a finite number of coordinate charts. Hence for any given ε' we can cover $\mathcal{X}$ by the C^r images of coordinate cubes of size ε' so that the multiplicity of the cover is bounded by a constant $\mathfrak{K}$ which is independent of ε'.

Now, let $\varepsilon' = \delta \varepsilon^{1/\alpha}$ where δ is so small that the diameter of each partition element is smaller than $\varepsilon^{1/\alpha}/2$. Consider the cover of $\mathcal{X}$ described above and let A be the union of all cover elements Q such that $\varphi(x) \ge t + \frac{\varepsilon}{2}$ for each $x \in Q$ and $\mathcal{S}$ be the union of all partition elements which intersect ∂A. By the Isoperimetric Inequality,

$$\text{Area}(\partial A) \ge \frac{\mathfrak{h}}{\mathfrak{K}} \min(\nu(A), \nu(A^c)) \ge \frac{\mathfrak{h}}{\mathfrak{K}} \min(\nu(\varphi \ge t + \varepsilon), \nu(\varphi \le t))$$

where $\mathfrak{h}$ is the Cheeger constant of $\mathcal{X}$. On the other hand, there exists a constant κ which does not depend on ε or α so that

$$\text{Area}(\partial A) \leq \text{Area}(\partial \mathcal{S}) \leq \mathfrak{K}\kappa\varepsilon^{1/\alpha}\nu(\mathcal{S})$$

since for each cover element $Q \subset \mathcal{S}$ we have

$$\text{Area}(\partial S \bigcap \partial Q) \leq \kappa\varepsilon^{1/\alpha}\nu(Q).$$

Since $\varphi \in [t, t+\varepsilon]$ on $\mathcal{S}$ the result follows. □

7.2. **Proof of Theorem 6.** For the rest of Section 7 we consider the following setting. Let $\{X_n\}$ evolve on a compact Riemannian manifold M with transition densities $p_n(x,y)$ bounded and bounded away from 0. Let us assume that $f_n : M \times M \to \mathbb{R}$ satisfy $\|f_n\|_\alpha := \max(\sup|f_n|, v_\alpha(f_n)) \leq 1$ for some $0 < \alpha \leq 1$. Denote $\Phi_N(\xi) = \mathbb{E}(e^{\xi S_N})$.

34. **Proposition.** *For all $0 < \alpha \leq 1$ and $\delta > 0$ there exists $C_1(\alpha,\delta), c_1 = c_1(\alpha,\delta) > 0$ so that for every $n \in \mathbb{N}$ and $\xi \in \mathbb{R}$ with $|\xi| \geq \delta$ we have*

$$|\Phi_N(\xi)| \leq C_1 \exp\left(-c_1 V_N |\xi|^{1-\frac{1}{\alpha}}\right).$$

Theorem 6 follows by Proposition 34 together with Corollary 31.

The main step in the proof of Proposition 34 is the following.

35. **Lemma.** *For every Riemanian manifold $\mathcal{X}$ for every $\delta > 0$ there is a constant $\hat{c}$ such that for each real-valued function φ on $\mathcal{X}$ with $\|\varphi\|_\alpha \leq 1$ and each ξ such that $|\xi| \geq \delta$,*

$$\iint \sin^2\left(\frac{\xi[\varphi(x_1) - \varphi(x_2)]}{2}\right)\nu(x_1)d\nu(x_2) \leq \hat{c}|\xi|^{1-(1/\alpha)} \iint [\varphi(x_1) - \varphi(x_2)]^2 \nu(x_1)d\nu(x_2).$$

where ν is the normalized Riemannian volume on $\mathcal{X}$.

The lemma will be proven in §7.3, here we complete the proof of the proposition based on the lemma.

Let μ denote the normalized Riemannian volume on M. Let us fix some $n \in \mathbb{N}$ and consider a random hexagon $P_n = (x_{n-2}, x_{n-1}, x_n; y_{n-1}, y_n, y_{n+1})$ based at n.

Recall (3.5) and (3.7). By uniform ellipticity we have

$$u_n^2 \asymp \int \Gamma^2(P_n) d\mu^6(P_n), \quad d_n^2(\xi) \asymp \int \sin^2\left(\frac{\xi\Gamma(P_n)}{2}\right) d\mu^6(P_n). \tag{7.1}$$

where

$$\Gamma(P_n) = f_{n-2}(x_{n-2}, x_{n-1}) + f_{n-1}(x_{n-1}, x_n) + f_n(x_n, y_{n+1})$$
$$-f_{n-2}(x_{n-2}, y_{n-1}) - f_{n-1}(y_{n-1}, y_n) - f_n(y_n, y_{n+1})$$

is the balance of P_n.

Applying Lemma 35 with $\mathcal{X} = M \times M$ and

$$\phi_{x_{n-2}, y_{n+1}}(x_{n-1}, x_n) = f_{n-2}(x_{n-2}, x_{n-1}) + f_{n-1}(x_{n-1}, x_n) + f_n(x_n, y_{n+1})$$

and integrating with respect to x_{n-2} and y_{n+1} we obtain $d_n^2(\xi) \geq C\xi^{1-(1/\alpha)} u_n^2$.

Now Proposition 34 follows from (3.8) □

7.3. **The proof of Lemma 35.** Set

$$\Delta(x_1,x_2)=|\varphi(x_1)-\varphi_n(x_2)|, \quad \varepsilon=\xi^{-1}, \quad \mathfrak{u}^2=\iint \Delta^2(x_1,x_2)\nu(x_1)d\nu(x_2),$$

$$\mathfrak{d}^2(\xi)=\iint \sin^2\left(\frac{\Delta(x_1,x_2)}{2\varepsilon}\right)\nu(x_1)d\nu(x_2).$$

Decompose $\mathcal{X}\times\mathcal{X}=\mathcal{A}_1\cup\mathcal{A}_2$ where $\mathcal{A}_1=\left\{(x_1,x_2):\Delta\leq\frac{\varepsilon}{8}\right\}$ and $\mathcal{A}_2$ is its complement. We split the proof of Lemma 35 into two cases.

Case 1. If the integral of Δ^2 over $\mathcal{A}_1$ is larger than the integral over $\mathcal{A}_2$ then using that $\left|\frac{\sin t}{t}\right|\geq c$ for $|t|\leq 1/8$ we get

$$\mathfrak{d}^2(\xi)\geq\iint_{\mathcal{A}_1}\sin^2\frac{\Delta(x_1,x_2)}{2\varepsilon}d\nu(x_1)d\nu(x_2)\geq\frac{c^2\xi^2}{4}\mathfrak{u}^2.$$

Case 2. Now we assume that the integral over $\mathcal{A}_2$ is larger. Let

$$l_k=2^k\varepsilon, \quad k^*=\operatorname{argmax}\left[l_k(\nu\times\nu)(\Delta\in[l_k,2l_k))\right], \quad \mathfrak{l}=l_{k^*}$$

and

$$\mathfrak{v}=\mathfrak{l}(\nu\times\nu)(\Delta\in[\mathfrak{l},2\mathfrak{l})).$$

Note that under the assumptions of Case 2 we have

$$\mathfrak{u}^2\leq C_0\sum_{k=-3}^{\log_2(1/\varepsilon)}\sum_k l_k^2(\nu\times\nu)(\Delta\in[l_k,2l_k))\leq C_0\mathfrak{v}\sum_{k=-3}^{\log_2(1/\varepsilon)}l_k\leq C\mathfrak{v}. \tag{7.2}$$

Next, let $\mathfrak{m}$ denote a median of φ with respect to ν, $\tilde\varphi=\varphi-\mathfrak{m}$ and

$$\Omega_1=\{\tilde\varphi\leq\mathfrak{l}/2\}, \quad \Omega_2=\{\tilde\varphi\in(-\mathfrak{l}/2,\mathfrak{l}/2)\}, \quad \Omega_3=\{\tilde\varphi\geq\mathfrak{l}/2\}.$$

Let us assume that $\mu(\Omega_3)\geq\mu(\Omega_1)$, the case where the opposite inequality holds being similar. Since $\Delta(x_1,x_2)<\mathfrak{l}$ for $(x_1,x_2)\in\Omega_2\times\Omega_2$ we have

$$(\nu\times\nu)(\Delta\geq\mathfrak{l})\leq 2\left[\nu(\Omega_1)+\nu(\Omega_3)\right]\leq 4\nu(\Omega_3).$$

Let

$$\Omega_j'=\{\tilde\varphi\in[(j+0.1)\varepsilon,(j+0.2)\varepsilon]\} \quad \Omega_j''=\{\tilde\varphi\in[(j+0.3)\varepsilon,(j+0.4)\varepsilon]\}.$$

Since $\mathfrak{m}$ is a median, $\nu(\Omega_1\cup\Omega_2)\geq\frac{1}{2}$. Hence Lemma 33 shows that that for $j\leq\frac{\mathfrak{l}}{4\varepsilon}$ we have

$$\nu(\Omega_j')\geq c\varepsilon^{1/\alpha}\nu(\Omega_3), \quad \nu(\Omega_j'')\geq c\varepsilon^{1/\alpha}\nu(\Omega_3). \tag{7.3}$$

On the other hand there is a constant $\delta_0>0$ such that for each $x_1\in X$ we have that $\sin^2\left(\frac{\Delta(x_1,x_2)}{2\varepsilon}\right)\geq\delta_0$ either for all j and all $x_2\in\Omega_j'$ or for all j for all $x_2\in\Omega_j''$. It follows that if $\mathcal{A}_2$ dominates then

$$\mathfrak{d}^2(\xi)\geq\delta_0\min\left(\sum_{j=1}^{\mathfrak{l}/4\varepsilon}\nu(\Omega_j'),\sum_{j=1}^{\mathfrak{l}/4\varepsilon}\nu(\Omega_j'')\right)\geq\hat c\mathfrak{l}\mu(\Omega_3)=\tilde c\varepsilon^{1/\alpha-1}\mathfrak{l}(\nu\times\nu)(\Delta\in[\mathfrak{l},2\mathfrak{l}))=\tilde c\varepsilon^{1/\alpha-1}\mathfrak{v}.$$

Combining this with (7.2) we obtain that if $\mathcal{A}_2$ dominates then $\mathfrak{d}^2(\xi) \geq c\varepsilon^{1/\alpha-1}\mathfrak{u}^2$.

Combining the estimates of cases 1 and 2 we obtain the result. $\square$

7.4. **Cantor functions.** In order to show the optimality of Theorem 6 we need to consider a function f for which the estimate of Lemma 35 is optimal. Moreover, we want f to grow on a set of small Hausdorff dimension and we want the distribution of f to have atoms at values which are commensurable with each other. It turns out that Cantor functions studied in [32, 22] satisfy these conditions. So in this subsection we describe briefly the construction and properties of Cantor functions.

Let us fix some integers $p \geq 3$, $k \geq 1$ and let $q = (p-1)k$. Set

$$\alpha_{p,p+q} = \frac{1}{\log_p(q+p)} = \frac{\ln p}{\ln(p+q)}.$$

On $[0,1]$, let $C_{p,p+q}$ (where $q = (p-1)k$) be the Cantor set of all numbers of the form $x = \sum_{j=1}^{\infty} \frac{(k+1)a_j}{(p+q)^j}, a_j = 0, 1, ..., p-1$. In other words $C_{p,p+q}$ consists of all number in $[0,1]$ which can be written in base $p+q$ so that all its digits are divisible by $k+1$.

Let f be the corresponding Cantor function ([32]). Namely, for $x \in C_{p,p+q}$ we have

$$f(x) = \sum_j \frac{a_j}{p^j}, \quad \text{if} \quad x = \sum_j \frac{(k+1)a_j}{(p+q)^j},$$

while outside $C_{p,p+q}$ we have

$$f(x) = \sup_{y \in C_{p,p+q},\, y \leq x} f(y) = \sum_{j=1}^{n} \frac{b_j}{p^j} \quad \text{where} \quad x = \sum_j \frac{x_j}{(p+q)^j}, \quad b_j = \left[\frac{x_j}{k+1}\right] + 1$$

and n is the first index so that x_n is not divisible by $k+1$. By [32, Theorem 2] (see also [22]), f is Hölder continuous with exponent $\alpha_{p,q}$, which is also the the Hausdorff dimension of $C_{p,q+p}$. Note that f is increasing (see [32, Theorem 1]) and that $f(0) = 0$ and $f(1) = 1$.

36. **Lemma.** *For each $n \in \mathbb{N}$*

$$\text{Leb}\{x \in [0,1] : p^n f(x) \notin \mathbb{Z}\} = \left(\frac{p}{p+q}\right)^n. \tag{7.4}$$

Proof. To prove the lemma we explain the inductive construction of f by following the recursive construction of the set $C_{p,q+p}$. First, we split $[0,1]$ into $p+q$ closed intervals $I_1, I_2, ..., I_{p+q}$ of the same length $\frac{1}{p+q}$ so that I_s is to the left of I_{s+1} for each s. Next, define intervals $J_1, J_2, ..., J_{2p+1}$ as follows: we define $J_1 = I_1$, and then inductively $J_{2l+1} = I_{s_l+k+1}$, if $J_{2l-1} = I_{s_l}$. For $1 \leq l < p$ we define and J_{2l} to be the union of the intervals I_s between J_{2l-1} and J_{2l+1}. On J_{2l} we define $f|_{J_{2l}} = \frac{l}{p}$.

The reconstruction of the function f now proceeds by induction. Suppose that at the n-th step of the construction f was additionally defined on a union of closed intervals $U_1, ..., U_{j_n}$, $j_n = (p-1)p^{n-1}$ of length $k(p+q)^{-n}$ so that $f|_{U_j} = jp^{-n}$, U_j is to the left of U_{j+1}, and the gap between U_j and U_{j+1} is $(p+q)^{-n}$, where $U_0 = \{0\}$ and

$U_{j_n+1} = \{1\}$. Let us split the interval between U_j and U_{j+1} into equal $p+q$ intervals $I_{1,j,n+1}, I_{2,j,n+1}, ..., I_{p+q,j,n+1}$ of length $(p+q)^{-n-1}$ so that $I_{s,j,n+1}$ is to the left of $I_{s+1,j,n+1}$ for each s. In the $(n+1)$-th step the intervals $J_{1,j,n+1}, J_{2,j,n+1}, ..., J_{2p+1,j,n+1}$ are defined as follows: we define $J_{1,j,n+1} = I_{1,j,n+1}$, and then inductively $J_{2l+1,j,n+1} = I_{s_l+k+1,j,n+1}$, if $J_{2l-1,j,n+1} = I_{s_l,j,n+1}$. For $1 \leq l < p$ we define and $J_{2l,j,n+1}$ to be the union of the intervals $I_{s,j,n+1}$ between $J_{2l-1,j,n+1}$ and $J_{2l+1,j,n+1}$. On $J_{2l,j,n+1}$ we define

$$f|_{J_{2l,j,n+1}} = \frac{jp+l}{p^{n+1}} = \frac{j}{p^n} + \frac{l}{p^{n+1}}.$$

In view of the above recursive construction of f, we obtain (7.4) since in the $(n+1)$-th step there are p^n intervals of length $(p+q)^{-n}$ on which f has not been defined yet, and the values of f in all the steps proceeding the n-th step do not have the form s/p^n for $s \in \mathbb{Z}$. □

7.5. **Optimality.**

Proof of Theorem 7. We first observe that it is enough to prove Theorem 7 for a dense set of numbers α in $(0,1)$. Indeed, if the theorem holds for α belonging to a dense set A, given $\alpha_0 \in (0,1)$ and $r > \frac{\alpha_0+1}{1-\alpha_0}$, we can find $\alpha \in A$ so that $\alpha > \alpha_0$ and $r > \frac{\alpha+1}{1-\alpha}$. Now, the α-Hölder continuous function we get from Theorem 7 with this α is also α_0-Hölder continuous so the result follows.

Next, let us consider the set

$$A = \left\{ \frac{\ln p}{\ln(p+q)} : p, q \in \mathbb{N}, p \geq 3, \ q|(p-1) \right\}.$$

This set is dense in $(0,1)$. Indeed, let $0 < a < b < 1$. Then, using that $\frac{\ln p}{\ln(q+p)} = \frac{1}{\log_p(q+p)}$, for all $p \geq 3$ and denoting $k = \frac{q}{(p-1)}$, $\quad k \in \mathbb{N}$ we have

$$\frac{1}{\log_p(q+p)} \in (a,b) \iff p^{1/b-1} < k+1-\frac{1}{p} < p^{1/a-1}.$$

Since $\lim_{p\to\infty} p^{1/a-1} - p^{1/b-1} = \infty$, we can find a number k satisfying the above inequality provided that p is large enough.

Thus we fix some integers $p \geq 3$, $k \geq 1$ and let $q = (p-1)k$. Set

$$\alpha = \alpha_{p,p+q} = \frac{1}{\log_p(q+p)} = \frac{\ln p}{\ln(p+q)}.$$

Let $f : [-1,1] \to [-1,1]$ be the odd function whose restriction to $[0,1]$ is the Cantor function from §7.4. We will now show that $S_n f$ does not obey Edgworth expansions of any order $r > \frac{\alpha+1}{\alpha-1}$. Let $r = r(\alpha)$ be the smallest integer so that $r > \frac{\alpha+1}{\alpha-1}$, where $\alpha = \alpha_{p,q}$. Let us take $\frac{\alpha}{1-\alpha} < c < \frac{r}{2} - \frac{1}{2}$ and set $k_N = p^{[c\log_p N]}$. Then

$$\mathbb{P}(k_N S_N \notin \mathbb{Z}) \leq N\mathbb{P}(k_N f \notin \mathbb{Z}) = N\left(\frac{p}{p+q}\right)^{[c\log_p N]} = O\left(N^{1-[(1/\alpha)-1]c}\right) = o_{N\to\infty}(1)$$

where the second step follows from Lemma 36 and the last step follows since

$$c\left(\frac{1}{\alpha}-1\right)=\frac{c(1-\alpha)}{\alpha}>1.$$

Let $p_N=p^{[c\log_p N]}\sigma_N=k_N\sigma_N$ which is of order $N^{c+1/2}$. Then

$$\lim_{N\to\infty}\mathbb{P}(S_N/\sigma_N\in(p_N)^{-1}\mathbb{Z})=1.$$

Thus, by considering points in $(p_N)^{-1}\mathbb{Z}$ which are of order 1, we find that if C is large enough then denoting

$$m_N=\text{argmax}\{\mathbb{P}(S_N/\sigma_N=k/p_N):|k/p_N|\leq C\}$$

and recalling that $c+\frac{1}{2}>r$ we have

$$\mathbb{P}(S_N/\sigma_N=m_N/p_N)\geq C_1p_N^{-1}\geq C_2N^{-c-1/2}\geq C_3\sigma_N^{-r} \tag{7.5}$$

where C_1,C_2 and C_3 are positive constants. On the other hand, if S_N obeyed expansions of order r then

$$\mathbb{P}\left(\frac{S_N}{\sigma_N}=\frac{m_N}{p_N}\right)\leq\limsup_{\delta\to0^+}\left[\mathbb{P}\left(\frac{S_N}{\sigma_N}\leq\frac{m_N}{p_N}\right)-\mathbb{P}\left(\frac{S_N}{\sigma_N}\leq\frac{m_N}{p_N}-\delta\right)\right]=o(\sigma_N^{-r})$$

which is inconsistent with (7.5). □

References

[1] J. Angst, G. Poly, *A weak Cramér condition and application to Edgeworth expansions*, Electron. J. Probab. **22** (2017) # 59, 1-24.

[2] P. Baldi, Y. Rinott *On normal approximations of distributions in terms of dependency graphs,* Ann. Prob. **17** (1989) 1646–1650.

[3] A.D. Barbour, *Asymptotic expansions based on smooth functions in the central limit theorem*, Probab. Th. Rel. Fields **72** (1986) 289–303.

[4] A.D. Barbour and L.H.Y. Chen, *An Introduction to Stein's Method*, Singapore University Press, Singapore, 2005.

[5] V. Bentkus, F. Gotze, W. Van Zwet *An Edgeworth Expansion for Symmetric Statistics.* Ann. Stat. **25** (1997) 851–896.

[6] W. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Trans. AMS **49** (1941) 122–136.

[7] R.N. Bhattacharya, R. Ranga Rao *Normal Approximation and Asymptotic Expansions,* Wiley, New York-London-Sydney-Toronto (1976) xiv+274 pp.

[8] P. Bickel P., F. Gotze, W. Van Zwet *The Edgeworth Expansion for U-Statistics of Degree Two,* Ann. Stat. **14** (1986) 1463–1484.

[9] E. Breuillard, *Distributions diophantiennes et theoreme limite local sur* $\mathbb{R}^d$, Probab. Th. Rel. Fields **132** (2005) 39–73.

[10] H. Callaert, P. Janssen, *The Berry-Esseen theorem for U statistics*, Ann. Stat. **6** (1978) 417–421.

[11] H. Callaert, P. Janssen, N. Veraverbeke *An Edgeworth Expansion for U-Statistics.* Ann. Stat. **8** (1980) 299–312.

[12] L.H.Y Chen, Q.M. Shao, *Normal approximation under local dependence*, Ann. Probab. **32** (2004) 1985–2028.

[13] Z. Coelho, W. Parry *Central limit asymptotics for shifts of finite type*, Israel J. Math. **69** (1990) 235–249.

[14] H. Cramer *On the composition of elementary errors,* Skand. Aktuarietid skr. **1** (1928) 13–74; 141–180.

[15] R. Dobrushin *Central limit theorems for non-stationary Markov chains I, II.* Theory Probab. Appl. **1** (1956) 65–80, 329–383.

[16] D. Dolgopyat *On decay of correlations in Anosov flows,* Ann. Math. **147** (1998) 357–390.

[17] D. Dolgopyat *Prevalence of rapid mixing in hyperbolic flows,* Erg. Th. Dynam. Sys. **18** (1998) 1097–1114.

[18] D. Dolgopyat, K. Fernando *An error term in the Central Limit Theorem for sums of discrete random variables,* preprint.

[19] D. Dolgopyat, Y. Hafouta *Edgeworth expansion for indepenent bounded integer valued random variables,* arXiv:2011.14852.

[20] D. Dolgopyat, O. Sarig, *Local limit theorems for inhomogeneous Markov chains*, arXiv:2109.05560.

[21] H. Döring, P. Eichelsbacher, *Moderate deviations via cumulants*, J. Theor. Probab. **26** (2013) 360–385.

[22] O. Dovgoshey, O. Martio, V. Ryazanov, M. Vuorinen *The Cantor function*, Expositiones Math. **24** (2006) 1-37.

[23] P. Doukhan, M.H. Neumann, *Probability and moment inequalities for sums of weakly dependent random variables, with applications*, Stoch. Proc. App. **117** (2007) 878–903.

[24] D. Dragičević, Y. Hafouta, *Limit theorems for random expanding or Anosov dynamical systems and vector valued observables*, Ann. Henri Poincaré **21** (2020) 3869–3917.

[25] L. Dubois *Projective metrics and contraction principles for complex cones,* J. Lond. Math. Soc. **79** (2009) 719–737.

[26] L. Dubois *An explicit Berry-Esseen bound for uniformly expanding maps on the interval,* Israel J. Math. **186** (2011) 221–250.

[27] C.-G. Esseen *Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law,* Acta Math. **77** (1945) 1–125.

[28] C.-G. Esseen, *A moment inequality with an application to the central limit theorem"*, Skand. Aktuarietidskr. 39, (1956) 160–170.

[29] W. Feller, *An introduction to probability theory and its applications*, Vol. II., 2d edition, John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[30] K. Fernando, C. Liverani *Edgeworth expansions for weakly dependent random variables*, Ann. de l'Institut Henri Poincare Prob. & Stat. **57** (2021) 469–505.

[31] K. Fernando, F. Pène *Expansions in the local and the central limit theorems for dynamical systems*, arXiv:2008.08726.

[32] R.E. Gilman, *A class of functions continuous but not absolutely continuous*, Ann. Math. **33** (1932) 433–442.

[33] B.V. Gnedenko, A.N. Kolmogorov *Limit distributions for sums of independent random variables,* Cambridge, Addison-Wesley, 1954.

[34] W.F. Grams, R.J. Serfling, *Convergence rates for U-statistics and related statistics*, Ann. Stat. **1** (1973) 153–160.

[35] S. Gouëzel. *Berry–Esseen theorem and local limit theorem for non uniformly expanding maps*, Annales IHP Prob. & Stat. **41** (2005), 997–1024.

[36] S. Gouëzel. *Limit theorems in dynamical systems using the spectral method,* Proc. Sympos. Pure Math. **89** (2015) 161–193.

[37] Y. Guivarc'h, J. Hardy. *Théorèmes limites pour une classe de chaînes de markov et applications aux difféomorphismes d'anosov,* Annales IHP Prob. & Stat. **24** (1988), 73–98.

[38] Y. Hafouta, *On the asymptotic moments and Edgeworth expansions of some processes in random dynamical environment*, J. Stat Phys **179** (2020) 945–971.

[39] Y. Hafouta, *Limit theorems for some time dependent expanding dynamical systems*, Nonlinearity **33** (2020) 6421–6460.

[40] Y. Hafouta and Yu. Kifer, *Nonconventional limit theorems and random dynamics*, World Scientific, Singapore, 2018.

[41] P.Hall, *Edgeworth expansion for Student's t statistic under minimal moment conditions*, Ann. Probab. **15** (1987) 920–931.

[42] H. Hennion, L. Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Springer Lecture Notes in Math. **1766** (2001).

[43] L. Hervé, F. Pène, *The Nagaev–Guivarch method via the Keller-Liverani theorem,* Bull. Soc. Math. France **138** (2010) 415–489.

[44] I. A. Ibragimov, Yu. V. Linnik *Independent and stationary sequences of random variables,* Wolters-Noordhoff Publishing, Groningen, 1971. 443 pp.

[45] M. Jirak *Berry-Esseen theorems under weak dependence*, Ann. Prob. **44** (2016) 2024–2063.

[46] M. Jirak *A Berry–Esseen bound with (almost) sharp dependence conditions,* arXiv:2008.07332.

[47] Z. Lin, L. Bai, *Probability inequalities*, Springer, 2010.

[48] C. Liverani *Decay of correlations,* Ann. Math. **142** (1995) 239–301.

[49] S.V. Nagaev, *Some limit theorems for stationary Markov chains*, Theory Probab. Appl. **2** (1957) 378–406.

[50] S.V. Nagaev, *More exact statements of limit theorems for homogeneous Markov chains*, Theory Probab. Appl. **6** (1961), 62- -81.

[51] L. Saulis, V.A. Statulevicius, *Limit Theorems for Large Deviations*, Kluwer, Dordrecht, Boston, 1991.

[52] I. Shevtsova, *An Improvement of Convergence Rate Estimates in the Lyapunov Theorem*, Doklady Mathematics. 82 (3) 862–864.

[53] Yu. V. Prokhorov, *On a local limit theorem/or lattice distributions*, Dokl. Akad. Nauk SSSR, **98** (1954) 535–538.

[54] E. Rio, *Sur le théorème de Berry-Esseen pour les suites faiblement dṕendantes*, Probab. Th. Relat. Fields 104 (1996), 255-282.

[55] Rinott, Y. and Rotar V. *On Edgeworth expansions for dependency-neighborhoods chain structures and Stein's method,* Probab. Theory Rel. Fields **126** (2003) 528–570.

[56] J. Rousseau-Egele.*Un théoreme de la limite locale pour une classe de transformations dilatantes et monotones par morceaux,* Ann. Prob. **11** (1983) 772–788.

[57] Yu. A. Rozanov, *On a local limit theorem for lattice distributions*, Th. Prob. & Appl. **2** (1957) 275–281.

[58] H.H. Rugh, *Cones and gauges in complex spaces: Spectral gaps and complex Perron-Frobenius theory*, Ann. Math. **171** (2010), 1707–1752.

[59] S. Sethuraman, S. R. S. Varadhan *A martingale proof of Dobrushin's theorem for non-homogeneous Markov chains,* Electron. J. Probab. **10** (2005) paper 36, 1221–1235.

University of Maryland and The Ohio State University