

## LARGE DEVIATION PRINCIPLES FOR LACUNARY SUMS

CHRISTOPH AISTLEITNER, NINA GANTERT, ZAKHAR KABLUCHKO,  
JOSCHA PROCHNO, AND KAVITA RAMANAN

*Dedicated to the memory of Nicole Tomczak-Jaegermann*

ABSTRACT. Let  $(a_k)_{k \in \mathbb{N}}$  be an increasing sequence of positive integers satisfying the Hadamard gap condition  $a_{k+1}/a_k > q > 1$  for all  $k \in \mathbb{N}$ , and let

$$S_n(\omega) = \sum_{k=1}^n \cos(2\pi a_k \omega), \quad n \in \mathbb{N}, \omega \in [0, 1].$$

Then  $S_n$  is called a lacunary trigonometric sum, and can be viewed as a random variable defined on the probability space  $\Omega = [0, 1]$  endowed with Lebesgue measure. Lacunary sums are known to exhibit several properties that are typical for sums of independent random variables. For example, a central limit theorem for  $(S_n)_{n \in \mathbb{N}}$  has been obtained by Salem and Zygmund, while a law of the iterated logarithm is due to Erdős and Gál. In this paper we study large deviation principles for lacunary sums. Specifically, under the large gap condition  $a_{k+1}/a_k \rightarrow \infty$ , we prove that the sequence  $(S_n/n)_{n \in \mathbb{N}}$  does indeed satisfy a large deviation principle with speed  $n$  and the same rate function  $\tilde{I}$  as for sums of independent random variables with the arcsine distribution. On the other hand, we show that the large deviation principle may fail to hold when we only assume the Hadamard gap condition. However, we show that in the special case when  $a_k = q^k$  for some  $q \in \{2, 3, \dots\}$ ,  $(S_n/n)_{n \in \mathbb{N}}$  satisfies a large deviation principle (with speed  $n$ ) and a rate function  $I_q$  that is different from  $\tilde{I}$ , and describe an algorithm to compute an arbitrary number of terms in the Taylor expansion of  $I_q$ . In addition, we also prove that  $I_q$  converges pointwise to  $\tilde{I}$  as  $q \rightarrow \infty$ . Furthermore, we construct a random perturbation  $(a_k)_{k \in \mathbb{N}}$  of the sequence  $(2^k)_{k \in \mathbb{N}}$  for which  $a_{k+1}/a_k \rightarrow 2$  as  $k \rightarrow \infty$ , but for which at the same time  $(S_n/n)_{n \in \mathbb{N}}$  satisfies a large deviation principle with the same rate function  $\tilde{I}$  as in the independent case, which is surprisingly different from the rate function  $I_2$  one might naïvely expect. We relate this fact to the number of solutions of certain Diophantine equations. Together, these results show that large deviation principles for lacunary trigonometric sums are very sensitive to the arithmetic properties of the sequence  $(a_k)_{k \in \mathbb{N}}$ .

---

Received by the editors November 29, 2021, and, in revised form, June 28, 2022.

2020 *Mathematics Subject Classification*. Primary 42A55, 60F10, 11L03; Secondary 37A05, 11D45, 11K70.

*Key words and phrases*. Lacunary series, lacunary trigonometric sums, large deviations, Hadamard gap condition, large gap condition, hyperbolic dynamics, Diophantine equations, normal numbers.

The first author was supported by the Austrian Science Fund (FWF), projects F-5512, I-3466, I-4945 and Y-901. The third author was supported by the German Research Foundation under Germany's Excellence Strategy EXC 2044 – 390685587, Mathematics Münster: Dynamics - Geometry - Structure. The fourth author was supported by the Austrian Science Fund (FWF), projects P32405 and the Special Research Program F5508-N26. The fifth author was supported by the National Science Foundation (NSF) Grant DMS-1954351 and the Roland George Dwight Richardson Chair at Brown University and a Simon Guggenheim Fellowship. The authors were also supported by the Oberwolfach Research Institute for Mathematics (MFO), where they initiated their collaboration during the workshop on “New Perspectives and Computational Challenges in High Dimensions” (Workshop ID 2006b) in February 2020.

This is particularly noteworthy since no such arithmetic effects are visible in the central limit theorem or in the law of the iterated logarithm for lacunary trigonometric sums. Our proofs use a combination of tools from probability theory, harmonic analysis, and dynamical systems.

### 1. INTRODUCTION

The study of lacunary series is a classical and still flourishing topic in harmonic analysis that has attracted considerable attention. In the article [48] published in 1922, Rademacher studied the convergence behavior of series of the form

$$\sum_{k=1}^{\infty} b_k r_k(\omega),$$

where  $\omega \in [0, 1]$ ,  $b = (b_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ , and  $r_k$  denotes the  $k^{\text{th}}$  Rademacher function, that is,  $r_k(\omega) = \text{sign}(\sin(2^k \pi \omega))$ . He proved that such a series converges for almost every  $\omega \in [0, 1]$  if  $\sum_{k \in \mathbb{N}} |b_k|^2 < \infty$ , or equivalently,  $b \in \ell_2$ . The necessity of square summability was obtained shortly after by Khintchine and Kolmogorov in 1925 [35], thereby establishing an interesting  $\ell_2$ -dichotomy in the convergence behavior of such series. Note that by the structure of the Rademacher functions one has

$$\sum_{k=1}^{\infty} b_k r_k(\omega) = \sum_{k=1}^{\infty} b_k r_1(2^{k-1} \omega),$$

where on the right-hand side we have a series of dilates of a fixed function, with an exponentially growing dilation factor. This leads naturally to the study of similar questions for lacunary trigonometric series, that is, series of the form

$$\sum_{k=1}^{\infty} b_k \cos(2\pi a_k \omega) \quad \text{and} \quad \sum_{k=1}^{\infty} b_k \sin(2\pi a_k \omega),$$

where  $\omega \in [0, 1]$ ,  $b = (b_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ , and  $(a_k)_{k \in \mathbb{N}}$  is a sequence of positive integers that is lacunary, in the sense that it satisfies the Hadamard gap condition

$$(1.1) \quad \frac{a_{k+1}}{a_k} \geq q > 1, \quad \text{for every } k \in \mathbb{N}.$$

Interestingly, results similar to the Rademacher case were obtained for such series. Kolmogorov showed in [37] that the square summability of  $b$  is sufficient for the almost everywhere convergence of lacunary series and Zygmund proved in [60] that the square summability condition was necessary, again establishing the same  $\ell_2$ -dichotomy as for Rademacher series.

An important property of the Rademacher functions is that they form a system of independent random variables. More precisely, if  $(b_k)_{k \in \mathbb{N}}$  is a sequence of real numbers, then the weighted Rademacher functions  $b_k r_k$ ,  $k \in \mathbb{N}$ , form a sequence of independent and centered random variables with  $\text{Var}(b_k r_k) = b_k^2$ . One readily checks that Lindeberg’s condition is satisfied whenever both  $b \notin \ell_2$  and  $\max_{1 \leq k \leq n} |b_k| = o(\|(b_k)_{k=1}^n\|_2)$ . This means that under these two conditions we have, for every  $t \in \mathbb{R}$ , the central limit theorem (CLT)

$$\lim_{n \rightarrow \infty} \lambda \left( \left\{ \omega \in [0, 1] : \sum_{k=1}^n b_k r_k(\omega) \leq t \|(b_k)_{k=1}^n\|_2 \right\} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-y^2/2} dy,$$

which in particular holds when  $b_k = 1$  for every  $k \in \mathbb{N}$ . In 1939 Kac proved an analogous central limit theorem in the lacunary case for integer sequences  $(a_k)_{k \in \mathbb{N}}$  with very large gaps, that is, those for which  $a_{k+1}/a_k \rightarrow \infty$ , as  $k \rightarrow \infty$ . The case of Hadamard lacunary sequences  $(a_k)_{k \in \mathbb{N}}$  was settled in 1947 when Salem and Zygmund established in [52] that, for all  $t \in \mathbb{R}$ ,

$$(1.2) \quad \lim_{n \rightarrow \infty} \lambda \left( \left\{ \omega \in [0, 1] : \sum_{k=1}^n \cos(2\pi a_k \omega) \leq t\sqrt{n/2} \right\} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-y^2/2} dy.$$

These results suggest that lacunary trigonometric sums behave in many ways like sums of independent random variables, and in fact, this has become a classical heuristic that has been confirmed in many settings. Indeed, under the Hadamard gap condition, the sequence of scaled partial sums of the functions  $\cos(2\pi a_k x)$ ,  $k \in \mathbb{N}$ , not only satisfies the central limit theorem in (1.2), but, as Salem and Zygmund [53] and Erdős and Gál [23] showed, also satisfies a law of the iterated logarithm (LIL), that is, for almost every  $\omega \in [0, 1]$ ,

$$\limsup_{n \rightarrow \infty} \frac{\sum_{k=1}^n \cos(2\pi a_k \omega)}{\sqrt{n \log \log n}} = 1.$$

A generalization to non-integral sequences  $(a_k)_{k \in \mathbb{N}}$  was also later established in [57].

A natural question is to ask whether the CLT and LIL still hold under the Hadamard gap condition when the function  $\omega \mapsto \cos(2\pi\omega)$  is replaced by an arbitrary 1-periodic function  $f$ . A famous example of Erdős and Fortet (see, e.g., [32]) shows that this is not true in general. However, under the additional condition that the function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is of bounded variation on  $[0, 1]$  and satisfies both

$$(1.3) \quad f(\omega + 1) = f(\omega) \quad \text{and} \quad \int_0^1 f(\omega) d\omega = 0,$$

Kac was able to show in [31] that a central limit theorem holds for scaled partial sums of the functions  $\omega \mapsto f(2^k \omega)$ ,  $k \in \mathbb{N}$ , but in this case the variance of the Gaussian limit law is

$$(1.4) \quad \sigma^2 = \int_0^1 f(\omega)^2 d\omega + 2 \sum_{k=1}^{\infty} \int_0^1 f(\omega) f(2^k \omega) d\omega,$$

rather than  $\int_0^1 f(\omega)^2 d\omega$ , as one would have in the independent case, namely for the sequence of partial sums  $\sum_{k=1}^n f(2^k U_k)$ , where  $\{U_k\}_{k \in \mathbb{N}}$  are independent and identically distributed (i.i.d.) random variables distributed uniformly on  $(0, 1)$ . This shows that general lacunary function systems possess a more complicated dependence structure than lacunary trigonometric function systems, and that in the general case, the arithmetic structure of the lacunary integer sequence plays a crucial role. Gapoškin found a remarkable relation between the existence of a CLT and the number of solutions to a certain Diophantine equation [29]. It was only more recently, in 2010, that Aistleitner and Berkes improved Gapoškin’s result and provided the precise condition for the central limit theorem to hold in the general lacunary framework [4, Theorem 1.1].

While we have seen that the probabilistic behavior of lacunary series is quite well understood on the scales of both the CLT and LIL, this is not the case for large deviations. Specifically, large deviation principles (LDPs) seem to have not been studied at all in the lacunary setting. In contrast to the CLT, which captures

universal behavior in the sense that the limits are insensitive to details of the distribution beyond the first and second moments, probabilities of (large) deviations on the scale of laws of large numbers are non-universal and describe the asymptotic likelihood of rare events. More precisely, LDPs are sensitive to the distribution of the underlying random variables and their non-universality is reflected in the so-called rate function and/or the speed, which together define the asymptotic exponential decay rate of large deviation probabilities. The most classical result in this direction is Cramér’s theorem [15] (see also [16] and [17, Theorem 2.2.3]), which guarantees that if  $X, X_1, X_2, \dots$  are i.i.d. random variables with cumulant (or log-moment) generating function  $\Lambda(u) := \log \mathbb{E}[e^{uX}] < \infty$  for  $u$  in a neighborhood of zero, then one has

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(X_1 + \dots + X_n \geq nt) = -\Lambda^*(t),$$

for all  $t > \mathbb{E}[X]$ , where  $\Lambda^*$  is the Legendre-Fenchel transform of  $\Lambda$  given by

$$\Lambda^*(t) = \sup_{\theta \in \mathbb{R}} [\theta t - \Lambda(\theta)].$$

LDPs in the spirit of Donsker and Varadhan, who initiated a systematic study (see [17, 56] and the references cited therein), generalize the idea behind Cramér’s theorem. Loosely speaking, a sequence  $(X_n)_{n \in \mathbb{N}}$  of random variables in  $\mathbb{R}^d$  is said to satisfy an LDP with speed  $s_n \uparrow \infty$  and a rate function  $\mathbb{I} : \mathbb{R}^d \rightarrow [0, \infty]$  if for sufficiently large  $n \in \mathbb{N}$  and  $A \subset \mathbb{R}^d$  sufficiently regular,

$$\mathbb{P}(X_n \in A) \approx e^{-s_n \inf_{x \in A} \mathbb{I}(x)}.$$

More precisely, a sequence  $(X_n)_{n \in \mathbb{N}}$  of random variables in  $\mathbb{R}^d$  is said to satisfy an LDP with speed  $s_n$  and rate function  $\mathbb{I} : \mathbb{R}^d \rightarrow [0, \infty]$  if  $\mathbb{I} : \mathbb{R}^d \rightarrow [0, \infty]$  is lower-semicontinuous and for every Borel measurable set  $A \subset \mathbb{R}^d$ ,

$$(1.5) \quad - \inf_{x \in A^\circ} \mathbb{I}(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{s_n} \log \mathbb{P}(X_n \in A) \leq \limsup_{n \rightarrow \infty} \frac{1}{s_n} \log \mathbb{P}(X_n \in A) \leq - \inf_{x \in \bar{A}} \mathbb{I}(x),$$

where  $A^\circ$  and  $\bar{A}$ , respectively, denote the interior and closure of the set  $A$ .

In this paper we study large deviations for lacunary sums, thereby complementing existing limit theorems like the CLT and LIL mentioned above. More precisely, if  $(a_k)_{k \in \mathbb{N}}$  is a lacunary sequence, that is, a sequence of positive integers satisfying the Hadamard gap condition (1.1), we study the tail behavior of the associated sequence of lacunary sums, namely partial sums of the sequence  $X_k(\omega) := \cos(2\pi a_k \omega)$ ,  $\omega \in [0, 1]$ ,  $k \in \mathbb{N}$ , viewed as real-valued random variables on the space  $[0, 1]$  equipped with the Borel  $\sigma$ -algebra  $\mathcal{B}([0, 1])$  and Lebesgue measure  $\lambda$ . Our results reveal an interesting and surprising behavior, showing how sometimes – depending on arithmetic properties of the lacunary sequence  $(a_k)_{k \in \mathbb{N}}$  – the large deviations behavior of the associated sequence of lacunary trigonometric sums resembles that of partial sums of independent and identically distributed random variables, whereas in other situations it does not. This is particularly interesting since no such influence of the arithmetic structure of the lacunary sequence is visible under the Hadamard gap condition when considering lacunary trigonometric sums, neither in the case of the CLT nor in the case of the LIL.

We present precise statements of our main findings in the next section, with the proofs presented in the following section.

2. MAIN RESULTS

We now present the main results of this paper. Let  $U \sim \text{Unif}(0, 1)$  be a random variable with the uniform distribution on the interval  $[0, 1]$ . Given a sequence  $(a_k)_{k \in \mathbb{N}}$  of positive integers, define the random variables

$$(2.1) \quad X_k := \cos(2\pi a_k U), \quad k \in \mathbb{N},$$

and their partial sums

$$(2.2) \quad S_n := \sum_{k=1}^n X_k = \sum_{k=1}^n \cos(2\pi a_k U), \quad n \in \mathbb{N}.$$

These random variables are most conveniently defined on the probability space  $\Omega = [0, 1]$  endowed with the Borel  $\sigma$ -algebra  $\mathcal{B}([0, 1])$  and standard Lebesgue measure  $\lambda$ , which we shall sometimes also denote by  $\mathbb{P}$ . As a function on  $\Omega$ ,  $X_k$  is then given by  $X_k(\omega) = \cos(2\pi a_k \omega)$ , for  $\omega \in [0, 1]$  and  $k \in \mathbb{N}$ . Note that the random variables  $X_1, X_2, \dots$  are identically distributed and (if all  $a_k, k \in \mathbb{N}$ , are distinct) uncorrelated. To see that the correlations vanish when  $(a_k)_{k \in \mathbb{N}}$  are distinct, recall that  $\cos(\alpha) \cdot \cos(\beta) = 2^{-1}[\cos(\alpha - \beta) + \cos(\alpha + \beta)]$  and hence, whenever  $k \neq \ell$ , we have

$$\begin{aligned} & \int_0^1 \cos(2\pi a_k \omega) \cos(2\pi a_\ell \omega) \, d\omega \\ &= \frac{1}{2} \int_0^1 \cos(2\pi(a_k - a_\ell)\omega) \, d\omega + \frac{1}{2} \int_0^1 \cos(2\pi(a_k + a_\ell)\omega) \, d\omega = 0. \end{aligned}$$

However, the elements of the sequence  $(X_k)_{k \in \mathbb{N}}$  are not independent and in fact, the sequence is in general not even stationary.

**2.1. Behavior as in the independent case.** Our aim is to prove LDPs for the sequence  $(S_n/n)_{n \in \mathbb{N}}$ . It is natural to try to compare the behavior of  $S_n/n$  to the behavior of partial sums of *independent* random variables with the same distribution as  $X_1$ , the common distribution of  $X_k, k \in \mathbb{N}$ . To this end, consider the random variables

$$(2.3) \quad \tilde{X}_k := \cos(2\pi U_k), \quad k \in \mathbb{N},$$

where  $(U_k)_{k \in \mathbb{N}}$  are i.i.d. random variables with the same distribution as  $U$ , and define their partial sums

$$(2.4) \quad \tilde{S}_n := \sum_{k=1}^n \tilde{X}_k, \quad n \in \mathbb{N}.$$

By Cramér’s classical theorem (see, e.g., [17, Theorem 2.2.3]),  $(\tilde{S}_n/n)_{n \in \mathbb{N}}$  satisfies an LDP with speed  $n$  and rate function  $\tilde{I} : \mathbb{R} \rightarrow [0, +\infty]$  given by the Legendre-Fenchel transform of the logarithmic moment generating function, that is,

$$(2.5) \quad \tilde{I}(x) = \sup_{\theta \in \mathbb{R}} \left[ \theta x - \tilde{\Lambda}(\theta) \right],$$

where

$$(2.6) \quad \tilde{\Lambda}(\theta) := \log \mathbb{E}[e^{\theta \tilde{X}_1}], \quad \theta \in \mathbb{R}.$$

The function  $\tilde{\Lambda}$  can be computed explicitly. The common distribution of the random variables  $\tilde{X}_k, k \in \mathbb{N}$ , is the arcsine law on  $(-1, 1)$  with Lebesgue density

$$f(x) = \frac{1}{\pi\sqrt{1-x^2}}, \quad |x| < 1.$$

The moment generating function of  $\tilde{X}_1$  is accordingly given by

$$\begin{aligned} \mathbb{E}[e^{\theta\tilde{X}_1}] &= \int_{-1}^1 e^{\theta x} \frac{1}{\pi\sqrt{1-x^2}} dx = \sum_{m=0}^{\infty} \frac{\theta^{2m}}{(2m)!} \int_{-1}^1 \frac{x^{2m}}{\pi\sqrt{1-x^2}} dx \\ &= \sum_{m=0}^{\infty} \frac{\theta^{2m}}{(2m)!} \frac{\Gamma(m+1/2)}{\Gamma(m+1)\sqrt{\pi}} \\ &= \sum_{m=0}^{\infty} \frac{\theta^{2m}}{(2m)!} \frac{(2m)!}{2^{2m}m!m!} \\ &= \sum_{m=0}^{\infty} \frac{\theta^{2m}}{2^{2m}(m!)^2}. \end{aligned} \tag{2.7}$$

Note that the right-hand side equals the modified Bessel function  $B_0(\theta)$  of the first kind. When combined, the above calculations yield

$$\tilde{\Lambda}(\theta) = \log \sum_{m=0}^{\infty} \frac{\theta^{2m}}{2^{2m}(m!)^2}, \quad \theta \in \mathbb{R}. \tag{2.8}$$

Since  $\tilde{X}_1$  is supported on the interval  $[-1, 1]$ , the function  $\tilde{I}$  equals  $+\infty$  outside  $[-1, 1]$ . Moreover, the asymptotics of the modified Bessel function  $B_0$  given in [1, p. 377, 9.7.1] imply that

$$\tilde{\Lambda}(\theta) = \theta - \frac{1}{2} \log(2\pi\theta) + O\left(\frac{1}{\theta}\right), \quad \text{as } \theta \rightarrow +\infty,$$

which, after taking the Legendre-Fenchel transform, yields that  $\tilde{I}(\pm 1) = +\infty$ . On the interval  $(-1, 1)$ , the function  $\tilde{I}$  is finite.

Now, let us finally turn to the partial sums  $(S_n)_{n \in \mathbb{N}}$  defined in (2.2). Our first result states that when  $(a_k)_{k \in \mathbb{N}}$  satisfies the so-called ‘‘large gap condition’’, the associated sequence of lacunary sums  $(S_n/n)_{n \in \mathbb{N}}$  satisfies an LDP with the same speed and the same rate function  $\tilde{I}$  as in the truly independent case, that is, as  $(\tilde{S}_n/n)_{n \in \mathbb{N}}$ .

**Theorem A.** *Suppose that  $(a_k)_{k \in \mathbb{N}}$  is a sequence of positive integers that satisfies the ‘‘large gap condition’’*

$$\frac{a_{k+1}}{a_k} \rightarrow \infty \quad \text{as } k \rightarrow \infty.$$

*Then  $(S_n/n)_{n \in \mathbb{N}}$  satisfies an LDP with speed  $n$ , and rate function  $\tilde{I}$ .*

The proof of Theorem A is given in Section 3.2, after an instructive special case has been treated in Section 3.1.

*Remark 2.1.* In this paper, we discuss only sequences that satisfy Hadamard’s gap condition. If  $(a_k)_{k \in \mathbb{N}} \subset \mathbb{N}$  is a sublacunary sequence, that is, increasing and

$\log(a_k)/k \rightarrow 0$  as  $k \rightarrow \infty$ , then for  $z \in (0, 1)$ , we argue below that

$$(2.9) \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left( \left\{ \omega \in [0, 1] : \frac{1}{n} \sum_{k=1}^n \cos(2\pi a_k \omega) \geq z \right\} \right) = 0,$$

which says that in contrast to the lacunary case, these probabilities decay slower than exponentially in  $n$ . To show (2.9), fix  $z \in (0, 1)$  and choose  $\delta = \delta(z) > 0$  such that  $\cos(2\pi\omega) \geq z$  for  $|\omega| \leq \delta$ . Then,  $\cos(2\pi a_k \omega) \geq z$  for all  $k \in \{1, \dots, n\}$  if  $|\omega| \leq \delta/a_n$ . But  $\mathbb{P} \left( \omega \leq \frac{\delta}{a_n} \right) = \frac{\delta}{a_n}$  and

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left( \left\{ \omega \in [0, 1] : \frac{1}{n} \sum_{k=1}^n \cos(2\pi a_k \omega) \geq z \right\} \right) \\ & \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} (\{ \omega \in [0, 1] : \cos(2\pi a_k \omega) \geq z, \forall k \in \{1, 2, \dots, n\} \}) \\ & \geq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left( \omega \leq \frac{\delta}{a_n} \right) \\ & = 0, \end{aligned}$$

where the last equality uses the assumption that  $(a_k)_{k \in \mathbb{N}}$  is sublacunary. Since the opposite inequality follows trivially, this proves (2.9).

**2.2. The case of geometric progressions**  $a_k = q^k$ . Let us now consider the case when there exists  $q \in \{2, 3, \dots\}$  such that  $a_k = q^k$  for  $k \in \mathbb{N}$ . Contrary to the case of a large gap condition (see Theorem A), we now obtain LDPs with the same speed  $n$ , but with rate functions that are different from  $\tilde{I}$ , and depend on the particular value of the growth factor  $q$ . Our main findings in this case are summarized in Theorem B, whose proof is given in Section 3.3. Let us point out that the main aspect of Theorem B is not the LDP itself, which follows from standard results (see, e.g., [8, 45, 59]), but rather the properties of the rate function which are specified in (i)–(iv).

**Theorem B.** *Fix  $q \in \{2, 3, \dots\}$ . Let  $a_k = q^k$  for  $k \in \mathbb{N}$ , and let  $S_n$  be the partial sum defined in (2.2). Then the following limit exists:*

$$(2.10) \quad \Lambda_q(\theta) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{\theta S_n}],$$

with the convergence holding uniformly for  $\theta$  in compact subsets of an open set  $\mathcal{D}$  in the complex plane such that  $\mathbb{R} \subset \mathcal{D}$ . Moreover,  $(S_n/n)_{n \in \mathbb{N}}$  satisfies an LDP with speed  $n$  and rate function  $I_q$ , which is the Legendre-Fenchel transform of  $\Lambda_q$ , that is,

$$I_q(x) = \sup_{\theta \in \mathbb{R}} [\theta x - \Lambda_q(\theta)], \quad x \in \mathbb{R}.$$

Furthermore, each  $I_q$  satisfies  $I_q(x) > 0$  for  $x \neq 0$  and  $I_q$  is equal to  $+\infty$  outside  $[-1, 1]$ , and the family of rate functions  $I_q, q \in \{2, 3, \dots\}$ , has the following properties:

- (i) For every  $q \in \{2, 3, \dots\}$ , we have  $I_q(1) \leq \tilde{I}(1)$  and  $I_q(x) < \tilde{I}(x)$  for  $x \in (0, 1)$ , where  $\tilde{I}$  is defined in (2.5). In particular, the functions  $I_q$  and  $\tilde{I}$  are different.
- (ii) The limit  $\lim_{q \rightarrow \infty} I_q(x) = \tilde{I}(x)$  holds uniformly on compact subsets of the interval  $(-1, 1)$ .

- (iii) *There is a smooth transition of  $I_q$  towards  $\tilde{I}$  as  $q \rightarrow \infty$ , in the following sense: for all  $m \in \{1, \dots, q\}$ , we have*

$$\left(\frac{d}{d\theta}\right)^m \Lambda_q(\theta)\Big|_{\theta=0} = \left(\frac{d}{d\theta}\right)^m \tilde{\Lambda}(\theta)\Big|_{\theta=0} = \kappa_m(\tilde{X}_1),$$

where  $\tilde{\Lambda}$  is defined in (2.6),  $\tilde{X}_1$  in (2.3),  $\kappa_m(\tilde{X}_1)$  is the  $m^{\text{th}}$  cumulant of  $\tilde{X}_1$ . Further,

$$\left(\frac{d}{dx}\right)^m I_q(x)\Big|_{x=0} = \left(\frac{d}{dx}\right)^m \tilde{I}(x)\Big|_{x=0}.$$

In other words, the coefficients of  $1, \theta, \dots, \theta^q$  in the Taylor expansions of  $I_q(\theta)$  and  $\tilde{I}(\theta)$  coincide at the origin.

- (iv) *Whereas in (iii) the first  $q$  derivatives of  $\Lambda_q$  and  $\tilde{\Lambda}$  coincide, this is no longer true for the derivative of order  $q + 1$ . In particular,*

$$\left(\frac{d}{d\theta}\right)^{q+1} \Lambda_q(\theta)\Big|_{\theta=0} = \left(\frac{d}{d\theta}\right)^{q+1} \tilde{\Lambda}(\theta)\Big|_{\theta=0} + \frac{q+1}{2^q} > \left(\frac{d}{d\theta}\right)^{q+1} \tilde{\Lambda}(\theta)\Big|_{\theta=0}.$$

Indeed, we will see in Proposition 2.4 that not only is  $I_q \neq \tilde{I}$  but it is also true that  $I_{q_1} \neq I_{q_2}$  if  $q_1 \neq q_2$ .

We comment on Theorem B. As mentioned above, the fact that  $(S_n/n)_{n \in \mathbb{N}}$  satisfies an LDP will be deduced from general results on thermodynamic formalism and expanding maps of the interval  $[0, 1]$  (see the proof of the theorem in Section 3.3). The key takeaways of the theorem are the properties of the rate functions in (i)–(iv). Parts (ii) and (iii) state that the rate functions  $I_q$  converge towards  $\tilde{I}$  as  $q \rightarrow \infty$ , which is in accordance with the “limiting case” of Theorem A where the ratio of  $a_{k+1}/a_k$  diverges to  $+\infty$ , and where the rate function for the lacunary sums coincides with the one for the truly independent case. The rate functions  $\tilde{I}$  and  $I_q$  for small  $q$  are illustrated in Figure 1.

Note also that as a consequence of conclusion (i) of Theorem B, the probability of large deviations of the lacunary sums  $S_n$  is (asymptotically) greater than the large deviation probability for the corresponding partial sum of independent random variables  $\tilde{S}_n$  defined in (2.4). However, since the statement (i) only applies to positive values of  $x$ , this conclusion is only valid for large *positive* deviations of the lacunary sum. In the case of large *negative* deviations there seems to be an interesting dichotomy. When  $q$  is *odd*, then the lacunary sums have a distribution symmetric around 0, which is a consequence of the fact that the mapping  $\omega \mapsto \omega + 1/2 \pmod{1}$  transforms the function  $\cos(2\pi q^k \omega)$  into  $\cos(2\pi q^k(\omega + 1/2)) = -\cos(2\pi q^k \omega)$ . Accordingly, the probabilities of large positive and large negative deviations are equal, and we have  $I_q(-x) = I_q(x)$ , so that in the odd case we have  $I_q(x) \leq \tilde{I}(x)$  for all  $x \neq 0$ , with  $I_q(x) < \tilde{I}(x)$  for all sufficiently small  $|x|$  due to (iv). In contrast, when  $q$  is *even*, there is no such symmetry. In fact, for even  $q$ , it follows from (iii) and (iv) that  $\Lambda_q(\theta) < \tilde{\Lambda}(\theta)$  for  $\theta < 0$  sufficiently close to 0 (because the coefficient of  $\theta^{q+1}$ , an odd power, in the Taylor series of  $\Lambda_q(\theta)$  is larger than that of  $\tilde{\Lambda}(\theta)$ , while the smaller powers coincide). By taking the corresponding Legendre-Fenchel transforms, it follows that  $I_q(x) > \tilde{I}(x)$  for  $x < 0$  sufficiently close to 0. We believe the above inequalities hold without restricting  $|x|$  to be sufficiently small, as stated in Conjecture 2.2.



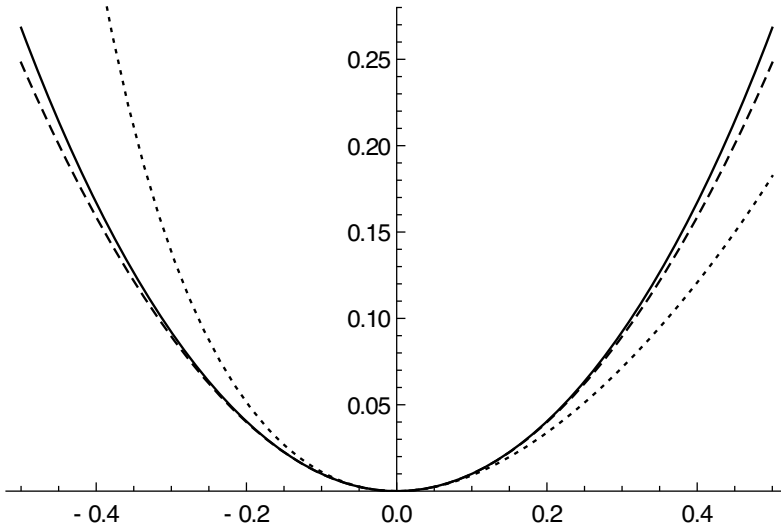


FIGURE 1. The rate function  $\tilde{I}$  (solid line), corresponding to the case of independent random variables, in comparison with the rate functions  $I_2$  (dotted) and  $I_3$  (dashed), whose existence is established in Theorem B. The plot is based on the Taylor approximation of the rate functions as given by Proposition 2.4. The rate functions  $I_4, I_5, \dots$  are not plotted, as they would be indistinguishable from  $\tilde{I}$  on the plot. The plot supports the heuristics that  $I_q$  should approach  $\tilde{I}$  as  $q \rightarrow \infty$ , as indeed turns out to be the case due to part (ii) of our Theorem B. Note also that  $\tilde{I}$  and  $I_3$  are symmetric, while  $I_2$  is not (cf. the discussion before the statement of Conjecture 2.2).

**Conjecture 2.2.** *Let  $q \in \{2, 3, \dots\}$  and let  $a_k = q^k$  for  $k \in \mathbb{N}$ . Let  $I_q$  be the rate function in the LDP for  $(S_n/n)_{n \in \mathbb{N}}$  (which exists by Theorem B). Then, if  $q$  is odd,*

$$I_q(x) < \tilde{I}(x) \quad \text{for all } x \in (-1, 1) \setminus \{0\}.$$

*On the other hand, if  $q$  is even, then*

$$I_q(x) < \tilde{I}(x) \quad \text{for all } 0 < x < 1 \quad \text{and} \quad I_q(x) > \tilde{I}(x) \quad \text{for all } -1 < x < 0.$$

As argued above, we have  $\tilde{I}(\pm 1) = +\infty$ . Since  $|S_n/n| \leq 1$ , it is clear that  $I_q(x) = +\infty$  for  $|x| > 1$ . Lemma 2.3 whose proof is given at the end of Section 3.3, states that  $I_q(+1)$  is finite.

**Lemma 2.3.** *For all  $q \in \{2, 3, \dots\}$  we have  $I_q(+1) \leq \log q$ .*

As explained above, for odd  $q$  we have  $I_q(-1) = I_q(+1)$ . For even  $q$ , it remains unclear whether  $I_q(-1)$  is finite (and in fact, it is not even clear whether  $I_q(x)$  is finite for all  $-1 < x < 0$ ).

The functions  $\Lambda_q$  and  $I_q$  appearing in Theorem B are not really explicit. In fact, the only known formula for  $\Lambda_q$  seems to be its representation as the logarithm of the largest eigenvalue of a certain Perron-Frobenius operator (see the proof of

Theorem B in Section 3.3). Proposition 2.4 identifies the first few terms in the Taylor expansions of  $I_q$  for  $q \in \{2, 3, 4\}$ . Before stating it, let us look at the Taylor series of the rate function  $\tilde{I}$ . From the expression for  $\tilde{I}$  in (2.5) and properties of the Legendre-Fenchel transform, it follows that the derivative  $\tilde{I}'$  of  $\tilde{I}$  is the inverse function of the derivative  $\tilde{\Lambda}'$  of  $\tilde{\Lambda}$ , see [50, Corollary 23.5.1, p. 219]. Using this fact together with the expression for  $\tilde{\Lambda}$  in (2.8), which yields the series expansion

$$(2.11) \quad \tilde{\Lambda}(\theta) = \frac{\theta^2}{4} - \frac{\theta^4}{64} + \frac{\theta^6}{576} - \frac{11\theta^8}{49152} + \frac{19\theta^{10}}{614400} - \frac{473\theta^{12}}{106168320} + O(\theta^{14}), \quad \text{as } \theta \rightarrow 0,$$

one can easily compute the first few terms in the Taylor series of  $\tilde{I}$  near 0:

$$(2.12) \quad \tilde{I}(z) = z^2 + \frac{z^4}{4} + \frac{5z^6}{36} + \frac{19z^8}{192} + \frac{143z^{10}}{1800} + \frac{1769z^{12}}{25920} + O(z^{14}), \quad \text{as } z \rightarrow 0.$$

**Proposition 2.4.** *In the case when  $a_k = 2^k$  for all  $k \in \mathbb{N}$ , the Taylor expansion of the rate function around 0 is given by*

$$I_2(z) = z^2 - z^3 + \frac{3z^4}{2} - \frac{13z^5}{6} + \frac{29z^6}{9} - \frac{23z^7}{5} + \frac{1127z^8}{180} - \frac{29083z^9}{3780} + \frac{12077z^{10}}{1575} + O(z^{11}).$$

*In the case when  $a_k = 3^k$  for all  $k \in \mathbb{N}$ , the rate function satisfies*

$$I_3(z) = z^2 - \frac{z^4}{12} + \frac{z^6}{6} - \frac{39z^8}{320} + \frac{18113z^{10}}{100800} + O(z^{12}),$$

*whereas when  $a_k = 4^k$  for all  $k \in \mathbb{N}$ , we have*

$$I_4(z) = z^2 + \frac{z^4}{4} - \frac{z^5}{12} + \frac{5z^6}{36} + O(z^7).$$

*In particular, the functions  $I_2, I_3, I_4$ , and  $\tilde{I}$  all differ from each other.*

In fact, in the proof of Proposition 2.4, which is deferred to Appendix A, we describe an algorithm to compute an arbitrary number of terms in the Taylor expansion of  $I_q$  for every  $q \in \{2, 3, \dots\}$ . The algorithm, as well as the proof of properties (i)–(iv) in Theorem B, is based on an analysis of the number of representations of 0 as a sum of  $m$  terms of the form  $\pm q^1, \pm q^2, \dots, \pm q^n$ . Denoting this number by  $A_m(n)$ , we prove in Proposition A.1 that for fixed  $m \in \mathbb{N}$ , it is a polynomial in  $n$  for all  $n \geq m - 2$ . This fact allows us to compute the first few moments of  $S_n$  and prove the above expansions.

Let us recall from (1.2) that  $(S_n/\sqrt{n})_{n \in \mathbb{N}}$  satisfies a central limit theorem under the Hadamard gap condition (1.1). Theorem C states that, perhaps surprisingly, the LDP does not hold in the same generality. More precisely, by mixing up powers of 2 and 3 we shall obtain an example of a Hadamard gap sequence  $(a_k)_{k \in \mathbb{N}}$  for which the corresponding scaled partial sums  $(S_n/n)_{n \in \mathbb{N}}$  fail to satisfy an LDP. This is stated in the following result, which is proved in Section 3.4.

**Theorem C.** *There exists a sequence of positive integers  $(a_k)_{k \in \mathbb{N}}$  satisfying  $a_{k+1}/a_k \geq q$  for some  $q > 1$  and all  $k \in \mathbb{N}$ , for which  $(S_n/n)_{n \in \mathbb{N}}$  does not satisfy an LDP with speed  $n$ . More precisely, for this sequence  $(a_k)_{k \in \mathbb{N}}$  there exists  $\bar{x}_0 \in (0, 1)$  such that for all  $x_0 \in (0, \bar{x}_0)$ ,*

$$0 < \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(S_n > nx_0) < \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(S_n > nx_0) < \infty.$$

Note that if  $(S_n/n)_{n \in \mathbb{N}}$  did satisfy an LDP with speed  $n$  and rate function  $I$ , then the last display would imply that  $\inf_{x \in (x_0, 1)} I(x) \neq \inf_{x \in [x_0, 1)} I(x)$  for all  $x_0 \in (0, \bar{x}_0)$ . In particular, each such  $x_0$  must be a jump-discontinuity, because the function

$$(0, \bar{x}_0) \ni x_0 \mapsto M(x_0) := \inf_{t \in (x_0, 1)} I(t)$$

is monotone increasing in the variable  $x_0$ . But, as a monotone function,  $M$  can only have countably many jump discontinuities. This is a contradiction.

**2.3. Randomized perturbation.** Theorem A and Theorem B might together give the impression that the existence of a limit of  $a_{k+1}/a_k$  as  $k \rightarrow \infty$  ensures an LDP and its value determines the rate function. In particular, one might be tempted to conjecture that the condition  $a_{k+1}/a_k \rightarrow \infty$  is necessary for the rate function in the LDP to coincide with  $\tilde{I}$ , the corresponding rate function for the independent case. However, this is not true. As Theorem D shows, it is possible to construct a randomly perturbed version  $(a_k)_{k \in \mathbb{N}}$  of the sequence  $(2^k)_{k \in \mathbb{N}}$  with  $\lim_{k \rightarrow \infty} a_{k+1}/a_k = 2$ , for which the corresponding sequence  $(S_n/n)_{n \in \mathbb{N}}$  almost surely satisfies an LDP with speed  $n$  and rate function  $\tilde{I}$ . This shows that a random perturbation may completely destroy the underlying dependence, at least at the large deviation scale, and, as also further elaborated in Section 2.4, rather than the asymptotic growth rate of the lacunary sequence  $(a_k)_{k \in \mathbb{N}}$ , what seems to determine the form of the rate function (when an LDP holds) is the fine arithmetic structure of  $(a_k)_{k \in \mathbb{N}}$ .

**Theorem D.** *Suppose we are given a sequence  $Y = (Y_k)_{k \in \mathbb{N}}$  of independent random variables, with each  $Y_k$  uniformly distributed on the discrete set*

$$(2.13) \quad \mathcal{D}_k := \left\{ h2^{\lceil k^{2/3} \rceil} : h \in \mathbb{Z}, 0 \leq h \leq 2^{\lceil k^{2/3} \rceil} \right\}, \quad k \in \mathbb{N},$$

*all supported on a common probability space  $(\mathcal{Y}, \mathcal{A}, \mathbb{P}_Y)$ , and an independent random variable  $U \sim \text{Unif}(0, 1)$ . Also, for each  $y \in \mathcal{D}_\infty := \otimes_{k \in \mathbb{N}} \mathcal{D}_k = \{(y_k)_{k \in \mathbb{N}} : y_k \in \mathcal{D}_k\}$ , define  $a_k^y := 2^k + y_k$  for all  $k \in \mathbb{N}$ , and let*

$$S_n^y := \sum_{k=1}^n \cos(2\pi a_k^y U), \quad n \in \mathbb{N}.$$

*Then, for  $\mathbb{P}_Y \circ Y^{-1}$ -almost every  $y \in \mathcal{D}_\infty$ , the sequence  $(S_n^y/n)_{n \in \mathbb{N}}$  satisfies an LDP with speed  $n$  and rate function  $\tilde{I}$ .*

In large deviation parlance, the LDP in Theorem D is often referred to as a “quenched LDP” since the LDP is conditional on the realization of the sequence  $y$ , and not averaged over the randomness of  $Y$ . Note however, that although the sequence  $(S_n^y/n)_{n \in \mathbb{N}}$  depends on the choice of  $y = (y_k)_{k \in \mathbb{N}}$ , the rate function  $\tilde{I}$  of the LDP (which holds for  $\mathbb{P}_Y \circ Y^{-1}$ -almost every  $y$ ) does not. The proof of Theorem D is given in Section 3.5.

*Remark 2.5.* Note that for every realization  $y$  of  $Y$  in Theorem D, we have  $2/(1 + 2^{-k^{1/3}}) \leq a_{k+1}^y/a_k^y \leq 2(1 + 2^{-(k+1)^{1/3}})$ , which implies that, as  $k \rightarrow \infty$ ,  $a_{k+1}^y/a_k^y \rightarrow 2$ . Thus, Theorem D proves that there exist lacunary sequences with proper exponential growth (as opposed to super-exponential growth as in Theorem A) that satisfy the LDP with rate function  $\tilde{I}$ .

Also, by interleaving the sequence  $(2^k)_{k \in \mathbb{N}}$  with the sequence  $(a_k^y)_{k \in \mathbb{N}}$  constructed in Theorem D (in the same way as in the proof of Theorem C), and using the fact that  $I_2$  does not coincide with  $\tilde{I}$ , it is possible to construct a sequence  $(b_k)_{k \in \mathbb{N}}$  such that as  $k \rightarrow \infty$ ,  $b_{k+1}/b_k \rightarrow 2$ , but the corresponding lacunary sums do not satisfy an LDP. This sharpens Theorem C. Somewhat surprisingly, even the randomized construction in Theorem D seems to be quite sensitive. While Theorem D can certainly be generalized in many directions, it appears to be more challenging to prove an analogue when each random component  $Y_k$  is sampled from  $\{0, 1, 2, \dots, k\}$ , or when it is sampled from  $\{0, 2^k, 2 \cdot 2^k, \dots, k \cdot 2^k\}$ . As elaborated in the next section, this can be related to the number of solutions of certain Diophantine equations (see also the proof of Theorem D in Section 3.5).

## 2.4. Concluding remarks and further open questions.

**2.4.1. Connection between LDPs and Diophantine equations.** Our results (in particular, Theorems B, C, and D together) show that only knowing that  $\lim_{k \rightarrow \infty} a_{k+1}/a_k = \eta$  for some  $\eta > 1$  does in general not allow one to determine the rate function in the LDP for the lacunary sum, or even conclude the existence of an LDP. The proofs of these results, which are presented in Section 3, often involve approximating the exponential function in  $\mathbb{E}[e^{\theta S_n}]$  via a Taylor polynomial in  $\theta$ . In turn, keeping in mind from (2.1) and (2.2) that  $S_n = \sum_{k=1}^n X_k$  is a finite sum of trigonometric functions, this entails estimates of integrals of products of trigonometric polynomials. Due to the orthogonality of the trigonometric system, calculation of these integrals leads to counting the number of solutions to certain Diophantine equations (with the  $a_k$ 's as variables). The reason why the rate function in the LDP (when it exists) for some lacunary sequences differs from the one for the independent case may be attributed to the existence of too many solutions to these Diophantine equations. For example, when  $a_k = 2^k$  for all  $k$ , then the equation  $2a_k - a_\ell = 0$  holds for many combinations of  $\ell, k$ , namely  $\ell = k + 1$  for all  $k$ . The Diophantine equations that appear in this context are always linear homogeneous Diophantine equations with integer coefficients. Thus, there are many more solutions to such equations when the sequence  $(a_k)_{k \in \mathbb{N}}$  allows many quotients  $a_\ell/a_k$  that are integers. In contrast, when the quotients  $a_\ell/a_k$  are bounded away from any integer (and any rational with a small denominator), then these Diophantine equations would have fewer solutions. Thus, while specific random perturbations such as the one chosen in Theorem D may drastically diminish the number of solutions, any generalization of Theorem D would require determining precisely how the Diophantine structure is altered by an arbitrary random perturbation, which appears to be highly non-trivial. This also suggests that there may still be some information that can be gleaned from the existence of the limit  $\lim_{k \rightarrow \infty} a_{k+1}/a_k = \eta$ , but only when  $\eta$  is a number that is not well approximated by rationals with small denominators, and when the same is true for  $\eta^2, \eta^3, \dots$ , which correspond to the limits of  $a_{k+2}/a_k, a_{k+3}/a_k$ , and so on. We formulate this as an open problem.

**Problem 2.6.** Let  $(a_k)_{k \in \mathbb{N}}$  be a lacunary sequence and assume that  $a_{k+1}/a_k \rightarrow \eta$  for a transcendental number  $\eta > 1$  (i.e.,  $\eta$  is not the root of a non-zero polynomial with integer coefficients). Is it true that  $(S_n/n)_{n \in \mathbb{N}}$  satisfies an LDP with speed  $n$  and rate function  $\tilde{I}$  (i.e., with the same rate function as in the independent case, that is, as for  $(\tilde{S}_n)_{n \in \mathbb{N}}$ )?

It may be that stronger assumptions on  $\eta$ , such as a condition on the irrationality measure of  $\eta$  and its powers, are necessary to derive the desired conclusion. However, we think that such an additional assumption should not be required. On the other hand, we believe that just assuming  $a_{k+1}/a_k \rightarrow \eta$  for an algebraic irrational  $\eta > 1$  will not be sufficient to deduce an LDP with rate function  $\tilde{I}$ .

Since the Diophantine structure of the sequence  $(a_k)_{k \in \mathbb{N}}$  plays such a key role in establishing the LDP for lacunary trigonometric sums, it would be very interesting to study this phenomenon in more detail. A natural candidate to analyze is the sequence  $a_k = 2^k + 1$ ,  $k \geq 1$ , known from the Erdős-Fortet example mentioned in Section 1. In the context of the CLT and LIL, this sequence and its generalizations have received widespread attention. The different type of behavior resulting from “pure” geometric progressions such as  $a_k = 2^k$ ,  $k \geq 1$ , on the one hand, and “perturbed” sequences such as  $a_k = 2^k + 1$ ,  $k \geq 1$ , on the other hand, can be explained analytically in terms of Fourier analysis. However, there is also a very interesting dynamical perspective, where the pure geometric progressions allow a natural interpretation as an ergodic sum, while the perturbed sequences have been interpreted as modified ergodic sums; see for example, [14, 28, 54].

**2.4.2. Normal number theory.** From a number theoretic perspective, sequences of the form  $(q^k \omega)_{k \in \mathbb{N}}$  are associated with the notion of normal numbers (in base  $q$ ), as introduced by Borel in 1909. It is well known that almost all numbers are normal in any base. The degree of normality of a number can be quantified using uniform distribution theory and discrepancy theory, which by Weyl’s criterion and the Erdős-Turán inequality naturally leads to trigonometric sums such as the ones studied in the present paper (see [20, 38] for general background on uniform distribution modulo one and discrepancy theory). LDPs for such sums can thus be viewed as quantifying the relative fraction of “non-normal” or “abnormal” numbers in a certain base, that is, numbers whose digital structure very significantly deviates from “normal” behavior. Such non-normal numbers have been intensively studied in the number theory literature, see for example [5, 41, 42]. A particularly challenging and interesting topic in normal number theory is questions concerning simultaneous normality resp. non-normality in two or more different bases (see for example [12, 47]). In terms of the large deviation problems studied in the present paper, it would be interesting to quantify the proportion of numbers that are non-normal in two or more different bases. For example, one could try to establish an LDP to estimate the probability of the set where two lacunary sums arising from the sequences  $(q^k)_{k \in \mathbb{N}}$  and  $(r^k)_{k \in \mathbb{N}}$  (for two different bases  $q, r \geq 2$ ) are both large.

In the context of normal numbers, the case of general sequences  $(a_k)_{k \in \mathbb{N}}$  satisfying  $\frac{a_{k+1}}{a_k} \in \mathbb{Z}_{>2}$ ,  $k \geq 1$ , corresponds to normality with respect to so-called Cantor expansions. This is a topic that has been pioneered by Erdős and Rényi [22, 49], and received strong attention in recent years; see for example [2, 3, 25] for recent work, and cf. also our proof of a special case of Theorem A in Section 3.1.

**2.4.3. More general lacunary sums.** We finally recall from Section 1 that the theory of lacunary trigonometric sums is structurally relatively simple in comparison with the theory of general lacunary sums, where interesting new phenomena show up even in the CLT setting. In light of this, it would be interesting to study the LDP

for

$$S_n(\omega) = \sum_{k=1}^n f(a_k\omega),$$

where  $f$  is a centered 1-periodic function (possibly satisfying some regularity assumptions). Already when  $f$  is a 2-term trigonometric polynomial (as in the Erdős-Fortet example alluded to above) there can be additional arithmetic effects in comparison to the simple case of pure trigonometric sums. It would certainly be interesting to investigate LDPs in this general lacunary setup.<sup>1</sup> A further challenging step would be to go beyond lacunary sums for a single fixed function  $f$  and investigate LDPs for the discrepancy (which is defined as a supremum over indicator functions), in the spirit of Philipp’s [46] resolution of the Erdős–Gál conjecture and Fukuyama’s [27] very precise results for the LIL for geometric progressions  $a_k = q^k$ .

### 3. PROOFS

In our proofs we will make use of the Gärtner-Ellis theorem, which we require in the following form. For a reference, see, for example, [17, Theorem 2.3.6].

**Theorem 3.1** (Gärtner-Ellis theorem). *Let  $(S_n)_{n \in \mathbb{N}}$  be a sequence of real-valued random variables. Suppose that the limit*

$$\Lambda(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{\theta S_n}]$$

*exists for all  $\theta \in \mathbb{R}$ . Assume furthermore that the function  $\theta \mapsto \Lambda(\theta)$  is differentiable for all  $\theta \in \mathbb{R}$ . Then  $(S_n/n)_{n \in \mathbb{N}}$  satisfies an LDP with speed  $n$  and convex rate function  $I$ , which can be expressed as the Legendre-Fenchel transform of  $\Lambda$ , that is,*

$$I(x) = \sup_{\theta \in \mathbb{R}} [\theta x - \Lambda(\theta)] \in \mathbb{R} \cup \{+\infty\}, \quad x \in \mathbb{R}.$$

**3.1. Proof of Theorem A in a simple special case.** We first give a proof of Theorem A in a special case, the justification being two-fold: we believe that the proof helps the intuition of the reader, but we also point out that it goes through if we replace  $\cos(2\pi \cdot)$  by any Lipschitz continuous function  $f$  that also satisfies (1.3) (i.e., is 1-periodic and centered). We consider a sequence  $(a_k)_{k \in \mathbb{N}}$  of positive integers such that  $a_1 = 1$ ,

$$m_k := \frac{a_{k+1}}{a_k} \in \{2, 3, \dots\}, k \in \mathbb{N}, \quad \text{and} \quad \lim_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} = +\infty.$$

The assumption  $a_1 = 1$  is without loss of generality, but the assumption that  $m_k, k \in \mathbb{N}$ , are integers will be essential for the following argument. By the Gärtner-Ellis theorem, it suffices to show that for all  $\theta \in \mathbb{R}$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{\theta S_n}] = \tilde{\Lambda}(\theta),$$

with  $\tilde{\Lambda}$  as defined in (2.6). To this end, we shall approximate each  $S_n$  by a random variable  $T_n$  that is easier to deal with, in the sense that it can be written as a sum of independent random variables expressed, as defined below, in terms of certain

---

<sup>1</sup>Note added during revision stage: a first major step towards a theory of LDPs for general lacunary sums was taken in a recent paper of Frühwirth, Juhos and Prochno; see [26]. They obtained a perfect analogue of Theorem A of the present paper, and a version of our Theorem B in which the rate function depends in a delicate way on an interplay of properties of the function  $f$  and the growth factor  $q$ .

conditional expectations. First, recall that the Cantor series expansion  $(\xi_k)_{k \in \mathbb{N}}$  of  $\omega \in [0, 1)$  associated with  $(m_k)_{k \in \mathbb{N}} \subset \{2, 3, \dots\}$  is given as follows:

$$\omega = \sum_{k=1}^{\infty} \frac{\xi_k(\omega)}{m_1 m_2 \cdots m_k} = \sum_{k=1}^{\infty} \frac{\xi_k(\omega)}{a_{k+1}},$$

with  $\xi_k(\omega) \in \{0, 1, \dots, m_k - 1\}$  for every  $k \in \mathbb{N}$ . This expansion was first introduced by Cantor in [10], and the investigation of its probabilistic properties appears to have been initiated by Erdős, Rényi, and Turán [22, 49, 55]. Our construction uses the following key property established by Rényi in [49] (see also (7) of [44]): the image of Lebesgue measure on  $[0, 1)$ , under the correspondence  $\omega \mapsto (\xi_k(\omega))_{k \in \mathbb{N}}$  makes  $(\xi_k)_{k \in \mathbb{N}}$  a sequence of independent integers with each  $\xi_k$  uniformly distributed on  $\{0, 1, \dots, m_k - 1\}$ .<sup>2</sup>

Given this property, now consider the filtration  $\mathcal{F}_2 \subset \mathcal{F}_3 \subset \dots$ , where the  $\sigma$ -algebras are defined by

$$\mathcal{F}_{k+1} := \sigma\left(J_{k+1,i}, i = 0, \dots, a_{k+1} - 1\right), \quad k \in \mathbb{N},$$

where for  $k \in \mathbb{N}$ ,

$$J_{k+1,i} := \left[\frac{i}{a_{k+1}}, \frac{i+1}{a_{k+1}}\right), \quad i = 0, \dots, a_{k+1} - 1.$$

We now define certain conditional expectations:

$$Y_k := \mathbb{E}[X_k | \mathcal{F}_{k+1}] = \sum_{i=0}^{a_{k+1}-1} \mathbb{E}[X_k | J_{k+1,i}] \mathbb{1}_{J_{k+1,i}}, \quad k \in \mathbb{N}.$$

In particular, we see that by construction, for every  $k \in \mathbb{N}$  and  $i = 0, \dots, a_{k+1} - 1$ ,

$$Y_k = \sum_{i=0}^{a_{k+1}-1} c_{k+1,i} \mathbb{1}_{J_{k+1,i}},$$

where  $c_{k+1,i}$ , the constant representing the value of  $Y_k$  on  $J_{k+1,i}$ , is given by

$$c_{k+1,i} := a_{k+1} \int_{J_{k+1,i}} X_k(\omega) \lambda(d\omega).$$

Since the function  $X_k$  is  $(1/a_k)$ -periodic, for any  $k \in \mathbb{N}$ ,  $c_{k+1,i} = c_{k+1,i'}$  whenever  $|i - i'|$  is a multiple of  $m_k = a_{k+1}/a_k$ . Hence, for each  $k \in \mathbb{N}$ , the random variable  $Y_k$  is only a function of  $\xi_k$ . Since the  $\{\xi_k\}_{k \in \mathbb{N}}$  are independent, the random variables  $\{Y_k\}_{k \in \mathbb{N}_0}$  are also independent.

We now show that the approximation of  $X_k$  by  $Y_k$  is sufficiently good, more precisely, for each  $k \in \mathbb{N}$  and  $i = 0, \dots, a_{k+1} - 1$ , using the mean-value theorem

---

<sup>2</sup>This is simply a generalization of the possibly more familiar result going back to Borel [7], where  $m_k = r$  for all  $k$  and the correspondence between elements of  $[0, 1]$  and their  $r$ -ary expansions maps Lebesgue measure on  $[0, 1]$  to the Bernoulli product measure on the space of  $\{0, 1, \dots, r - 1\}$ -valued sequences, with uniform marginals (see also [39, Section 2.3] for a more detailed exposition).

and the fact that  $X_k$  has Lipschitz constant  $2\pi a_k$ , we have

$$\begin{aligned} \max_{\omega \in J_{k+1,i}} |X_k(\omega) - Y_k(\omega)| &= \max_{\omega \in J_{k+1,i}} \left| X_k(\omega) - a_{k+1} \int_{J_{k+1,i}} X_k(y) dy \right| \\ &= \max_{\omega \in J_{k+1,i}} |X_k(\omega) - X_k(\omega_0)| \\ &\leq \max_{\omega \in J_{k+1,i}} 2\pi a_k |\omega - \omega_0| \\ &\leq 2\pi \frac{a_k}{a_{k+1}}, \end{aligned}$$

where  $\omega_0 = \omega_{0,k,i} \in J_{k+1,i}$  is obtained from the mean value theorem. Taking the maximum over all  $i = 0, \dots, a_{k+1} - 1$  yields

$$(3.1) \quad \|X_k - Y_k\|_\infty \leq 2\pi \frac{a_k}{a_{k+1}}.$$

In particular, this means that if  $S_n := \sum_{k=1}^n X_k$  and  $T_n := \sum_{k=1}^n Y_k$ , then

$$\|S_n - T_n\|_\infty \leq 2\pi \sum_{k=1}^n \frac{a_k}{a_{k+1}} = o(n), \quad n \rightarrow \infty,$$

because by assumption  $a_k/a_{k+1} \rightarrow 0$  as  $k \rightarrow \infty$ . For fixed  $\theta \in \mathbb{R}$  we obtain

$$\mathbb{E} [e^{\theta S_n}] = \mathbb{E} [e^{\theta T_n} e^{\theta(S_n - T_n)}] \leq e^{|\theta| \|S_n - T_n\|_\infty} \mathbb{E} [e^{\theta T_n}].$$

We also have the analogous lower bound

$$\mathbb{E} [e^{\theta S_n}] = \mathbb{E} [e^{\theta T_n} e^{\theta(S_n - T_n)}] \geq e^{-|\theta| \|S_n - T_n\|_\infty} \mathbb{E} [e^{\theta T_n}].$$

Altogether, taking into account that  $\|S_n - T_n\|_\infty = o(n)$ , we obtain

$$(3.2) \quad \mathbb{E} [e^{\theta S_n}] = e^{o(n)} \mathbb{E} [e^{\theta T_n}], \quad n \rightarrow \infty.$$

Since  $T_n = Y_1 + \dots + Y_n$  is a sum of independent random variables, it follows that

$$(3.3) \quad \frac{1}{n} \log \mathbb{E} [e^{\theta S_n}] = o(1) + \frac{1}{n} \log \mathbb{E} [e^{\theta T_n}] = o(1) + \frac{1}{n} \sum_{k=1}^n \log \mathbb{E} [e^{\theta Y_k}].$$

Similarly, in view of (3.1) and the fact that  $(X_k)_{k \in \mathbb{N}}$  are identically distributed, we have

$$\log \mathbb{E} [e^{\theta Y_k}] = o(1) + \log \mathbb{E} [e^{\theta X_k}] = o(1) + \log \mathbb{E} [e^{\theta X_1}], \quad k \rightarrow \infty.$$

Inserting this into (3.3) and recalling that the usual convergence implies convergence of arithmetic means to the same limit, the fact that  $X_1$  and  $\tilde{X}_1$  are identically distributed and the definition (2.6) of  $\tilde{\Lambda}$ , we arrive at

$$\frac{1}{n} \log \mathbb{E} [e^{\theta S_n}] = o(1) + \log \mathbb{E} [e^{\theta X_1}] = o(1) + \log \mathbb{E} [e^{\theta \tilde{X}_1}] = o(1) + \tilde{\Lambda}(\theta),$$

as desired. Since the function  $\tilde{\Lambda}$  is differentiable by (2.8), the Gärtner-Ellis theorem (reproduced as Theorem 3.1 herein) can be applied and the proof of Theorem A in the case when  $a_{k+1}/a_k$  are positive integer numbers tending to infinity is complete.



**3.2. Proof of Theorem A in full generality.** Fix  $\theta \in \mathbb{R}$  and a sufficiently small  $\varepsilon \in (0, 1)$ . As in the simple case, we wish to apply the Gärtner-Ellis theorem, but this time the analysis is more delicate. In contrast to the proof for the simple case, which relied on a reduction to the independent setting, our proof for the general case uses harmonic analysis methods and is more in the spirit of the classical works of Salem and Zygmund, Kac, and others mentioned in Section 1. Recall from (2.1) and (2.2) that

$$(3.4) \quad \mathbb{E} [e^{\theta S_n}] = \int_0^1 e^{\theta \sum_{k=1}^n \cos(2\pi a_k \omega)} d\omega = \int_0^1 \prod_{k=1}^n e^{\theta \cos(2\pi a_k \omega)} d\omega.$$

We start with an elementary lemma on the approximation of the exponential function by a Taylor polynomial of length  $d$ . For  $d \in \mathbb{N}$ , define

$$(3.5) \quad p_d(x) := \sum_{m=0}^d \frac{x^m}{m!}, \quad x \in \mathbb{R}.$$

**Lemma 3.2.** *There exists  $d : (0, 1) \rightarrow \mathbb{N}$  with  $d(\varepsilon) \rightarrow \infty$  as  $\varepsilon \rightarrow 0$  such that the polynomial  $p := p_{d(\varepsilon)}$  satisfies for every  $k \in \mathbb{N}$ ,*

$$(3.6) \quad 1 - \varepsilon \leq \frac{p(\theta \cos(2\pi a_k \omega))}{e^{\theta \cos(2\pi a_k \omega)}} \leq 1 + \varepsilon, \quad \omega \in [0, 1].$$

*Proof.* Fix  $d \in \mathbb{N}$  and  $p = p_d$ . Then, by the classical theory of Taylor approximation, the remainder in the Lagrange form satisfies

$$|e^x - p(x)| \leq \frac{e^{\max\{0, x\}}}{(d + 1)!} |x|^{d+1},$$

and hence,

$$1 - \frac{e^{\max\{0, x\}} |x|^{d+1}}{(d + 1)! e^x} \leq \frac{p(x)}{e^x} \leq 1 + \frac{e^{\max\{0, x\}} |x|^{d+1}}{(d + 1)! e^x}.$$

Noting that in our situation we have  $|\theta \cos(2\pi a_k \omega)| \leq |\theta|$ , this implies that for every  $k \in \mathbb{N}$ ,

$$(3.7) \quad 1 - \frac{e^{|\theta|} |\theta|^{d+1}}{(d + 1)!} \leq \frac{p(\theta \cos(2\pi a_k \omega))}{e^{\theta \cos(2\pi a_k \omega)}} \leq 1 + \frac{e^{|\theta|} |\theta|^{d+1}}{(d + 1)!}.$$

Recall that  $\theta$  is assumed to be fixed throughout the section. By picking  $d = d(\varepsilon) \in \mathbb{N}$  sufficiently large such that  $\frac{e^{|\theta|} |\theta|^{d+1}}{(d+1)!} < \varepsilon$ , we have  $d(\varepsilon) \rightarrow \infty$  as  $\varepsilon \rightarrow 0$ , and (3.6) follows. □

As an immediate corollary we see that for every  $\varepsilon > 0$ , we have  $d = d(\varepsilon) \in \mathbb{N}$  such that for every  $n \in \mathbb{N}$ ,  $p = p_d$  satisfies

$$(3.8) \quad (1 - \varepsilon)^n \leq \frac{\int_0^1 \prod_{k=1}^n p(\theta \cos(2\pi a_k \omega)) d\omega}{\int_0^1 \prod_{k=1}^n e^{\theta \cos(2\pi a_k \omega)} d\omega} \leq (1 + \varepsilon)^n.$$

Let  $k_0 = k_0(d)$  be a positive integer such that  $a_{k+1}/a_k > 2d$  for all  $k \geq k_0$ ; such an index must exist since we assumed that  $a_{k+1}/a_k \rightarrow \infty$  as  $k \rightarrow \infty$ . For  $n > k_0$ , we split

$$\prod_{k=1}^n p(\theta \cos(2\pi a_k \omega)) = \left( \prod_{k=1}^{k_0} p(\theta \cos(2\pi a_k \omega)) \right) \left( \prod_{k=k_0+1}^n p(\theta \cos(2\pi a_k \omega)) \right),$$

and, taking into account that  $p(\theta \cos(2\pi a_k \omega)) > 0$  by (3.6), note that

$$(3.9) \quad m_0 \left( \prod_{k=k_0+1}^n p(\theta \cos(2\pi a_k \omega)) \right) \leq \prod_{k=1}^n p(\theta \cos(2\pi a_k \omega)) \leq M_0 \left( \prod_{k=k_0+1}^n p(\theta \cos(2\pi a_k \omega)) \right),$$

where

$$m_0 := \inf_{\omega \in [0,1]} \left( \prod_{k=1}^{k_0} p(\theta \cos(2\pi a_k \omega)) \right) \quad \text{and} \quad M_0 := \sup_{\omega \in [0,1]} \left( \prod_{k=1}^{k_0} p(\theta \cos(2\pi a_k \omega)) \right).$$

We now state an important estimate on the integral of the common product that is on both sides of the inequality (3.9).

**Lemma 3.3.** *Fix  $d \in \mathbb{N}$  and  $p = p_d$  as in (3.5). Then for any  $\theta, x \in \mathbb{R}$ ,*

$$(3.10) \quad p_d(\theta \cos x) = \sum_{j=0}^d b_j(\theta) \cos(jx),$$

where the coefficients  $b_j(\theta) = b_j(\theta; d), j = 0, 1, \dots, d$ , are real numbers with

$$(3.11) \quad b_0(\theta) = b_0(\theta; d) = \sum_{0 \leq m \leq \lfloor d/2 \rfloor} \frac{\theta^{2m}}{2^{2m} (m!)^2},$$

and for  $j = 1, \dots, d, b_j(\theta) \geq 0$  when  $\theta > 0$ . Furthermore, given  $k_0 = k_0(d) \in \mathbb{N}$  as above, for every  $\theta \in \mathbb{R}$ , and all  $n > k_0$ ,

$$(3.12) \quad \int_0^1 \prod_{k=k_0+1}^n p(\theta \cos(2\pi a_k \omega)) d\omega = b_0(\theta)^{n-k_0}.$$

We first show how Theorem A follows from Lemma 3.3, and then provide the proof of the lemma. Integrating each term in the inequalities in (3.9) with respect to Lebesgue measure over the interval  $[0, 1]$ , and applying (3.12), we obtain

$$m_0 b_0(\theta)^{n-k_0} \leq \int_0^1 \prod_{k=1}^n p(\theta \cos(2\pi a_k \omega)) d\omega \leq M_0 b_0(\theta)^{n-k_0}.$$

Combining these inequalities with (3.8) and (3.4), we arrive at

$$\frac{1}{(1 + \varepsilon)^n} m_0 (b_0(\theta))^{n-k_0} \leq \mathbb{E} [e^{\theta S_n}] \leq \frac{1}{(1 - \varepsilon)^n} M_0 (b_0(\theta))^{n-k_0}.$$

Taking the natural logarithm of each term, dividing by  $n$  and letting  $n \rightarrow \infty$ , while keeping all other variables fixed, we obtain

$$\begin{aligned} \log b_0(\theta) - \log(1 + \varepsilon) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{\theta S_n}] \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{\theta S_n}] \leq \log b_0(\theta) - \log(1 - \varepsilon). \end{aligned}$$

Recall that  $\varepsilon \in (0, 1)$  was fixed but arbitrary, that  $d = d(\varepsilon)$  depends on  $\varepsilon$  and satisfies  $d(\varepsilon) \rightarrow \infty$  as  $\varepsilon \rightarrow 0$ , and that  $b_0(\theta) = b_0(\theta; d(\varepsilon))$  depends on this choice of  $d$ . Note that  $b_0(\theta)$  is a finite partial sum of the series expansion for the moment generating function of the arcsine distribution on the interval  $(-1, 1)$ , which we

derived in (2.7). Since the logarithm is a continuous function, (3.11) and (2.8) yield

$$\lim_{\varepsilon \rightarrow 0} \log b_0(\theta; d(\varepsilon)) = \tilde{\Lambda}(\theta).$$

Thus, the last two displays together yield the limit

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{\theta S_n}] = \tilde{\Lambda}(\theta),$$

as desired. Since the function  $\tilde{\Lambda}$  is differentiable, the Gärtner-Ellis theorem can be applied, and the proof of Theorem A is complete, given Lemma 3.3.

To complete the proof of Theorem A, it only remains to establish Lemma 3.3.

*Proof of Lemma 3.3.* For every fixed  $k \in \mathbb{N}$ , the function  $\omega \mapsto p(\theta \cos(2\pi a_k \omega))$  is a polynomial of degree  $d$  in  $\theta \cos(2\pi a_k \omega)$ . A standard trigonometric formula asserts that for even  $m = 2n$ ,  $(\cos x)^m$  can be expressed as a linear combination of  $1, \cos(2x), \cos(4x), \dots, \cos(mx)$ , more precisely, for every  $n \in \mathbb{N}$  and  $x \in \mathbb{R}$ ,

$$(3.13) \quad (\cos x)^{2n} = \frac{1}{2^{2n}} \binom{2n}{n} + \frac{1}{2^{2n-1}} \sum_{\ell=0}^{n-1} \binom{2n}{\ell} \cos(2(n-\ell)x).$$

For odd  $m$ ,  $(\cos x)^m$  can be expressed as a linear combination of  $\cos x, \cos(3x), \dots, \cos(mx)$ , more precisely, for every  $n \in \mathbb{N}$  and  $x \in \mathbb{R}$ ,

$$(3.14) \quad (\cos x)^{2n+1} = \frac{1}{4^n} \sum_{\ell=0}^n \binom{2n+1}{\ell} \cos((2n+1-2\ell)x).$$

The precise statements of the last two formulas will not be important to us; we will only use the fact that the coefficient of the constant term in the expansion of  $(\cos x)^m$  is  $\binom{m}{m/2} 2^{-m}$  when  $m$  is even and zero otherwise. So for  $d \in \mathbb{N}$  and  $\theta \in \mathbb{R}$ , it is possible to write  $p(\theta \cos x)$  for  $x \in \mathbb{R}$ , in the form

$$p(\theta \cos x) = b_0(\theta) + b_1(\theta) \cos x + b_2(\theta) \cos(2x) + \dots + b_d(\theta) \cos(dx),$$

where the coefficients  $b_k(\theta) = b_k(\theta; d), k = 1, \dots, d$ , depend on the coefficients of the polynomial  $p$  (and thus on  $d$ ) and on  $\theta$ , but not on  $x$ , and the zeroth coefficient  $b_0(\theta)$  takes the explicit form

$$(3.15) \quad \begin{aligned} b_0(\theta) &= \sum_{\substack{0 \leq m \leq d, \\ m \text{ even}}} \binom{m}{m/2} \frac{\theta^m}{2^m m!} \\ &= \sum_{0 \leq m \leq \lfloor d/2 \rfloor} \binom{2m}{m} \frac{\theta^{2m}}{2^{2m} (2m)!} \\ &= \sum_{0 \leq m \leq \lfloor d/2 \rfloor} \frac{(2m)!}{m! m!} \frac{\theta^{2m}}{2^{2m} (2m)!} \\ &= \sum_{0 \leq m \leq \lfloor d/2 \rfloor} \frac{\theta^{2m}}{2^{2m} (m!)^2}, \end{aligned}$$

which agrees with (3.11). This proves (3.10). Further, when  $\theta > 0$ , since the Taylor coefficients of the exponential function are all positive, and the coefficients in the trigonometric identities (3.17) and (3.13) are all non-negative, it follows that  $b_j(\theta) \geq 0$  for  $j = 1, \dots, d$ .

It only remains to show that when integrating the product-form integrand on the right-hand side of (3.12) only terms involving the zeroth coefficient remain. To this end, note that by (3.10) we have for  $\omega \in [0, 1]$ ,

$$\begin{aligned} & \prod_{k=k_0+1}^n p(\theta \cos(2\pi a_k \omega)) \\ &= \prod_{k=k_0+1}^n \left( b_0(\theta) + b_1(\theta) \cos(2\pi a_k \omega) + b_2(\theta) \cos(2\pi 2a_k \omega) + \dots + b_d(\theta) \cos(2\pi d a_k \omega) \right). \end{aligned}$$

When multiplying out this product, we obtain a constant term  $b_0(\theta)^{n-k_0}$  as well as a sum of many mixed terms of the form

$$b_0(\theta)^{n-k_0-\ell} \cdot b_{j_1}(\theta) \cos(2\pi j_1 a_{k_1} \omega) \cdot \dots \cdot b_{j_\ell}(\theta) \cos(2\pi j_\ell a_{k_\ell} \omega),$$

for some  $\ell \in \{1, \dots, n - k_0\}$ ,  $(j_1, \dots, j_\ell) \in \{1, \dots, d\}^\ell$ , and  $(k_1, \dots, k_\ell) \in \{k_0 + 1, \dots, n\}^\ell$  such that  $k_1 > \dots > k_\ell$ . Thus, to prove (3.12), it suffices to show that for any such configuration, we have

$$(3.16) \quad \int_0^1 \cos(2\pi j_1 a_{k_1} \omega) \cdot \dots \cdot \cos(2\pi j_\ell a_{k_\ell} \omega) d\omega = 0.$$

We now show that this follows because  $a_{k+1}/a_k > 2d$  for  $k \geq k_0$  (by the choice of  $k_0$ ). Indeed, recall the standard trigonometric identity

$$(3.17) \quad \cos x \cos y = \frac{1}{2} \left( \cos(x - y) + \cos(x + y) \right),$$

which implies that the product  $\cos(2\pi j_1 a_{k_1} \omega) \cdot \dots \cdot \cos(2\pi j_\ell a_{k_\ell} \omega)$  can be written as a linear combination of cosine functions  $\cos(2\pi m \omega)$  with frequencies of the form

$$m = j_1 a_{k_1} \pm \dots \pm j_\ell a_{k_\ell}.$$

As already mentioned above, we have  $k_1 > k_2 > \dots > k_\ell > k_0$ . Then, taking into account that  $j_1 \geq 1$ , we have

$$j_1 a_{k_1} \pm \dots \pm j_\ell a_{k_\ell} \geq a_{k_1} - da_{k_2} - da_{k_3} - \dots - da_{k_\ell}.$$

The inequality  $a_{k+1}/a_k > 2d$  for all  $k \geq k_0$  then implies

$$\begin{aligned} a_{k_1} - da_{k_2} - da_{k_3} - \dots - da_{k_\ell} &\geq a_{k_1} \left( 1 - d \sum_{r=1}^{\ell} \frac{1}{(d+1)^r} \right) \\ &= a_{k_1} \underbrace{\left( 1 - d \left( \frac{1 - (d+1)^{-\ell}}{d} \right) \right)}_{>0} > 0. \end{aligned}$$

Consequently, the product  $\cos(2\pi j_1 a_{k_1} \omega) \cdot \dots \cdot \cos(2\pi j_\ell a_{k_\ell} \omega)$  can be written as a linear combination of cosine functions  $\cos(2\pi m \omega)$  that have *all non-zero* frequencies  $m \in \mathbb{N}$ . This clearly implies (3.16), and thus completes the proof.  $\square$

**3.3. Proofs of Theorem B and Lemma 2.3.** Let  $q \in \{2, 3, \dots\}$  be fixed, let  $a_k = q^k$  for each  $k \in \mathbb{N}$ , and let  $S_n$  be as defined in (2.2). We establish the LDP by first recasting  $S_n/n$  as a Birkhoff average (or time average) of a stationary sequence induced by the expanding piecewise continuous map  $\mathcal{T} : [0, 1] \rightarrow [0, 1]$  given by

$$(3.18) \quad \mathcal{T}\omega := q\omega \pmod{1} = q\omega - \lfloor q\omega \rfloor, \quad \omega \in [0, 1]$$

(which is merely the fractional part of  $q\omega$ ). Then, (2.2) and the identity  $a_k = q^k$  show that the lacunary sums of interest can be expressed as

$$(3.19) \quad S_n(\omega) = \sum_{k=0}^{n-1} X_1(\mathcal{T}^k\omega) = X_1(\omega) + \dots + X_n(\omega), \quad \omega \in [0, 1].$$

We can then apply tools from the theory of LDPs for (uniform and non-uniform) hyperbolic dynamics and mixing processes; see for example [8, 9, 13, 18, 19, 30, 34, 36, 40, 43, 58]. Since in some references (see, e.g., [11, p. 422] or [8, Thm. 10.8 on p. 90]), an LDP is stated only for some small neighborhood of 0, and since parts of the argument will be needed to prove property (iii) in the statement of Theorem B, we provide a sketch of the full proof in Section 3.3.1. The proofs of properties (i)–(iv), which are the main message of Theorem B, are presented in Section 3.3.3. They rely on additional estimates that are first obtained in Section 3.3.2. Finally, the proof of Lemma 2.3 is given in Section 3.3.4.

**3.3.1. Proof of the LDP in Theorem B.** By the Gärtner-Ellis Theorem, to prove the LDP it suffices to show that the limit  $\Lambda_q(\theta) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{\theta S_n}]$  exists for all  $\theta \in \mathbb{R}$  and is differentiable in  $\theta$ . We now express  $e^{\theta S_n}$  in terms of a certain linear operator. Let  $\text{Lip}[0, 1]$  denote the Banach space of Lipschitz functions  $f : [0, 1] \rightarrow \mathbb{C}$ , endowed with the norm  $\|f\| := \|f\|_\infty + L(f)$ , where  $L(f)$  is the Lipschitz constant of  $f$ . Next, for  $\theta \in \mathbb{R}$ , consider the linear operator  $\Phi_{\theta,q} : \text{Lip}[0, 1] \rightarrow \text{Lip}[0, 1]$  defined, for  $g \in \text{Lip}[0, 1]$ , by

$$(3.20) \quad (\Phi_{\theta,q}g)(\omega) := \frac{1}{q} \sum_{j=0}^{q-1} e^{\theta X_1(\frac{\omega+j}{q})} g\left(\frac{\omega+j}{q}\right), \quad \omega \in [0, 1],$$

where we recall from (2.1) that  $X_1(\omega) = \cos(2\pi q\omega)$ ,  $\omega \in [0, 1]$ .

The proof of the LDP for  $(S_n/n)_{n \in \mathbb{N}}$  stated in Theorem B is a direct consequence of Proposition 3.4.

**Proposition 3.4.** *Fix  $q \in \{2, 3, \dots\}$  and  $\theta \in \mathbb{R}$ . Then*

$$\Lambda_q(\theta) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{\theta S_n}] = \log \lambda_{\theta,q},$$

where  $\lambda_{\theta,q}$  is the Perron-Frobenius eigenvalue of the operator  $\Phi_{\theta,q}$  defined in (3.20). Moreover, there exists an open domain  $\mathcal{D}$  of the complex plane that contains the real line  $\mathbb{R}$  such that the convergence above holds uniformly for  $\theta$  in any compact subset of  $\mathcal{D}$ . In particular,  $\theta \mapsto \Lambda_q(\theta)$  is differentiable.

In the language of thermodynamic formalism [59],  $\log \lambda_\theta$  is referred to as the pressure or the free energy of a one-dimensional lattice system, and its differentiability expresses the known fact that there are no phase transitions for such systems. (For more background on the spectral gap property of Perron-Frobenius transfer operators, the reader is referred to [6, 8, 51] and [59, Chapter 4].)

*Proof of Proposition 3.4.* Recall the definition of the map  $\mathcal{T} : [0, 1] \rightarrow [0, 1]$  given in (3.18) and note that Lebesgue measure is an invariant measure for  $\mathcal{T}$ , i.e.,  $\mathcal{T}$  maps the measure space  $([0, 1], \mathcal{B}([0, 1]), \lambda)$  to itself and satisfies  $\lambda(\mathcal{T}^{-1}(A)) = \lambda(A)$  for every  $A \in \mathcal{B}[0, 1]$ . Indeed (for simplicity we only consider  $q = 2$ ),

$$\mathcal{T}(x) = 2x \pmod{1} = 2x - [2x] = \begin{cases} 2x & : 0 \leq x < \frac{1}{2}, \\ 2x - 1 & : \frac{1}{2} \leq x \leq 1 \end{cases}$$

and so, for every positive and measurable function  $f : [0, 1] \rightarrow \mathbb{R}$ ,

$$\begin{aligned} \int_{[0,1]} f(\mathcal{T}(x)) \lambda(dx) &= \int_{[0,1/2]} f(2x) \lambda(dx) + \int_{[1/2,1]} f(2x - 1) \lambda(dx) \\ &= \frac{1}{2} \int_{[0,1]} f(y) \lambda(dy) + \frac{1}{2} \int_{[0,1]} f(y) \lambda(dy) = \int_{[0,1]} f(y) \lambda(dy). \end{aligned}$$

In ergodic theory parlance,  $(([0, 1], \mathcal{B}([0, 1]), \lambda); \mathcal{T})$  is a measure-preserving dynamical system and we refer the reader to [21] for further details. The Perron-Frobenius operator  $\Phi_q : \text{Lip}[0, 1] \rightarrow \text{Lip}[0, 1]$  associated with  $\mathcal{T}$  is defined by

$$(\Phi_q g)(\omega) = \frac{1}{q} \sum_{j=0}^{q-1} g\left(\frac{\omega + j}{q}\right), \quad \omega \in [0, 1], g \in \text{Lip}[0, 1],$$

where recall  $\text{Lip}[0, 1]$  is the space of Lipschitz functions defined above. Note that for any  $g \in \text{Lip}[0, 1]$

$$\int_0^1 (\Phi_q g)(\omega) d\lambda(\omega) = \frac{1}{q} \sum_{j=0}^{q-1} \int_0^1 g\left(\frac{\omega + j}{q}\right) d\lambda(\omega) = \int_0^1 g(\omega) d\lambda(\omega),$$

where the last equality uses the fact that  $\mathcal{T}$  is  $\lambda$ -preserving and  $\{(\omega + j)/q, j = 0, 1, \dots, q - 1\}$  is the preimage of  $\omega$  under  $\mathcal{T}$ . This shows that  $\Phi_q$  preserves the integral for any function  $g \in \text{Lip}[0, 1]$ . Next, for  $\theta \in \mathbb{R}$ , note that the operator  $\Phi_{\theta,q}$  defined in (3.20) can be viewed as a perturbation of the operator  $\Phi_q$  since for  $g \in \text{Lip}[0, 1]$ ,

$$(3.21) \quad (\Phi_{\theta,q} g)(\omega) = (\Phi_q [e^{\theta X_1} g]) (\omega) = \frac{1}{q} \sum_{j=0}^{q-1} e^{\theta X_1(\frac{\omega+j}{q})} g\left(\frac{\omega + j}{q}\right), \quad \omega \in [0, 1],$$

where once again recall from (2.1) that  $X_1(\omega) = \cos(2\pi q\omega)$ ,  $\omega \in [0, 1]$ . It is immediate from the definition that both  $\Phi_q$  and  $\Phi_{\theta,q}$  are linear operators. Denoting by  $\Phi_q^n$  and  $\Phi_{\theta,q}^n$  the  $n$ -fold composition of  $\Phi_q$  and  $\Phi_{\theta,q}$ , respectively, a straightforward inductive argument (see, e.g., [8, Proposition 5.1 (P4)]) shows that

$$(3.22) \quad \Phi_{\theta,q}^n g = \Phi_q^n [e^{\theta S_n} g], \quad \text{for every } n \in \mathbb{N}.$$

Let  $\mathbf{1}$  denote the constant function on  $[0, 1]$  that takes the value 1, and henceforth, denote  $d\lambda(\omega)$  just as  $d\omega$ . Then, the moment generating function of  $S_n$  can be expressed as

$$(3.23) \quad \mathbb{E}[e^{\theta S_n}] = \int_0^1 e^{\theta S_n(\omega)} d\omega = \int_0^1 \Phi_q^n [e^{\theta S_n}](\omega) d\omega = \int_0^1 (\Phi_{\theta,q}^n \mathbf{1})(\omega) d\omega,$$

where the second equality uses the fact that  $\Phi_q^n$  preserves the integral and the last equality uses (3.22) with  $g = \mathbf{1}$ .

We will now use the crucial fact that the operator  $\Phi_{\theta,q}$  has the spectral gap property; see, e.g., [59, Theorems 4.1 and 4.23] and [6, Theorem 1.5], where all essential arguments can be found. Namely, we use the well-known fact that for every  $\theta \in \mathbb{R}$ ,  $\Phi_{\theta,q}$  admits a decomposition

$$(3.24) \quad \Phi_{\theta,q} = \lambda_\theta Q_\theta + R_\theta,$$

where

$$(3.25) \quad \lambda_\theta = \lambda_{\theta,q} > 0$$

is a simple eigenvalue of  $\Phi_{\theta,q}$ ,  $Q_\theta = Q_{\theta,q}$  is a projection operator onto a line spanned by an eigenfunction  $h_\theta = h_{\theta,q} > 0$  associated with  $\lambda_\theta$ , and  $R_\theta = R_{\theta,q}$  is an operator whose spectral radius is strictly smaller than  $\lambda_\theta$ . More precisely, there is a probability measure  $\mu_\theta = \mu_{\theta,q}$  on  $[0, 1]$  such that for every  $f \in \text{Lip}[0, 1]$ ,

$$Q_\theta f = h_\theta \cdot \frac{\int_0^1 f(\omega) d\mu_\theta(\omega)}{\int_0^1 h_\theta(\omega) d\mu_\theta(\omega)} \quad \text{and} \quad R_\theta Q_\theta = Q_\theta R_\theta \equiv 0.$$

Continuing to omit the dependence of the quantities  $\lambda_\theta, R_\theta, Q_\theta, h_\theta$  and  $\mu_\theta$  on  $q$ , by raising the decomposition of  $\Phi_{\theta,q}$  to the  $n$ -th power, it follows that for any  $f \in \text{Lip}[0, 1]$ ,

$$\Phi_{\theta,q}^n f = \lambda_\theta^n Q_\theta^n f + R_\theta^n f = \lambda_\theta^n \cdot h_\theta \cdot \frac{\int_0^1 f(\omega) d\mu_\theta(\omega)}{\int_0^1 h_\theta(\omega) d\mu_\theta(\omega)} + R_\theta^n f.$$

Now, setting  $f = \mathbf{1}$ , taking the integral on both sides, and using (3.23), one obtains

$$(3.26) \quad \mathbb{E}[e^{\theta S_n}] = \int_0^1 (\Phi_{\theta,q}^n \mathbf{1})(\omega) d\omega = \lambda_\theta^n \cdot \frac{\int_0^1 h_\theta(\omega) d\omega}{\int_0^1 h_\theta(\omega) d\mu_\theta(\omega)} + \int_0^1 (R_\theta^n \mathbf{1})(\omega) d\omega.$$

Recalling that the spectral radius of  $R_\theta$  is strictly smaller than  $\lambda_\theta$ , one obtains

$$(3.27) \quad \lim_{n \rightarrow \infty} \frac{\mathbb{E}[e^{\theta S_n}]}{\lambda_\theta^n} = \frac{\int_0^1 h_\theta(\omega) d\omega}{\int_0^1 h_\theta(\omega) d\mu_\theta(\omega)}.$$

Note in passing that this shows that the sequence  $(S_n)_{n \in \mathbb{N}}$  satisfies some version of mod-phi convergence [24], but what is more pertinent, it implies the weaker statement

$$(3.28) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{\theta S_n}] = \log \lambda_\theta,$$

which proves the first assertion of the proposition.

We now turn to the proof of the remaining assertions, which we claim (and justify below) can be deduced from the perturbation theory of linear operators [33, Chapter 7, §3, p. 368], in particular the Kato-Rellich theorem, as stated in [59, Theorem 4.24]. Indeed, since the family of operators  $\Phi_{\theta,q}$  depends on  $\theta \in \mathbb{C}$  in an analytic way (see [8, Proposition 5.1 (P3)] and [33, Theorem 1.7, p.368]), the decomposition (3.24) continues to hold in some neighborhood  $\mathcal{D}$  of the real axis (with  $\lambda_\theta, h_\theta$  and  $\mu_\theta$  becoming complex-valued), with  $\lambda_\theta \neq 0$  and  $\lambda_\theta$  (as well as  $h_\theta, \mu_\theta, R_\theta$ ) being analytic on  $\mathcal{D}$ . Moreover,  $|\lambda_\theta|$  stays strictly smaller than the spectral radius of  $R_\theta$  if  $\mathcal{D}$  is sufficiently small, which, looking at (3.26), shows that convergence in (3.28) is uniform on compact subsets of  $\mathcal{D}$ .  $\square$

3.3.2. *Moment estimates for the partial sums  $S_n$  and  $\tilde{S}_n$ .* Let  $n \in \mathbb{N}$  and consider

$$(3.29) \quad \Lambda_{q,n}(\theta) := \frac{1}{n} \log \mathbb{E}[e^{\theta S_n}] = \frac{1}{n} \log \sum_{m=0}^{\infty} \frac{\theta^m}{m!} \mathbb{E}[S_n^m],$$

and

$$(3.30) \quad \tilde{\Lambda}_n(\theta) := \frac{1}{n} \log \mathbb{E}[e^{\theta \tilde{S}_n}] = \frac{1}{n} \log \sum_{m=0}^{\infty} \frac{\theta^m}{m!} \mathbb{E}[\tilde{S}_n^m],$$

where we recall that  $\tilde{S}_n = \sum_{j=1}^n \tilde{X}_j$ , and  $(\tilde{X}_j)_{j \in \mathbb{N}}$  are i.i.d. having the same distribution as  $X_1$ , as defined in (2.4) and (2.3), respectively. The proof of properties

(i)–(iv) in Theorem B, presented in the next section, involves a comparison of the coefficients in the Taylor expansions of (3.29) and (3.30) considered as functions of  $\theta$ , which in turn relies on estimates on the moments of  $S_n$  and  $\tilde{S}_n$ , obtained in Lemmas 3.5–3.7. We start with Lemma 3.5 on estimates of the moments of  $S_n$ .

**Lemma 3.5.** *Fix  $q \in \{2, 3, \dots\}$ , let  $a_k = q^k$  for all  $k \in \mathbb{N}$ , and let  $S_n$  be as defined in (2.2). Then, for every  $m, n \in \mathbb{N}$ , we have*

$$\mathbb{E}[S_n^m] = \frac{A_m(n)}{2^m},$$

where  $A_m(n)$  is the number of solutions to the equation  $\sum_{i=1}^m \varepsilon_i q^{k_i} = 0$  in the unknowns  $k_1, \dots, k_m \in \{1, \dots, n\}$  and  $\varepsilon_1, \dots, \varepsilon_m \in \{+1, -1\}$ .

*Proof.* For every  $m \in \mathbb{N}$ , we have

$$\begin{aligned} \mathbb{E}[S_n^m] &= \int_0^1 \left( \sum_{k=1}^n \cos(2\pi q^k \omega) \right)^m d\omega \\ &= \int_0^1 \sum_{k_1, \dots, k_m=1}^n \prod_{\ell=1}^m \cos(2\pi q^{k_\ell} \omega) d\omega \\ &= \frac{1}{2^m} \int_0^1 \sum_{k_1, \dots, k_m=1}^n \prod_{\ell=1}^m [e(q^{k_\ell} \omega) + e(-q^{k_\ell} \omega)] d\omega, \end{aligned}$$

where we write  $e(z) := e^{2\pi iz}$  for  $z \in \mathbb{R}$  and used that  $\cos z = (e^{iz} + e^{-iz})/2$ . By rewriting the product in the last line of the last display in terms of an exponential and using the elementary identity  $\int_0^1 e(k\omega) d\omega = 0$  for all integer  $k \neq 0$ , we see that

$$(3.31) \quad \mathbb{E}[S_n^m] = \frac{1}{2^m} \sum_{k_1, \dots, k_m=1}^n \sum_{\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\}} \mathbb{1}_{\{\varepsilon_1 q^{k_1} + \dots + \varepsilon_m q^{k_m} = 0\}}.$$

To complete the proof of the lemma, observe that the right-hand side equals  $A_m(n)/2^m$ . □

Next, we give a combinatorial interpretation of  $A_m(n)$  for  $m \leq q$ . Let  $B_m(n)$  be the number of simple random walk paths in  $\mathbb{Z}^n$  of length  $m$  that return to the origin, which is sometimes also referred to as the number of bridges of length  $m$  in  $\mathbb{Z}^n$ .

**Lemma 3.6.** *For all  $n, m \in \mathbb{N}$ , we have  $A_m(n) \geq B_m(n)$  and, if  $m \leq q$ , then  $A_m(n) = B_m(n)$ .*

*Proof.* We start with the proof of the second statement. Let  $m \leq q$ . We first claim (and justify below) that

$$(3.32) \quad \sum_{\ell=1}^m \varepsilon_\ell q^{k_\ell} = 0, \quad k_\ell \in \{1, \dots, n\}, \varepsilon_\ell \in \{-1, 1\}$$

if and only if for every  $k \in \{1, \dots, n\}$ ,

$$(3.33) \quad H_k = \sum_{\ell=1}^m \mathbb{1}_{\{k_\ell=k\}} \cdot \varepsilon_\ell = 0.$$



In other words, (3.32) can hold only if every term  $+q^k$  is canceled by a term  $-q^k$  at some other place. One direction of the claim is immediate. We note that

$$(3.34) \quad \sum_{\ell=1}^m \varepsilon_\ell q^{k_\ell} = \sum_{k=1}^n \left( \sum_{\ell=1}^m \mathbb{1}_{\{k_\ell=k\}} \cdot \varepsilon_\ell \right) q^k = \sum_{k=1}^n H_k q^k$$

and therefore if all  $H_k$  vanish, then  $\sum_{\ell=1}^m \varepsilon_\ell q^{k_\ell} = 0$ . For the opposite direction, suppose  $\sum_{\ell=1}^m \varepsilon_\ell q^{k_\ell} = 0$ . Then, due to the identity in (3.34),

$$\sum_{k=1}^n H_k q^k = 0.$$

We first show that this, along with the fact that  $m \leq q$ , implies  $H_1 = 0$ . First, dividing everything by  $q \geq 2$ , we obtain

$$(3.35) \quad \sum_{k=1}^n H_k q^{k-1} = 0,$$

which clearly implies divisibility of  $H_1$  by  $q$ . Now, if  $m < q$ , then  $|H_1| \leq m < q$  by definition. Hence,  $H_1 = 0$ . If  $m = q$ , then either  $H_1 = 0$  or  $H_1 = \pm q$  and the latter case only occurs if all  $\varepsilon_1, \dots, \varepsilon_m$  are equal and  $k_\ell = 1$  for all  $\ell \in \{1, \dots, m\}$ . In this case, the condition  $\sum_{\ell=1}^m \varepsilon_\ell q^{k_\ell} = 0$  is violated. Hence, for  $m \leq q$ , we have  $H_1 = 0$ . Now dividing (3.35) by  $q$  and repeating the argument, it follows that  $H_2 = \dots = H_n = 0$  as well. This completes the proof of the claim of equivalence between the conditions (3.33) and (3.32).

Next, note that the conditions (3.33) on  $H_k$  may be interpreted as follows: for given  $\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\}$  and  $k_1, \dots, k_m \in \{1, \dots, n\}$ , we consider the nearest neighbor path of length  $m$  in  $\mathbb{Z}^n$  whose  $\ell^{\text{th}}$  step is equal to  $\varepsilon_\ell \vec{e}_{k_\ell}$  for  $\vec{e}_1, \dots, \vec{e}_n$  the standard vector basis in  $\mathbb{R}^n$ . Clearly, condition (3.33) is satisfied if and only if the path returns to its starting point. It follows that  $A_m(n) = B_m(n)$ , which proves the second assertion of the lemma.

To prove the first assertion, note that if  $m \in \mathbb{N}$  is arbitrary, then the solutions of (3.32) can be divided into the trivial ones (i.e., those for which  $H_1 = \dots = H_n = 0$ ), and the non-trivial ones (such as  $q^2 - q - \dots - q = 0$  for  $m = q + 1$ , where the term  $-q$  appears  $m$  times). Since the number of trivial solutions is  $B_m(n)$ , and (by definition)  $A_m(n)$  is the total number of solutions, the claim  $A_m(n) \geq B_m(n)$  follows.  $\square$

Taken together, Lemmas 3.5 and 3.6 show that, for each  $m \leq q$ ,

$$\mathbb{E}[S_n^m] = \frac{B_m(n)}{2^m}.$$

Let us turn to the computation of  $\mathbb{E}[\tilde{S}_n^m]$ , where we shall prove that the analogous identity holds, this time for all  $m \in \mathbb{N}$ .

**Lemma 3.7.** *Recall that  $\tilde{S}_n = \tilde{X}_1 + \dots + \tilde{X}_n$ , where  $\tilde{X}_1, \tilde{X}_2, \dots$  are i.i.d. random variables with the arcsine distribution on  $(-1, 1)$ . Then, for all  $m, n \in \mathbb{N}$ , we have*

$$\mathbb{E}[\tilde{S}_n^m] = \frac{B_m(n)}{2^m}.$$

*Proof.* Recalling that  $(U_k)_{k \in \mathbb{N}}$  is a sequence of i.i.d. random variables with the same uniform distribution as  $U$ , we can write

$$\tilde{X}_k = \cos(2\pi U_k) = \frac{e(U_k) + e(-U_k)}{2},$$

where we again write  $e(z) := e^{2\pi iz}$ . For  $m \in \mathbb{N}$ , we have

$$\begin{aligned} \mathbb{E}[\tilde{S}_n^m] &= \frac{1}{2^m} \mathbb{E} \left[ \left( \sum_{k=1}^n [e(U_k) + e(-U_k)] \right)^m \right] \\ &= \frac{1}{2^m} \sum_{k_1, \dots, k_m=1}^n \sum_{\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\}} \mathbb{E} \left[ \prod_{\ell=1}^m e(\varepsilon_\ell U_{k_\ell}) \right] \\ &= \frac{1}{2^m} \sum_{k_1, \dots, k_m=1}^n \sum_{\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\}} \mathbb{E} \left[ e \left( \sum_{\ell=1}^m \varepsilon_\ell U_{k_\ell} \right) \right] \\ &= \frac{1}{2^m} \sum_{k_1, \dots, k_m=1}^n \sum_{\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\}} \mathbb{E} \left[ e \left( \sum_{\ell=1}^n H_\ell U_\ell \right) \right], \end{aligned}$$

where for any fixed  $k_i \in \{1, \dots, n\}$ ,  $\varepsilon_i \in \{-1, 1\}$ ,  $i = 1, \dots, m$ , we set

$$H_\ell := \sum_{i=1}^m \mathbb{1}_{\{k_i=\ell\}} \cdot \varepsilon_i, \quad \ell = 1, \dots, n.$$

Then we have

$$\mathbb{E} \left[ e \left( \sum_{\ell=1}^n H_\ell U_\ell \right) \right] = \prod_{\ell=1}^n \mathbb{E} \left[ e \left( H_\ell U_\ell \right) \right] = \mathbb{1}_{\{H_1=\dots=H_n=0\}}.$$

Since  $H_1 = \dots = H_n = 0$  if and only if the associated nearest neighbor path of length  $m$  in  $\mathbb{Z}^n$ , whose  $\ell^{th}$  step is equal to  $\varepsilon_i \vec{e}_{k_i}$ , with  $\vec{e}_1, \dots, \vec{e}_n$  the standard basis in  $\mathbb{Z}^n$ , returns to its starting point, we have shown that

$$\mathbb{E}[\tilde{S}_n^m] = \frac{1}{2^m} \sum_{k_1, \dots, k_m=1}^n \sum_{\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\}} \mathbb{1}_{\{H_1=\dots=H_n=0\}} = \frac{B_m(n)}{2^m},$$

which completes the proof. □

**3.3.3. Proof of properties (i)–(iv) of Theorem B.** We now complete the proof of Theorem B. First, note that the function  $\Lambda_q$ , as a uniform limit of analytic functions, is itself analytic for all  $\theta \in \mathbb{C}$ ,  $|\theta| < \varepsilon_0$ , for a sufficiently small  $\varepsilon_0 > 0$ .

*Proof of (i).* First, let us observe that the proof that  $I_q \leq \tilde{I}$  on the positive real axis is simple. Indeed, Lemmas 3.5–3.7 imply that for all  $m, n \in \mathbb{N}$ ,

$$\mathbb{E}[S_n^m] \geq \mathbb{E}[\tilde{S}_n^m].$$

When combined with (3.29) and (3.30), it follows that for every  $n \in \mathbb{N}$  and  $\theta > 0$ ,

$$\Lambda_{q,n}(\theta) \geq \tilde{\Lambda}_n(\theta).$$

Passing to the limit as  $n \rightarrow \infty$  on both sides, and noting that both limits exist and are equal to  $\Lambda_q(\theta)$  and  $\tilde{\Lambda}_q(\theta)$ , respectively, due to the proof in Section 3.3.1 and the independence of  $(\tilde{X}_k)_{k \in \mathbb{N}}$ , we conclude that  $\Lambda_q(\theta) \geq \tilde{\Lambda}_q(\theta)$  for all  $\theta > 0$ . Passing to the Legendre-Fenchel transform we then obtain  $I_q(x) \leq \tilde{I}(x)$  for all  $x > 0$ .

The proof of the strict inequality  $I_q(x) < \tilde{I}(x)$  for  $x > 0$  is more delicate. Assume that  $q \geq 2$  and  $\theta > 0$  are fixed. We choose a large integer  $d > q$ ; at the end of the proof we will let  $d \rightarrow \infty$ . As in the proof of Theorem A, we approximate the exponential function by a Taylor polynomial  $p = p_d$  of degree  $d$ , and by (3.7), we have

$$\mathbb{E} [e^{\theta S_n}] \geq \left(1 + \frac{e^\theta \theta^{d+1}}{(d+1)!}\right)^{-n} \int_0^1 \prod_{k=1}^n p(\theta \cos(2\pi q^k \omega)) d\omega.$$

We recall from Lemma 3.3 that we can write  $p(\theta \cos(2\pi q^k \omega))$  in the form

$$(3.36) \quad b_0(\theta) + b_1(\theta) \cos(2\pi q^k \omega) + b_2(\theta) \cos(2\pi 2q^k \omega) + \dots + b_d(\theta) \cos(2\pi dq^k \omega),$$

where  $b_0 = b_0(\theta; d)$  is given by (3.15) and  $b_j = b_j(\theta; d) \geq 0$  for  $j = 1, \dots, d$ . Since  $d > q$  by assumption, the  $q$ -th term in the Taylor expansion for  $p(\theta \cos(2\pi q^k x))$  is  $(\theta \cos(2\pi q^k x))^q / q!$ . From (3.13) and (3.14) we see that the expansion of  $(\cos y)^q$  into a linear combination of cosine functions contains the term  $2^{-q+1} \cos(qy)$ . We emphasize again that all coefficients, in the Taylor expansion of  $e^y$  as well as in (3.13) and (3.14), are non-negative. Thus the coefficient  $b_q(\theta)$  in (3.36) is at least as large as the contribution coming from  $(\theta \cos(2\pi q^k x))^q / q!$ , and so we have

$$(3.37) \quad b_q = b_q(\theta) \geq \frac{\theta^q}{q!} \frac{1}{2^{q-1}}.$$

By a similar reasoning the coefficient  $b_1(\theta)$  in (3.36) is at least as large as the contribution coming from the linear term in the Taylor expansion, which is simply  $\theta \cos(2\pi q^k x)$ . Thus we have  $b_1(\theta) \geq \theta$ . Once again using the fact that all coefficients are non-negative, in (3.36) as well as in (3.17), (3.13) and (3.14), we have

$$\int_0^1 \prod_{k=1}^n p(\theta \cos(2\pi q^k \omega)) d\omega \geq \int_0^1 \prod_{k=1}^n (b_0 + b_1 \cos(2\pi q^k \omega) + b_q \cos(2\pi q^{k+1} \omega)) d\omega.$$

Now the point is that there will always be interference between the term  $b_q \cos(2\pi q^{k+1} x)$  coming from index  $k$ , and the term  $b_1 \cos(2\pi q^{k+1} x)$  coming from index  $k + 1$ . Let us assume that  $n$  is even. Always combining two consecutive factors together, we have

$$\begin{aligned} & \int_0^1 \prod_{k=1}^n (b_0 + b_1 \cos(2\pi q^k \omega) + b_q \cos(2\pi q^{k+1} \omega)) d\omega \\ &= \int_0^1 \prod_{\ell=1}^{n/2} (b_0 + b_1 \cos(2\pi q^{2\ell-1} \omega) + b_q \cos(2\pi q^{2\ell} \omega)) (b_0 + b_1 \cos(2\pi q^{2\ell} \omega) \\ & \quad + b_q \cos(2\pi q^{2\ell+1} \omega)) d\omega \\ &\geq \int_0^1 \prod_{\ell=1}^{n/2} (b_0^2 + b_q \cos(2\pi q^{2\ell} \omega) b_1 \cos(2\pi q^{2\ell} \omega)) d\omega \\ &\geq \prod_{\ell=1}^{n/2} \left( b_0^2 + \frac{b_1 b_q}{2} \right) \\ &\geq \prod_{\ell=1}^{n/2} \left( b_0^2 + \frac{\theta q^{q+1}}{q! 2^q} \right), \end{aligned}$$

where the last inequality uses (3.37) and  $b_1(\theta) > \theta$ . Consequently, we have

$$\Lambda_q(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{\theta S_n}] \geq \log \left( \underbrace{\left( 1 + \frac{e^\theta \theta^{d+1}}{(d+1)!} \right)^{-1}}_{\rightarrow 1 \text{ as } d \rightarrow \infty} \sqrt{b_0^2 + \frac{\theta^{q+1}}{q!2^q}} \right).$$

Recall that  $b_0$  depends on  $d$  and  $\theta$ , and that we have  $\log b_0 \rightarrow \tilde{\Lambda}(\theta)$  as  $d \rightarrow \infty$ . For every fixed  $\theta > 0$  since the logarithm is a strictly increasing function, the term  $\log \left( \sqrt{b_0^2 + \frac{\theta^{q+1}}{q!2^q}} \right)$  converges to a quantity that is strictly larger than  $\tilde{\Lambda}(\theta)$  as  $d \rightarrow \infty$ . Consequently, we have

$$\Lambda_q(\theta) > \tilde{\Lambda}(\theta), \quad \text{for all } \theta > 0.$$

From the properties of the Bessel function  $B_0(\theta)$  it is easily seen that for  $x \in (0, 1)$  the supremum in the definition of  $\tilde{I}(x)$  is actually a maximum, and is attained at some (finite) value  $\theta_x > 0$ . Consequently, we have

$$I_q(x) = \sup_{\theta > 0} [\theta x - \Lambda_q(\theta)] = \theta_x x - \Lambda_q(\theta_x) < \theta_x x - \tilde{\Lambda}(\theta_x) = \tilde{I}(x).$$

Thus, we have  $I_q(x) < \tilde{I}(x)$  for all  $x \in (0, 1)$ .

In conclusion, we note that we can make the difference between  $\Lambda_q$  and  $\tilde{\Lambda}$  quantitative. Recall that  $\theta > 0$  by assumption. Since  $b_0(\theta)$  is a partial sum of  $B_0(\theta)$ , we have  $b_0(\theta) \leq B_0(\theta)$ . Furthermore, from the series expansion for  $B_0(\theta)$  it is easily seen that  $B_0(\theta) \leq e^\theta$ . Thus  $b_0^2(\theta) \leq e^{2\theta}$ , and  $b_0^2 + \frac{\theta^{q+1}}{q!2^q} \geq b_0^2 \left( 1 + \frac{\theta^{q+1}}{q!2^q e^{2\theta}} \right)$ . Thus, letting  $d \rightarrow \infty$ , we deduce that

$$\Lambda_q(\theta) - \tilde{\Lambda}(\theta) \geq \frac{1}{2} \log \left( 1 + \frac{\theta^{q+1}}{q!2^q e^{2\theta}} \right).$$

*Proof of (ii).* It follows from Proposition 3.4 that for  $\theta \in \mathbb{R}$ ,  $\Lambda_q(\theta) = \log \lambda_{\theta,q}$ , where  $\lambda_{\theta,q} > 0$  is the largest eigenvalue of the Perron-Frobenius transfer operator defined in (3.20). Fixing  $\theta \in \mathbb{R}$  and sending  $q \rightarrow \infty$ , the Riemann sums converge on the right-hand side of the definition in (3.20) converge to the corresponding Riemann integrals; hence this sequence of operators converges in the norm topology to the operator

$$(\tilde{\Phi}_\theta g)(\omega) = \int_0^1 e^{\theta X_1(z)} g(z) dz = \tilde{\lambda}_\theta \frac{\int_0^1 e^{\theta X_1(z)} g(z) dz}{\int_0^1 e^{\theta X_1(z)} dz} \cdot \mathbf{1},$$

where  $\tilde{\lambda}_\theta := \int_0^1 e^{\theta X_1(z)} dz = e^{\tilde{\Lambda}(\theta)}$ , where  $\tilde{\Lambda}$  is defined as in (2.6). Thus,  $\tilde{\Phi}_\theta / \tilde{\lambda}_\theta$  is a projection onto the line spanned by the function  $\mathbf{1}$ . The Perron-Frobenius eigenvalue of  $\tilde{\Phi}_{\theta,q}$  is  $\tilde{\lambda}_\theta$ . Now, if  $\theta \in \mathbb{R}$  stays constant and  $q \rightarrow \infty$ , we can view  $\tilde{\Phi}_{\theta,q}$  as a perturbation of  $\tilde{\Phi}_\theta$ . By perturbation theory (see, e.g., [33]), we have the convergence of the Perron-Frobenius eigenvalues, that is,  $\lim_{q \rightarrow \infty} \lambda_{\theta,q} = \tilde{\lambda}_\theta$  for every  $\theta \in \mathbb{R}$ . Taking the logarithm, we get  $\lim_{q \rightarrow \infty} \Lambda_q(\theta) = \tilde{\Lambda}(\theta)$ . Since the involved functions are convex, the convergence is, in fact, uniform on compact intervals. By taking the Legendre-Fenchel transform, it follows that  $\lim_{q \rightarrow \infty} I_q(x) = \tilde{I}(x)$  locally uniformly on  $(-1, 1)$ .

*Proof of (iii).* Lemma 3.7 shows that for every  $n \in \mathbb{N}$ , whenever  $m \in \mathbb{N}$  satisfies  $m \leq q$ , one has

$$\mathbb{E}[S_n^m] = \mathbb{E}[\tilde{S}_n^m] = \frac{B_m(n)}{2^m},$$

or, in other words, the moments of  $S_n$  and  $\tilde{S}_n$  coincide for all  $m \leq q$ . Since cumulants of order less than or equal to  $q$  can be expressed in terms of moments of order less than or equal to  $q$ , we infer that as long as  $m \leq q$ ,

$$\kappa_m(S_n) = \kappa_m(\tilde{S}_n) := n \cdot \kappa_m(\tilde{X}_1),$$

where  $\kappa_m(Y)$  denotes the  $m$ th cumulant of a real-valued random variable  $Y$ . Hence, for  $m \leq q$  and every  $n \in \mathbb{N}$ , we have

$$\left( \frac{d}{d\theta} \right)^m \Lambda_{q,n}(\theta) \Big|_{\theta=0} = \frac{1}{n} \kappa_m(S_n) = \kappa_m(\tilde{X}_1) = \frac{1}{n} \kappa_m(\tilde{S}_n) = \left( \frac{d}{d\theta} \right)^m \tilde{\Lambda}(\theta) \Big|_{\theta=0}.$$

Now because the uniform convergence of the analytic functions  $\Lambda_{q,n} \rightarrow \Lambda_q$  (established in Proposition 3.4) implies the convergence of the derivatives, we obtain (iii).

*Proof of (iv).* In the case when  $m = q + 1$ , a slight modification of the argument used to prove Lemma 3.6 shows that any solution to (3.32) either satisfies  $H_1 = \dots = H_n = 0$ , or is a permutation of one of the solutions  $q^k + \dots + q^k - q^{k+1} = 0$  or  $q^{k+1} - q^k - \dots - q^k = 0$ , where  $k \in \{1, \dots, n - 1\}$ . The total number of such exceptional solutions is  $2(q + 1)(n - 1)$ , hence

$$A_{q+1}(n) = B_{q+1}(n) + 2(q + 1)(n - 1).$$

From Lemma 3.5 and Lemma 3.7 it follows that

$$\mathbb{E}[S_n^{q+1}] = \frac{A_{q+1}(n)}{2^{q+1}} = \frac{B_{q+1}(n) + 2(q + 1)(n - 1)}{2^{q+1}} = \mathbb{E}[\tilde{S}_n^{q+1}] + \frac{(q + 1)(n - 1)}{2^q}.$$

The cumulant  $\kappa_{q+1}(S_n)$  can be expressed as  $\mathbb{E}[S_n^{q+1}]$  plus some polynomial function of the lower moments  $\mathbb{E}[S_n^m]$  with  $m \leq q$ . A similar representation holds for the cumulant  $\kappa_{q+1}(\tilde{S}_n)$ , and the moments of all orders  $m \leq q$  of  $S_n$  coincide with those of  $\tilde{S}_n$  by part (ii) of Theorem B. It follows that

$$\kappa_{q+1}(S_n) = \kappa_{q+1}(\tilde{S}_n) + \frac{(q + 1)(n - 1)}{2^q}.$$

For the derivatives of order  $q + 1$  of  $\Lambda_{q,n}$  and  $\tilde{\Lambda}$  at  $\theta = 0$  we therefore obtain

$$\begin{aligned} \Lambda_{q,n}^{(q+1)}(0) &= \frac{1}{n} \kappa_{q+1}(S_n) = \frac{1}{n} \kappa_{q+1}(\tilde{S}_n) + \frac{(q + 1)(n - 1)}{2^q n} \\ &= \tilde{\Lambda}^{(q+1)}(0) + \frac{(q + 1)(n - 1)}{2^q n}. \end{aligned}$$

Letting  $n \rightarrow \infty$  and using that the uniform convergence of analytic functions  $\Lambda_{q,n} \rightarrow \Lambda_q$  implies convergence of their derivatives, we arrive at

$$\Lambda_q^{(q+1)}(0) = \tilde{\Lambda}^{(q+1)}(0) + \frac{q + 1}{2^q}.$$

This proves (iv).

3.3.4. *Proof of Lemma 2.3.* We now present the proof of Lemma 2.3. The idea is that in the lacunary sum  $S_n$  all cosine functions  $\cos(2\pi q^k U)$  are equal to 1 at  $U = 0$ . Thus,  $S_n$  is close to  $n$  if the uniform random variable  $U \sim \text{Unif}(0, 1)$  takes a value that is sufficiently close to 0. To make this precise, fix  $\varepsilon \in (0, 1)$ . We have  $\cos x \geq 1 - x^2/2$ . It follows that

$$X_k = \cos(2\pi q^k U) \geq 1 - \varepsilon \quad \text{provided that} \quad U \leq \frac{\sqrt{2\varepsilon}}{2\pi} q^{-k}.$$

Hence, if  $U \leq \frac{\sqrt{2\varepsilon}}{2\pi} q^{-n}$ , then we have  $S_n \geq (1 - \varepsilon)n$ . It follows that

$$\begin{aligned} I_q(1 - \varepsilon) &= - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq (1 - \varepsilon)n) \leq - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left(U \leq \frac{\sqrt{2\varepsilon}}{2\pi} q^{-n}\right) \\ &\leq - \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{\sqrt{2\varepsilon}}{2\pi} q^{-n}\right) \\ &= \log q. \end{aligned}$$

Since this holds for every  $\varepsilon \in (0, 1)$ , by the lower semicontinuity of  $I_q$ , it follows that

$$I_q(+1) \leq \liminf_{\varepsilon \rightarrow 0} I_q(1 - \varepsilon) \leq \log q.$$

This completes the proof.

3.4. **Proof of Theorem C.** We know from Theorem B and Proposition 2.4 that there exists some sufficiently small  $\bar{x}_0 > 0$  such that  $0 < I_2(x_0) < I_3(x_0)$  for every  $x_0 \in \mathbb{R}$  with  $0 < x_0 \leq \bar{x}_0$ . By interleaving the powers of 2 and 3 appropriately, we shall construct a Hadamard gap sequence  $(a_k)_{k \in \mathbb{N}}$  such that for all  $x_0 \in (0, \bar{x}_0)$ , the corresponding partial sums  $(S_n)_{n \in \mathbb{N}}$  satisfy

$$0 < \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(S_n \geq nx_0) < \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(S_n \geq nx_0) < \infty.$$

Since both  $I_2$  and  $I_3$  are continuous, there exist  $\varepsilon_0 \in (0, x_0)$  and  $\delta_0 > 0$  such that

$$\sup_{|x-x_0| \leq \varepsilon_0} I_2(x) + \delta_0 < \inf_{|x-x_0| \leq \varepsilon_0} I_3(x) \quad \text{and} \quad \inf_{|x-x_0| \leq \varepsilon_0} I_2(x) > \delta_0.$$

Our construction proceeds inductively. Assume that for some  $n \in \mathbb{N}$  we have constructed increasing positive integers  $a_1, \dots, a_n$  such that

$$-\frac{1}{n} \log \mathbb{P}(S_n \geq nx_0) > \inf_{|x-x_0| \leq \varepsilon_0} I_3(x) - \frac{\delta_0}{3} =: c_+.$$

We want to extend the sequence  $a_1, \dots, a_n$  to a longer sequence  $a_1, \dots, a_N$ , with  $N \in \mathbb{N}$ ,  $N > n$  in such a way that

$$-\frac{1}{N} \log \mathbb{P}(S_N \geq Nx_0) < \sup_{|x-x_0| \leq \varepsilon_0} I_2(x) + \frac{\delta_0}{3} =: c_-.$$

Note that  $0 < c_- < c_+$ . This is done as follows. We define  $a_{n+1} := 2^m$ , where  $m \in \mathbb{N}$  is any number such that  $2^m > 2a_n$  (to guarantee the Hadamard gap condition) and  $m > n + 1$ . Further, we define  $a_{n+\ell} := 2^{m+(\ell-1)}$  so that with  $N = n + (N - n)$ , we have  $a_N := 2^{N-n+m-1}$ . We choose  $N \in \mathbb{N}$  sufficiently large, in particular such that  $2m/N < \varepsilon_0/5$ . Clearly,

$$\left| \sum_{k=1}^n \cos(2\pi a_k x_0) \right| \leq n \quad \text{and} \quad \left| \sum_{k=1}^{m-1} \cos(2\pi 2^k x_0) \right| \leq m.$$

Therefore, by replacing the first  $n$  elements  $a_1, \dots, a_n$  by  $m - 1$  powers of 2, more precisely, by  $2, 2^2, \dots, 2^{m-1}$ , respectively, and using the specific choice of  $a_{n+1}, a_{n+2}, \dots$  together with the two estimates in the previous display (which guarantee that the replacement of  $n$  by  $m - 1$  cosine terms yields an error bounded above by  $n + m$ ), we obtain

$$\begin{aligned} & -\frac{1}{N} \log \mathbb{P}(S_N \geq Nx_0) \\ & \leq -\frac{1}{N} \log \mathbb{P}\left(\sum_{k=1}^{N-n+m-1} \cos(2\pi 2^k x_0) \geq Nx_0 + (n + m)\right) \\ & \leq -\frac{1}{N} \log \mathbb{P}\left(\sum_{k=1}^N \cos(2\pi 2^k x_0) \geq Nx_0 + (n + m) + (m - n - 1)\right) \\ & \leq -\frac{1}{N} \log \mathbb{P}\left(\sum_{k=1}^N \cos(2\pi 2^k x_0) \geq Nx_0 + 2m\right) \\ & \leq -\frac{1}{N} \log \mathbb{P}\left(\sum_{k=1}^N \cos(2\pi 2^k x_0) \geq N(x_0 + \varepsilon_0/5)\right), \end{aligned}$$

where we used that  $2m/N < \varepsilon_0/5$ . The latter expression converges, as  $N \rightarrow \infty$ , to  $I_2(x_0 + \varepsilon_0/5)$ . Hence, making  $N \in \mathbb{N}$  larger, if necessary, we obtain

$$-\frac{1}{N} \log \mathbb{P}(S_N \geq Nx_0) < I_2(x_0 + \varepsilon_0/5) + \frac{\delta_0}{3} \leq \sup_{|x-x_0| \leq \varepsilon_0} I_2(x) + \frac{\delta_0}{3} = c_-.$$

Now we can continue this argument back and forth, by adding strings of consecutive powers of 2 in odd steps and strings of powers of 3 in even steps, we can construct an infinite sequence  $(a_k)_{k \in \mathbb{N}}$  for which  $-\frac{1}{n} \log \mathbb{P}(S_n \geq nx_0)$  is infinitely often smaller than  $c_-$  and infinitely often larger than  $c_+$ .

**3.5. Proof of Theorem D.** Recall that the i.i.d. sequence  $Y = (Y_k)_{k \in \mathbb{N}}$  is defined on a common probability space  $(\mathcal{Y}, \mathcal{A}, \mathbb{P}_Y)$  with each  $Y_k$  uniformly distributed on the discrete set

$$(3.38) \quad \mathcal{D}_k := \left\{ h 2^{\lceil k^{2/3} \rceil} : h \in \mathbb{Z}, 0 \leq h \leq 2^{\lceil k^{2/3} \rceil} \right\}, \quad k \in \mathbb{N}.$$

Since by definition  $a_k^Y = 2^k + Y_k$ ,  $(a_k^Y)_{k \in \mathbb{N}}$  is also a sequence of independent random variables defined on  $(\mathcal{Y}, \mathcal{A}, \mathbb{P}_Y)$ . We also assume (without loss of generality) that the independent uniform random variable  $U$  is realized as the identity map on the space  $([0, 1], \mathcal{B}(0, 1), \lambda)$  and, since  $U$  and  $Y$  are independent, that both  $Y$  and  $U$  are defined on the product measure space  $(\mathcal{Y} \times [0, 1], \mathcal{A} \otimes \mathcal{B}(0, 1), \mathbb{P}_Y \otimes \lambda)$ . Throughout the argument, fix  $\theta \in \mathbb{R}$ . The proof proceeds in several steps.

*Step 1* (Construct a suitable partition of the integers). For any large  $n$ , we split the set of all positive integers into disjoint sets  $\Delta_1, \Delta_2, \dots$  and  $\Delta'_1, \Delta'_2, \dots$ , which are defined via the following recursive construction. First, set  $\Delta_1 := \{1, \dots, n^{1/2}\}$ , where for notational simplicity, we assume that  $n^{1/2}$  is an integer. Let the set  $\Delta'_1$  contain the next  $n^{2/5}$  smallest positive integers not already contained in  $\Delta_1$ , where (again for notational simplicity) we assume that  $n^{2/5}$  is also an integer. Then, for each  $i \in \mathbb{N}$ , we recursively define  $\Delta_{i+1}$  to contain the  $n^{1/2}$  smallest positive integers not already contained in  $\bigcup_{j=1}^i (\Delta_j \cup \Delta'_j)$ , and the set  $\Delta'_{i+1}$  to contain the  $n^{2/5}$

smallest positive integers not already contained in  $\Delta_{i+1} \cup \bigcup_{j=1}^i (\Delta_j \cup \Delta'_j)$ . This decomposition can be characterized by the following requirements:

- $\Delta_i < \Delta'_i < \Delta_{i+1}$  for all  $i \in \mathbb{N}$ , where the inequality is understood to hold element-wise
- $\bigcup_{i=1}^\infty (\Delta_i \cup \Delta'_i) = \mathbb{N}$ .
- $\#\Delta_i = n^{1/2}$  for all  $i \in \mathbb{N}$ , and  $\#\Delta'_i = n^{2/5}$  for all  $i \in \mathbb{N}$ .

The philosophy is that the primed index sets are sufficiently large to cause a strong “independence” between the trigonometric functions in the non-primed sets, while at the same time the total cardinality of the primed index sets is so small that they are asymptotically negligible. The precise choice of  $n^{1/2}$  and  $n^{2/5}$  for the cardinalities of the  $\Delta$  and  $\Delta'$  blocks is somewhat arbitrary, the relevant facts are that the one type of block is significantly larger than the other, and that both types of blocks are not too small in comparison with  $n$ .

For  $i \in \mathbb{N}$ , let  $\delta_{i,\min}$  and  $\delta_{i,\max}$  denote the smallest and largest integers in  $\Delta_i$ , respectively. Then our construction ensures that

$$(3.39) \quad \delta_{i,\max} + n^{2/5} < \delta_{i+1,\min}, \quad \forall i \in \mathbb{N}, \quad \text{and} \quad \max_{1 \leq i \leq M_n} \delta_{i,\max} + n^{2/5} \leq n,$$

where

$$(3.40) \quad M_n := \min \left\{ M \in \mathbb{N} : \{1, \dots, n\} \subset \bigcup_{i=1}^{M+1} (\Delta_i \cup \Delta'_i) \right\} \leq \sqrt{n},$$

with the last inequality being a simple consequence of the fact that  $|\Delta_i| = \sqrt{n}$  for each  $i \in \mathbb{N}$ .

*Step 2* (Bound the moment generating function in terms of polynomial integrals). Recall that  $\mathcal{D}_\infty = \otimes_{k \in \mathbb{N}} \mathcal{D}_k$ , where the definition of the discrete set  $\mathcal{D}_k$  was repeated again in (3.38), and for  $y \in \mathcal{D}_\infty$ ,  $a_k^y = 2^k + y_k$ . Recall also that  $S_n^y(\omega) = \sum_{k=1}^n \cos(2\pi a_k^y \omega)$ ,  $\omega \in [0, 1]$ .

**Lemma 3.8.** *Fix  $n \in \mathbb{N}$  sufficiently large such that  $n^{2/5} \leq M_n$ . Then, for  $y \in \mathcal{D}_\infty$  and  $\omega \in [0, 1]$ ,*

$$(3.41) \quad e^{-5\theta n^{9/10}} \int_0^1 H_n^y(\omega) d\omega \leq \int_0^1 e^{\theta S_n^y(\omega)} d\omega \leq e^{5\theta n^{9/10}} \int_0^1 H_n^y(\omega) d\omega,$$

where for  $\omega \in [0, 1]$

$$(3.42) \quad H_n^y(\omega) := \left( \prod_{n^{2/5} \leq i \leq M_n} \prod_{k \in \Delta_i} e^{\theta \cos(2\pi a_k^y \omega)} \right).$$

Consequently, for any  $\varepsilon > 0$ , there exists  $d = d(\varepsilon) \in \mathbb{N}$  such that the Taylor polynomial  $p = p_{d(\varepsilon)}$  of length  $d(\varepsilon)$  defined in (3.5) satisfies

$$(3.43) \quad (1 - \varepsilon)^n \int_0^1 H_n^y(\omega) d\omega \leq \int_0^1 \prod_{i=n^{2/5}}^{M_n} \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) d\omega \leq (1 + \varepsilon)^n \int_0^1 H_n^y(\omega) d\omega,$$

for every  $y \in \mathcal{D}_\infty$  and for all sufficiently large  $n \in \mathbb{N}$ .



*Proof.* Fix  $n \in \mathbb{N}$  as in the statement of the lemma. Also, fix  $y \in \mathcal{D}_\infty$  and for notational conciseness, omit all dependence on  $y$ . Then for any  $\omega \in [0, 1]$ , we can split

$$\begin{aligned}
 e^{\theta S_n(\omega)} &= \prod_{k=1}^n e^{\theta \cos(2\pi a_k \omega)} \\
 &= \underbrace{\left( \prod_{1 \leq i < n^{2/5}} \prod_{k \in \Delta_i \cup \Delta'_i} e^{\theta \cos(2\pi a_k \omega)} \right)}_{=: H^{(1)}(\omega)} \times \underbrace{\left( \prod_{n^{2/5} \leq i \leq M_n} \prod_{k \in \Delta_i} e^{\theta \cos(2\pi a_k \omega)} \right)}_{=: H^{(2)}(\omega)} \\
 &\quad \times \underbrace{\left( \prod_{n^{2/5} \leq i \leq M_n} \prod_{k \in \Delta'_i} e^{\theta \cos(2\pi a_k \omega)} \right)}_{=: H^{(3)}(\omega)} \times \underbrace{\left( \prod_{\substack{1 \leq k \leq n, \\ k \notin \bigcup_{i=1}^{M_n} (\Delta_i \cup \Delta'_i)}} e^{\theta \cos(2\pi a_k \omega)} \right)}_{=: H^{(4)}(\omega)}.
 \end{aligned}$$

We will show below that  $H^{(1)}, H^{(3)}$ , and  $H^{(4)}$  are all sub-exponential in  $n$  (that is, their logarithms are all sublinear in  $n$ ), and thus these three factors will be negligible in comparison with  $H^{(2)}$ , whose logarithm is linear in  $n$ . Indeed, first note that by construction  $H^{(1)}$  is a product of at most  $2n^{2/5}n^{1/2}$  factors, each of which is trivially between  $e^{-\theta}$  and  $e^\theta$ , so in total we have  $e^{-2\theta n^{9/10}} \leq H^{(1)}(\omega) \leq e^{2\theta n^{9/10}}$  for all  $\omega \in [0, 1]$ . Next, the product  $H^{(3)}$  contains all contributions coming from the complete short ‘‘primed’’ blocks  $\Delta'_i$ ; the purpose of these blocks was just to separate the longer blocks, and  $H^{(3)}$  is also small in comparison with  $H^{(2)}$ . Since the product  $H^{(3)}$  has a total of at most  $M_n n^{2/5} \leq n^{9/10}$  many factors, we have  $e^{-\theta n^{9/10}} \leq H^{(3)}(\omega) \leq e^{\theta n^{9/10}}$  for all  $\omega \in [0, 1]$ . Lastly, the product  $H^{(4)}$  is split off since it does not cover a full block; this is no problem, since  $H^{(4)}$  only has a small number of factors. More precisely, since by (3.40),  $n - M_n \leq M_{n+1} - M_n \leq 2\sqrt{n}$ , we have  $e^{-2\theta n^{1/2}} \leq H^{(4)}(\omega) \leq e^{2\theta n^{1/2}}$  for all  $\omega \in [0, 1]$ . Overall, this implies  $e^{-5\theta n^{9/10}} \leq H^{(1)}(\omega)H^{(3)}(\omega)H^{(4)}(\omega) \leq e^{5\theta n^{9/10}}$  for all  $\omega \in [0, 1]$ . When combined with the last display, and the observation that everything inside the integrals is positive, this yields (3.41) with  $H_n := H^{(2)}$ , which agrees with the expression in (3.42). The second estimate (3.43) is then a simple consequence of (3.41), (3.6) of Lemma 3.2 and the relations  $|\Delta_i| = \sqrt{n}$  for all  $i$  and  $M_n \leq \sqrt{n}$ .  $\square$

*Step 3* (Evaluate the integral  $\int_0^1 \prod_{i=n^{2/5}}^{M_n} \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) \, d\omega$  from (3.43)). The key idea is to first show that we can take the product  $\prod_{i=n^{2/5}}^{M_n}$  outside the integral; see (3.45). In other words, we show that there are no correlations between cosine functions with indices from different blocks  $\Delta_i$  and  $\Delta_j$ , for  $i, j, i \neq j$ , in the range, and thus, that it is possible to evaluate all integrals entirely within each block. Indeed, this was the purpose of the construction of  $\Delta_i$  and  $\Delta'_i$  in Step 1. Then we simplify each of the integrals in the product using the expansion for the polynomial  $p$  obtained in Lemma 3.3. Indeed, recall from (3.10) and (3.11) of Lemma 3.3 that for  $d \in \mathbb{N}$ , there exist non-negative coefficients  $b_j = b_j(\theta; d), j = 0, \dots, d$ , such that

for all  $k \in \mathbb{N}$ , the Taylor polynomial  $p = p_d$  satisfies

$$(3.44) \quad p(\theta \cos(2\pi a_k^y \omega)) = \sum_{j=0}^d b_j(\theta) \cos(2\pi j a_k^y \omega),$$

where the zeroth coefficient  $b_0 = b_0(\theta; d)$  is given explicitly by the finite series in (3.11). To shorten notations we suppress the dependence of  $b_0, b_1, \dots$  on  $\theta$  and  $d$  in the formulas below.

**Proposition 3.9.** *Fix  $d \in \mathbb{N}$  and  $p = p_d$  the Taylor polynomial of length  $d$ . Then, for all sufficiently large  $n$ , and every  $y \in \mathcal{D}_\infty$ ,*

$$(3.45) \quad \int_0^1 \prod_{i=n^{2/5}}^{M_n} \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) \, d\omega = \prod_{i=n^{2/5}}^{M_n} \int_0^1 \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) \, d\omega.$$

Furthermore, for  $i \in \{n^{2/5}, \dots, M_n\}$ ,

$$(3.46)$$

$$c_0^{(i)}(y) := \int_0^1 \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) \, d\omega$$

$$(3.47)$$

$$= b_0^{\sqrt{n}} + \sum \sum \sum b_0^{\sqrt{n}-\ell} b_{j_1} \cdots b_{j_\ell} \sum \frac{\ell!}{2^{\ell-1}} \mathbb{1}_{\{j_1 a_{k_1}^y + s_2 j_2 a_{k_2}^y + \cdots + s_\ell j_\ell a_{k_\ell}^y = 0\}},$$

where the four summations in the displayed formula above are taken over the ranges (in the order of appearance)

$$(3.48) \quad \sum_{1 \leq \ell \leq \sqrt{n}}, \quad \sum_{\substack{(k_1, \dots, k_\ell) \in \Delta_i, \\ k_1 > \cdots > k_\ell}}, \quad \sum_{\substack{(j_1, \dots, j_\ell) \in \{0, \dots, d\}^\ell, \\ (j_1, \dots, j_\ell) \neq (0, \dots, 0)}}, \quad \sum_{(s_2, \dots, s_\ell) \in \{-1, 1\}^{\ell-1}},$$

and the coefficients  $b_j = b_j(\theta; d)$ ,  $j = 0, \dots, d$ , are as in (3.44). Furthermore, for all sufficiently large  $i$ , given  $\ell$  and  $k_m, j_m, s_m, m = 1, \dots, \ell$  as in (3.48), we have

$$(3.49) \quad j_1 a_{k_1}^y + s_2 j_2 a_{k_2}^y + \cdots + s_\ell j_\ell a_{k_\ell}^y = 0 \quad \Rightarrow \quad j_1 y_{k_1} + s_2 j_2 y_{k_2} + \cdots + s_\ell j_\ell y_{k_\ell} = 0.$$

*Proof.* Fix  $y = (y_k)_{k \in \mathbb{N}} \in \mathcal{D}_\infty$ . We will start by establishing (3.47) and (3.49). Multiplying out the product  $\prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega))$  within a certain fixed block  $\Delta_i$ , using (3.44) and the cosine product trigonometric identity (3.17) we obtain

$$(3.50)$$

$$\prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) = b_0^{\sqrt{n}} + \sum \sum \sum b_0^{\sqrt{n}-\ell} b_{j_1} \cdots b_{j_\ell} \sum \frac{\ell!}{2^{\ell-1}} \cos(2\pi(j_1 a_{k_1}^y + s_2 j_2 a_{k_2}^y + \cdots + s_\ell j_\ell a_{k_\ell}^y) \omega),$$

where the four summations in the displayed formula above are taken over the ranges (in the order of appearance) in (3.48), and the power  $\sqrt{n}$  in the constant term  $b_0^{\sqrt{n}}$  and the coefficient  $b_0^{\sqrt{n}-\ell}$  arises from the fact that  $|\Delta_i| = \sqrt{n}$ . Note that (3.50) shows that  $\prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega))$  can be written as the sum of the constant term  $b_0^{\sqrt{n}}$  (which would be the contribution for the “independent” case; see (3.12) of Lemma 3.3) plus a linear combination of cosine functions with frequencies

$$(3.51) \quad j_1 a_{k_1}^y \pm \cdots \pm j_\ell a_{k_\ell}^y,$$

the latter following from the trigonometric identity (3.17). Assume that the expression in (3.51) is non-zero. Recall that  $a_k^y = 2^k + y_k$ , where  $y_k$  takes values in  $\mathcal{D}_k := \{h2^{k^{2/3}} : 0 \leq h \leq 2^{k^{2/3}}\}$ ; here and in the sequel we write  $2^{k^{2/3}}$  for  $2^{\lceil k^{2/3} \rceil}$  for notational conciseness. Substituting  $a_k^y = 2^k + y_k$  into (3.51), we can rewrite the frequency of the cosine function as

$$(3.52) \quad j_1(2^{k_1} + y_{k_1}) \pm \dots \pm j_\ell(2^{k_\ell} + y_{k_\ell}) = \underbrace{j_1 2^{k_1} \pm \dots \pm j_\ell 2^{k_\ell}}_{\text{fixed part}} + \underbrace{j_1 y_{k_1} \pm \dots \pm j_\ell y_{k_\ell}}_{\mathcal{D}_\infty\text{-dependent part}},$$

which is different from zero only if at least one of the parts is non-zero. Note that by (3.48), the absolute value of the fixed part in this expression, whenever it is non-zero, has a value between

$$2^{\delta_{i,\min}} \quad \text{and} \quad d\sqrt{n}2^{\delta_{i,\max}},$$

where recall  $\delta_{i,\min}$  and  $\delta_{i,\max}$ , respectively, are the smallest and largest elements of  $\Delta_i$ . Indeed, the upper bound is trivial and since  $k_1 > k_2 > \dots > k_\ell$ , we also obtain the lower bound:

$$|j_1 2^{k_1} \pm \dots \pm j_\ell 2^{k_\ell}| = |2^{k_\ell}| \cdot \underbrace{|j_1 2^{k_1 - k_\ell} \pm \dots \pm j_\ell|}_{\geq 1, \text{ since non-zero}} \geq 2^{\delta_{i,\min}}.$$

Similarly, recalling the structure of  $\mathcal{D}_k$  from (3.38), the  $\mathcal{D}_\infty$ -dependent part, whenever it is non-zero, has absolute value between

$$2^{\delta_{i,\min}^{2/3}} \quad \text{and} \quad d\sqrt{n}2^{\delta_{i,\max}^{2/3}}.$$

Thus (if both are non-zero), the absolute value of the sum of the fixed and  $\mathcal{D}_\infty$ -dependent parts always lies between

$$\frac{1}{2}2^{\delta_{i,\min}} \quad \text{and} \quad 2d\sqrt{n}2^{\delta_{i,\max}}.$$

Hence, the product  $\prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega))$  can be written as a constant term, plus a linear combination of cosine functions, each of which has a frequency that is contained in the range

$$(3.53) \quad \bar{\mathcal{R}}_i := \left[ 2^{\delta_{i,\min}^{2/3}}, d\sqrt{n}2^{\delta_{i,\max}^{2/3}} \right] \cup \left[ \frac{1}{2}2^{\delta_{i,\min}}, 2d\sqrt{n}2^{\delta_{i,\max}} \right].$$

In other words, from (3.50) and the above discussion, it follows that

$$(3.54) \quad \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) = c_0^{(i)}(y) + \sum_{m^{(i)} \in \mathcal{R}_i} c_m^{(i)}(y) \cos(2\pi m^{(i)} x)$$

for some appropriate set  $\mathcal{R}_i \subset \bar{\mathcal{R}}_i$  of positive integers, and appropriate coefficients  $c_m^{(i)}(y)$ . Note that  $c_0^{(i)}$  may differ from  $b_0^{\sqrt{n}}$  since some frequencies of the form (3.51) may vanish, and the coefficients of the corresponding cosines would then contribute to the constant term. Using (3.50) and the fact that the integral over  $[0, 1]$  of any cosine term in that expansion with a non-zero frequency vanishes, we obtain (3.47). Note that the dependence of  $c_0^{(i)} = c_0^{(i)}(y)$  on  $y$  arises because the value of the indicator

$$\mathbb{1}_{\{j_1 a_{k_1}^y + s_2 j_2 a_{k_2}^y + \dots + s_\ell j_\ell a_{k_\ell}^y = 0\}}$$

depends on  $(y_{k_1}, \dots, y_{k_\ell})$  via the values of  $a_{k_1}^y, \dots, a_{k_\ell}^y$ .

We now turn to the proof of (3.49). Recall that we constructed our blocks  $\Delta_i, \Delta'_i$  and defined  $M_n$  such that  $\delta_{i+1, \min} \geq \delta_{i, \max} + n^{2/5}$  and for  $i \leq M_n$ ,  $\delta_{i, \max} + n^{2/5} \leq n$ , see (3.39), which together with the mean-value theorem implies that

$$\delta_{i+1, \min}^{2/3} \geq \left(\delta_{i, \max} + n^{2/5}\right)^{2/3} \geq \delta_{i, \max}^{2/3} + \frac{2}{3} \frac{1}{(\delta_{i, \max} + n^{2/5})^{1/3}} n^{4/15} \geq \delta_{i, \max}^{2/3} + \frac{2}{3} n^{1/15},$$

and hence, for  $1 \leq i \leq M_n$ ,

$$(3.55) \quad 2^{\delta_{i+1, \min}} \geq 2^{n^{2/5} + \delta_{i, \max}} \quad \text{and} \quad 2^{\delta_{i+1, \min}^{2/3}} \geq 2^{\delta_{i, \max}^{2/3} + \frac{2}{3} n^{1/15}}.$$

Also, for sufficiently large  $n$  and  $n^{2/5} \leq i \leq M_n$ , note that  $\delta_{i, \min} \geq (n^{2/5})(n^{1/2}) = n^{9/10} = n^{7/30 + 2/3} \geq n^{7/30} + n^{2/3} \geq n^{7/30} + \delta_{M_n, \max}^{2/3}$ , and so

$$(3.56) \quad \frac{1}{2} 2^{\delta_{n^{2/5}, \min}} \geq \frac{2^{n^{7/30} - 1}}{d\sqrt{n}} \left(d\sqrt{n} 2^{\delta_{M_n, \max}^{2/3}}\right).$$

The last inequality shows that for all sufficiently large  $n \in \mathbb{N}$ , any “ $\mathcal{D}_\infty$ -dependent part” of a frequency that could originate from some product with indices in  $\Delta_i$  (with  $n^{2/5} \leq i \leq M_n$ ) is of a much smaller order than the smallest non-zero “fixed” part that we could encounter from such blocks, which proves (3.49). (This is why we split off the product  $H^{(1)}$  with the frequencies in  $\Delta_i \cup \Delta'_i$  for  $1 \leq i < n^{2/5}$  earlier in Lemma 3.8, since the frequencies there are so small that their fixed parts could cause correlations with the  $\mathcal{D}_\infty$ -dependent parts coming from blocks with higher indices.)

To complete the proof of the proposition, it only remains to prove (3.45). To show how our construction facilitates control of the value of the integral

$$(3.57) \quad \int_0^1 \prod_{i=n^{2/5}}^{M_n} \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) d\omega,$$

note that (3.54) implies that we have for  $y \in \mathcal{D}_\infty$  and  $\omega \in [0, 1]$ ,

$$(3.58) \quad \prod_{i=n^{2/5}}^{M_n} \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) = \prod_{i=n^{2/5}}^{M_n} \left( c_0^{(i)}(y) + \sum_{m^{(i)} \in \mathcal{R}_i} c_{m^{(i)}}^{(i)}(y) \cos(2\pi m^{(i)} \omega) \right),$$

with  $c_0^{(i)}$  as in (3.47) and  $c_{m^{(i)}}^{(i)}$  other coefficients as described above (whose precise values will not matter for what follows). When multiplying out the terms in the product on the right-hand side of (3.58), for each  $i$  in the range  $n^{2/5} \leq i \leq M_n$  we can either choose the factor  $c_0^{(i)}(y)$  or a factor of the form  $c_{m^{(i)}}^{(i)}(y) \cos(2\pi m^{(i)} \omega)$  for some  $m^{(i)} \in \mathcal{R}_i$ . That is, we can write the right-hand side of (3.58) as

$$\sum_{\mathcal{U}, \mathcal{V}} \left( \prod_{i \in \mathcal{U}} c_0^{(i)}(y) \right) \left( \prod_{i \in \mathcal{V}} \sum_{m^{(i)} \in \mathcal{R}_i} c_{m^{(i)}}^{(i)}(y) \cos(2\pi m^{(i)} \omega) \right),$$

where the sum is taken over all sets  $\mathcal{U}, \mathcal{V}$  that form a disjoint partition of  $\{n^{2/5}, \dots, M_n\}$ , i.e.,  $\mathcal{U} \cap \mathcal{V} = \emptyset$  and  $\mathcal{U} \cup \mathcal{V} = \{n^{2/5}, \dots, M_n\}$ . Assume that  $\mathcal{V}$  is non-empty. Then using the standard trigonometric identity (3.17) we can expand

$$\prod_{i \in \mathcal{V}} \sum_{m^{(i)} \in \mathcal{R}_i} c_{m^{(i)}}^{(i)}(y) \cos(2\pi m^{(i)} \omega)$$

into a linear combination of cosine-functions with frequencies of the form

$$\sum_{i \in \mathcal{V}} \pm m^{(i)}, \quad m^{(i)} \in \mathcal{R}_i.$$

Since  $\mathcal{R}_i$  is contained in the range set  $\bar{\mathcal{R}}_i$  defined in (3.53), and since we have the estimates (3.55) and (3.56) separating these respective ranges for different values of  $i$ , it is not possible that the linear combination equals zero (provided that  $n$  is large enough). Thus our construction ensures that all frequencies of cosine-functions in this linear combination are non-zero, which implies that their integrals vanish over  $[0, 1]$ , so that we have

$$\int_0^1 \prod_{i \in \mathcal{V}} \sum_{m^{(i)} \in \mathcal{S}_i} c_{m^{(i)}}^{(i)}(y) \cos(2\pi m^{(i)}\omega) d\omega = 0,$$

and consequently,

$$\int_0^1 \left( \prod_{i \in \mathcal{U}} c_0^{(i)}(y) \right) \left( \prod_{i \in \mathcal{V}} \sum_{m^{(i)} \in \mathcal{R}_i} c_{m^{(i)}}^{(i)}(y) \cos(2\pi m^{(i)}\omega) \right) d\omega = 0,$$

whenever  $\mathcal{V}$  is non-empty. Thus, the only term that actually contributes to the value of (3.57) is when all indices  $i$  are contained in  $\mathcal{U}$  and  $\mathcal{V} = \emptyset$ . The contribution of this case to the integral is

$$\int_0^1 \prod_{i=n^{2/5}}^I c_0^{(i)}(y) d\omega = \prod_{i=n^{2/5}}^I c_0^{(i)}(y),$$

so that in total we have for every  $y \in \mathcal{D}_\infty$ ,

$$\begin{aligned} \int_0^1 \prod_{i=n^{2/5}}^I \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) d\omega &= \prod_{i=n^{2/5}}^I c_0^{(i)}(y) \\ &= \prod_{i=n^{2/5}}^{M_n} \int_0^1 \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) d\omega. \end{aligned}$$

This is (3.45) and completes the proof of the proposition. □

*Step 4* (Give an explicit formula for  $\prod_{i=n^{2/5}}^{M_n} \int_0^1 \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^Y \omega)) d\omega$  which holds with large  $\mathbb{P}_Y$ -probability). We will prove the following result.

**Lemma 3.10.** *Let  $Y = (Y_k)_{k \in \mathbb{N}}$  be the sequence of independent random variables, with each  $Y_k$  uniformly distributed on the set  $\mathcal{D}_k$  defined in (3.38). Then for every  $d \in \mathbb{N}$ , with  $p = p_d$ , the Taylor polynomial of length  $d$ , and  $b_0 = b_0(\theta; d)$  as in (3.11), we have*

$$(3.59) \quad \mathbb{P}_Y \left( \prod_{i=n^{2/5}}^{M_n} \int_0^1 \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^Y \omega)) d\omega = \prod_{i=n^{2/5}}^{M_n} b_0^{\sqrt{n}} \right) \geq 1 - n^{-3/2},$$

for all sufficiently large  $n \in \mathbb{N}$ .

*Proof.* Fix  $d \in \mathbb{N}$  and set  $p = p_d$  to be the corresponding Taylor polynomial, and let  $b_j := b_j(\theta; d)$ ,  $j = 0, 1, \dots, d$ , be the associated coefficients as presented in equations (3.10) and (3.11) of Lemma 3.3 (see also (3.44)). For any  $y \in \mathcal{D}^\infty$ , let  $c_0^{(i)}(y)$  be defined as in (3.47). Combining (3.47) and (3.49) in Proposition 3.9 with the fact

that  $i \geq n^{2/5}$  in (3.47), we see that for all sufficiently large  $n \in \mathbb{N}$  and for all  $i$  in the range  $n^{2/5} \leq i \leq M_n$  we have

$$\begin{aligned}
 & \int_0^1 \prod_{k \in \Delta_i} p(\theta \cos(2\pi a_k^y \omega)) d\omega \\
 (3.60) \quad & = b_0^{\sqrt{n}} + \sum \sum \sum b_0^{\sqrt{n}-\ell} b_{j_1} \cdots b_{j_\ell} \sum \frac{\ell!}{2^{\ell-1}} \mathbb{1}_{\{j_1 y_{k_1} + s_2 j_2 y_{k_2} + \cdots + s_\ell j_\ell y_{k_\ell} = 0\}},
 \end{aligned}$$

with the summation ranges as specified in (3.48).

We now estimate the probability of the event  $\{j_1 Y_{k_1} + s_2 j_2 Y_{k_2} + \cdots + s_\ell j_\ell Y_{k_\ell} = 0\}$ . We recall that by assumption,  $Y_{k_1}, \dots, Y_{k_\ell}$  are independent discrete random variables, and that  $s_2, \dots, s_\ell$  are just some plus/minus signs. In principle the distribution of  $j_1 Y_{k_1} + s_2 j_2 Y_{k_2} + \cdots + s_\ell j_\ell Y_{k_\ell}$  could thus be calculated exactly by some convolution argument. However, for our purpose it suffices to establish a very crude bound. Observe from (3.48) that there is at least one value among  $j_1, \dots, j_\ell$  that is non-zero. Let us assume, without loss of generality, that  $j_\ell \neq 0$ . We split off the corresponding random variable  $Y_{k_\ell}$  in the indicator in (3.60), which is independent of  $Y_{k_1}, \dots, Y_{k_{\ell-1}}$  since  $k_1 > \cdots > k_\ell$ , and use the fact that by assumption  $Y_{k_\ell}$  is uniformly distributed among the  $2^{k_\ell^{2/3}} + 1$  different values in the set  $\mathcal{D}_{k_\ell}$  defined in (3.38), to obtain

$$\begin{aligned}
 & \mathbb{P}_{\mathcal{Y}}(j_1 Y_{k_1} + s_2 j_2 Y_{k_2} + \cdots + s_\ell j_\ell Y_{k_\ell} = 0) \\
 & = \sum_{a \in \mathbb{Z}} \left( \mathbb{P}_{\mathcal{Y}}(j_1 Y_{k_1} + s_2 j_2 Y_{k_2} + \cdots + s_{\ell-1} j_{\ell-1} Y_{k_{\ell-1}} = a) \underbrace{\mathbb{P}_{\mathcal{Y}}(s_\ell j_\ell Y_{k_\ell} = -a)}_{\leq 2^{-k_\ell^{2/3}}} \right) \\
 & \leq 2^{-k_\ell^{2/3}} \underbrace{\sum_{a \in \mathbb{Z}} \mathbb{P}_{\mathcal{Y}}(j_1 Y_{k_1} + s_2 j_2 Y_{k_2} + \cdots + s_{\ell-1} j_{\ell-1} Y_{k_{\ell-1}} = a)}_{=1} \\
 & \leq 2^{-\delta_{i, \min}^{2/3}},
 \end{aligned}$$

where the last inequality holds because  $(k_1, \dots, k_\ell) \in \Delta_i$ , and  $\delta_{i, \min}$  is by definition the smallest element of  $\Delta_i$ . In the quadruple sum in line (3.60) the total number of summands is at most  $\sqrt{n} \sqrt{n} \sqrt{n} (d+1)^{\sqrt{n}} 2^{\sqrt{n}}$ . Note that by construction  $\delta_{i, \min} \geq n^{9/10}$  for all  $i \geq n^{2/5}$ , so that  $2^{-\delta_{i, \min}^{2/3}} \leq 2^{(-n^{3/5})}$  for all  $i \geq n^{2/5}$ . Thus, by a union bound the  $\mathbb{P}_{\mathcal{Y}}$ -probability that there exists at least one configuration of  $\ell, (k_1, \dots, k_\ell), (j_1, \dots, j_\ell), (s_1, \dots, s_\ell)$  such that  $j_1 Y_{k_1} + s_2 j_2 Y_{k_2} + \cdots + s_\ell j_\ell Y_{k_\ell} = 0$  holds is bounded above by

$$\sqrt{n} \sqrt{n} \sqrt{n} (d+1)^{\sqrt{n}} 2^{\sqrt{n}} 2^{(-n^{3/5})}.$$

Observe that, since  $d$  is fixed, for sufficiently large  $n \in \mathbb{N}$ ,  $(d+1) \leq \sqrt{n}$  and

$$\sqrt{n} \sqrt{n} \sqrt{n} (d+1)^{\sqrt{n}} 2^{\sqrt{n}} 2^{(-n^{3/5})} \leq 2^{2\sqrt{n} \log_2(n) - n^{3/5}}$$

for which we can give the crude upper bound  $n^{-2}$  holding for all large enough  $n \in \mathbb{N}$ . Thus,

$$\mathbb{P}_{\mathcal{Y}} \left( c_0^{(i)}(Y) \neq b_0^{\sqrt{n}} \right) \leq \frac{1}{n^2}, \quad i \in \{n^{2/5}, \dots, M_n\},$$

for all sufficiently large  $n \in \mathbb{N}$ . Now, by (3.40),  $|\{n^{2/5}, \dots, M_n\}| \leq M_n \leq \sqrt{n}$ . Thus, we have

$$\mathbb{P}_Y \left( \bigcup_{i=n^{2/5}}^{M_n} \{c_0^{(i)}(Y) \neq b_0 \sqrt{n}\} \right) \leq \frac{1}{n^{3/2}},$$

for all large enough  $n \in \mathbb{N}$ , which implies the statement of the lemma. □

*Step 5* (Complete the proof of the LDP stated in Theorem D). By the definition of  $M_n$  in (3.40), we have the relation

$$\bigcup_{i=1}^{M_n+1} (\Delta_i \cup \Delta'_i) \supset \{1, \dots, n\}$$

and  $M_n \leq \sqrt{n}$ . Together with the fact that  $\Delta_i, \Delta'_i, i \in \mathbb{N}$ , are all disjoint,  $|\Delta_i| = \sqrt{n}$ , and  $|\Delta'_i| = n^{2/5}$ , this implies

$$\begin{aligned} \sum_{i=n^{2/5}}^{M_n} \sqrt{n} &\geq n - \sum_{i=n^{2/5}}^{M_n} |\Delta'_i| - \sum_{i=1}^{n^{2/5}-1} (|\Delta_i| + |\Delta'_i|) - (|\Delta_{M_n+1}| + |\Delta'_{M_n+1}|) \\ &\geq n - n^{2/5} \sqrt{n} - 2\sqrt{n} n^{2/5} - 2\sqrt{n} \\ &= n - 3n^{9/10} - 2\sqrt{n}, \end{aligned}$$

while in the other direction trivially  $\sum_{i=n^{2/5}}^{M_n} \sqrt{n} \leq n$ . Thus, for the factor  $\prod_{i=n^{2/5}}^{M_n} b_0 \sqrt{n}$  appearing in Lemma 3.10 we have the lower and upper bounds

$$(3.61) \quad b_0^{n-3n^{9/10}-2\sqrt{n}} \leq \prod_{i=n^{2/5}}^{M_n} b_0 \sqrt{n} \leq b_0^n.$$

Thus, for any fixed  $\theta \in \mathbb{R}$ , given any  $\varepsilon > 0$ , choosing  $d = d(\varepsilon)$  such that (3.43) of Lemma 3.8 holds with  $p = p_{d(\varepsilon)}$ , then invoking (3.41) as well as (3.45) of Proposition 3.9, next applying Lemma 3.10 with  $d = d(\varepsilon)$ ,  $b_0 := b_0(\theta; d(\varepsilon))$ , and finally using (3.61) we obtain

$$(3.62) \quad (1 + \varepsilon)^{-n} e^{-5\theta n^{9/10}} b_0^{n-3n^{9/10}-2\sqrt{n}} \leq \int_0^1 e^{\theta S_n^Y(\omega)} d\omega \leq b_0^n e^{5\theta n^{9/10}} (1 - \varepsilon)^{-n}$$

with  $\mathbb{P}_Y$ -probability at least  $1 - n^{-3/2}$ , for all sufficiently large  $n \in \mathbb{N}$ . Next, note that we have  $\log(1 + \varepsilon) \leq \varepsilon$ , and we can (and will) assume that  $\varepsilon > 0$  is so small that  $\log(1 - \varepsilon) \geq -2\varepsilon$ . We also have the trivial bound  $5\theta n^{9/10} + (3n^{9/10} + 2\sqrt{n}) \log b_0 = n^{9/10}(5\theta + 3 \log b_0) + (2\sqrt{n}) \log b_0 \leq n\varepsilon$  for all sufficiently large  $n \in \mathbb{N}$ . Thus, from (3.62) we can deduce that for sufficiently large  $n \in \mathbb{N}$ , with  $\mathbb{P}_Y$ -probability at least  $1 - n^{-3/2}$ ,

$$\begin{aligned} -2\varepsilon &\leq -\log(1 + \varepsilon) - \frac{1}{n} \left( 5\theta n^{9/10} + (3n^{9/10} + 2\sqrt{n}) \log b_0 \right) \\ &\leq \frac{1}{n} \log \left( \int_0^1 e^{\theta S_n^Y(\omega)} d\omega \right) - \log b_0 \\ &\leq \frac{1}{n} [5\theta n^{9/10}] - \log(1 - \varepsilon) \\ &\leq 3\varepsilon. \end{aligned}$$

This implies that for all sufficiently large  $n \in \mathbb{N}$ ,

$$\mathbb{P}_Y \left( \left| \frac{1}{n} \log \left( \int_0^1 e^{\theta S_n^Y(\omega)} d\omega \right) - \log b_0 \right| \leq 3\varepsilon \right) \geq 1 - n^{-3/2}.$$

By the Borel-Cantelli lemma, with  $\mathbb{P}_Y$ -probability equal to one only finitely many exceptional events occur. This implies that  $\mathbb{P}_Y$ -almost surely we have

$$(3.63) \quad \limsup_{n \rightarrow \infty} \left| \frac{1}{n} \log \left( \int_0^1 e^{\theta S_n^Y(\omega)} d\omega \right) - \log b_0 \right| \leq 3\varepsilon.$$

Recall from (3.15) that  $b_0(\theta; d(\varepsilon))$  is a finite polynomial approximation to the modified Bessel function  $B_0(\theta)$ , the moment generating function defined in (2.7), and that  $b_0(\theta; d(\varepsilon))$  can be made arbitrarily close to  $B_0(\theta)$  by choosing the degree  $d = d(\varepsilon)$  sufficiently large. Thus, letting  $\varepsilon \rightarrow 0$  (and hence  $d(\varepsilon) \rightarrow +\infty$ ) and using (3.63) together with (2.8), we derive, for every fixed  $\theta \in \mathbb{R}$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left( \int_0^1 e^{\theta \sum_{k=1}^n \cos(2\pi a_k \omega)} d\omega \right) = \log B_0(\theta) = \tilde{\Lambda}(\theta) \quad \mathbb{P}_Y\text{-a.s.}$$

Since  $\tilde{\Lambda}(\theta)$  is a continuous (in fact, differentiable) function in  $\theta$ , we can deduce that  $\mathbb{P}_Y$ -almost surely this result holds for all  $\theta \in \mathbb{R}$ : for  $\mathbb{P}_Y$ -almost all realizations of the random sequence  $Y$ , or equivalently,  $a_1^Y, a_2^Y, \dots$ , we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left( \int_0^1 e^{\theta S_n^Y(\omega)} d\omega \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left( \int_0^1 e^{\theta \sum_{k=1}^n \cos(2\pi a_k^Y \omega)} d\omega \right) = \tilde{\Lambda}(\theta)$$

for all  $\theta \in \mathbb{R}$ .

Together with the Gärtner-Ellis theorem, Theorem 3.1, this proves the desired result.

### APPENDIX A. PROOF OF PROPOSITION 2.4

Fix an integer  $q \in \{2, 3, \dots\}$ . For  $m \in \mathbb{N}$  and  $n \in \mathbb{N}$  recall that  $A_m(n)$  denotes the number of solutions to the equation

$$(A.1) \quad \sum_{i=1}^m \varepsilon_i q^{k_i} = 0$$

in the unknowns  $k_1, \dots, k_m \in \{1, \dots, n\}$  and  $\varepsilon_1, \dots, \varepsilon_m \in \{+1, -1\}$ .

**Proposition A.1.** *Fix  $m \in \mathbb{N}$ . Then, the function  $A_m(n)$  restricted to the values  $n \geq m - 2$  is a polynomial in  $n$  of degree at most  $\lfloor m/2 \rfloor$ .*

*Proof.* Let  $A_{m,p_1,p_2}(n)$  be the number of representations of zero as a sum of signed powers of 2 which begins with  $p_1$  terms of the form  $+q^1$  followed by  $p_2$  terms of the form  $-q^1$  and does not contain any more  $\pm q$ -terms. More precisely, for  $p_1, p_2 \in \mathbb{N}_0$  such that  $p_1 + p_2 \leq m$ , we define  $A_{m,p_1,p_2}(n)$  to be the number of solutions to (A.1) such that

$$\begin{aligned} k_1 &= \dots = k_{p_1+p_2} = 1, \\ \varepsilon_1 &= \dots = \varepsilon_{p_1} = +1, \\ \varepsilon_{p_1+1} &= \dots = \varepsilon_{p_1+p_2} = -1, \\ k_i &\in \{2, \dots, n\} \text{ for } i \in \{p_1 + p_2 + 1, \dots, n\}. \end{aligned}$$



Since in any general solution to (A.1) the terms  $\pm q$  can appear at arbitrary positions, we have

$$A_m(n) = \sum_{\substack{p_1, p_2 \geq 0 \\ p_1 + p_2 \leq m}} \binom{m}{p_1 + p_2} \binom{p_1 + p_2}{p_1} A_{m, p_1, p_2}(n).$$

To establish Proposition A.1 it suffices to prove the following two claims for all  $\ell \in \mathbb{N}$ :

- (a)  $A_{\ell, 0, 0}(n)$  is a polynomial in  $n$  of degree at most  $\lfloor \ell/2 \rfloor$  in the range  $n \geq \ell - 2$ .
- (b) For  $(p_1, p_2) \neq (0, 0)$ ,  $A_{\ell, p_1, p_2}(n)$  is a polynomial in  $n$  of degree at most  $\lfloor \ell/2 \rfloor - 1$  in the range  $n \geq \ell - 3$ .

First of all, observe that these claims are true for  $\ell = 1$  and  $\ell = 2$  because

$$A_{1, 0, 0}(n) = 0, \quad A_{2, 0, 0}(n) = 2n - 2, \quad A_{2, 1, 1}(n) = 2, \quad A_{2, 0, 1}(n) = A_{2, 1, 0}(n) = 0.$$

For larger values of  $\ell$ , we shall prove these claims by induction. The inductive argument is based on certain recurrence relations for the functions  $A_{m, p_1, p_2}(n)$  that we now derive.

*Case 1.* Let first  $p_1 = p_2 = p \in \mathbb{N}_0$ . Then, in (A.1) we can cancel the  $+q$ -terms with the  $-q$ -terms, which yields a representation of 0 as a sum of  $\pm q^2, \pm q^3, \dots, \pm q^n$ , the total number of terms being  $m - p_1 - p_2$ . Dividing all terms by  $q$ , we obtain a representation of 0 as a sum of  $m - p_1 - p_2$  terms of the form  $\pm q, \pm q^2, \dots, \pm q^{n-1}$ . The number of such representations is  $A_{m-p_1-p_2}(n-1)$ . Hence, we arrive at

$$(A.2) \quad A_{m, p, p}(n) = A_{m-2p}(n-1) = \sum_{\substack{r_1, r_2 \geq 0 \\ r_1 + r_2 \leq m-2p}} \binom{m-2p}{r_1 + r_2} \binom{r_1 + r_2}{r_1} A_{m-2p, r_1, r_2}(n-1).$$

*Case 2.* Let now  $p_1 > p_2$ . Then, in the representation (A.1) we can cancel  $p_2$  terms of the form  $+q^1$  with the same number  $p_2$  of terms of the form  $-q^1$ . The resulting representation of 0 contains  $p_1 - p_2 > 0$  terms of the form  $+q^1$  and  $m - p_1 - p_2$  terms of the form  $\pm q^2, \pm q^3, \dots, \pm q^n$ . If  $p_1 - p_2$  is not divisible by  $q$ , then  $A_{m, p_1, p_2}(n) = 0$  because the sum on the left-hand side of (A.1) is not divisible by  $q^2$ . So, assume that  $p_1 - p_2 = sq$  for some  $s \in \mathbb{N}$ . Divide the remaining  $p_1 - p_2$  terms of the form  $+q^1$  into  $s$  groups of the form  $+q^1 + \dots + q^1$ , each consisting of  $q$  terms, and replace each group by  $+q^2$ . We obtain  $s$  terms of the form  $q^2$ . However, we have also to take care of the terms of the form  $\pm q^2$  that can appear among the  $m - p_1 - p_2$  terms of the form  $\pm q^2, \pm q^3, \dots, \pm q^n$ . Let  $r_1$ , respectively,  $r_2$ , be the number of the terms  $+q^2$ , respectively,  $-q^2$ , among these  $m - p_1 - p_2$  terms. Dividing all terms by  $q$ , we obtain a representation of 0 starting with  $s = (p_1 - p_2)/q$  terms of the form  $+q^1$ , followed by a sum of  $m - p_1 - p_2$  terms of the form  $\pm q^1, \pm q^2, \dots, \pm q^{n-1}$ , among which  $r_1$  terms are of the form  $+q^1$  and  $r_2$  terms are of the form  $-q^1$ . Since the positions of these terms can be arbitrary among the  $m - p_1 - p_2$  terms, we arrive at the identity

$$(A.3) \quad A_{m, p_1, p_2}(n) = \sum_{\substack{r_1, r_2 \geq 0 \\ r_1 + r_2 \leq m - p_1 - p_2}} \binom{m - p_1 - p_2}{r_1 + r_2} \binom{r_1 + r_2}{r_1} A_{s+m-p_1-p_2, s+r_1, r_2}(n-1),$$

which holds if  $p_1 - p_2 = sq$  for  $s \in \mathbb{N}$ .

Case 3. Similar arguments show that in the case when  $p_1 < p_2$  we have  $A_{m,p_1,p_2}(n) = 0$  if  $p_2 - p_1$  is not divisible by  $q$  and

(A.4)

$$A_{m,p_1,p_2}(n) = \sum_{\substack{r_1, r_2 \geq 0 \\ r_1 + r_2 \leq m - p_1 - p_2}} \binom{m - p_1 - p_2}{r_1 + r_2} \binom{r_1 + r_2}{r_1} A_{s+m-p_1-p_2, r_1, s+r_2}(n-1),$$

if  $p_2 - p_1 = sq$  for some  $s \in \mathbb{N}$ .

We are now in position to prove claims (a) and (b) by induction. As already mentioned above, the claims are true for  $\ell = 1, 2$ . Assume that the claims are true for  $\ell = 1, \dots, m - 1$  with some  $m \in \{3, 4, \dots\}$ . We prove them for  $\ell = m$ .

Case A. Consider first the case when  $(p_1, p_2) \neq (0, 0)$ . Then, (A.2), (A.3), (A.4) yield a representation of  $A_{m,p_1,p_2}(n)$  as a linear combination of the terms  $A_{\ell, r_1, r_2}(n-1)$  with  $\ell < m$ . Applying the induction assumption, we obtain that  $A_{m,p_1,p_2}(n)$  is a polynomial in  $n$  of degree at most  $\lfloor m/2 \rfloor - 1$  in the range  $n \geq m - 3$ . In the individual cases, this can be seen as follows:

- Case 1: If  $p_1 = p_2 = p \neq 0$ , then from (A.2) we have  $\ell = m - 2p < m$ . By the induction assumptions (a) and (b), the terms  $A_{m-2p, r_1, r_2}(n-1)$  appearing in (A.2) are polynomials in  $(n-1)$  of degree at most  $\lfloor (m-2p)/2 \rfloor \leq \lfloor m/2 \rfloor - 1$  in the range  $n - 1 \geq m - 2p - 2$ , which lies in the range  $n \geq m - 3$ .
- Case 2: If  $p_1 > p_2$  and  $p_1 - p_2 = sq$  for  $s \in \mathbb{N}$ , then  $\ell = s + m - p_1 - p_2$ , which is strictly less than  $m$  since  $q \geq 2$ . By the induction assumption (b), the terms  $A_{s+m-p_1-p_2, s+r_1, r_2}(n-1)$  (for which we have  $s + r_1 > 0$  since  $s \in \mathbb{N}$ ) appearing in (A.3) are polynomials of  $(n-1)$  of degree at most  $\lfloor (s+m-p_1-p_2)/2 \rfloor - 1 \leq \lfloor m/2 \rfloor - 1$  in the range  $n - 1 \geq s + m - p_1 - p_2 - 3$ . This lies in the range  $n \geq m - 3$  since  $p_1 + p_2 - s = p + p_2 - (p_1 - p_2)/q > 0$  and hence, being integral, is greater than or equal to 1.
- Case 3: If  $p_2 > p_1$  and  $p_2 - p_1 = sq$  for  $s \in \mathbb{N}$ , then  $\ell = s + m - p_1 - p_2 < m$ . The remaining considerations are similar to Case 2.

In all three cases we obtain that (b) holds for  $\ell = m$ .

Case B. Consider now the case when  $p_1 = p_2 = 0$ . Then, (A.2) yields

$$A_{m,0,0}(n) = A_m(n-1) = \sum_{\substack{r_1, r_2 \geq 0 \\ r_1 + r_2 \leq m}} \binom{m-2p}{r_1+r_2} \binom{r_1+r_2}{r_1} A_{m, r_1, r_2}(n-1).$$

Separating the term with  $(r_1, r_2) = (0, 0)$ , we obtain

$$A_{m,0,0}(n) = A_{m,0,0}(n-1) + \sum_{\substack{r_1, r_2 \geq 0 \\ r_1 + r_2 \leq m \\ (r_1, r_2) \neq (0, 0)}} \binom{m-2p}{r_1+r_2} \binom{r_1+r_2}{r_1} A_{m, r_1, r_2}(n-1).$$

To each term in the sum on the right-hand side we can apply the same considerations as in Case A, due to the restriction  $(r_1, r_2) \neq (0, 0)$ . Thus, the sum on the right-hand side is a polynomial in  $n$  of degree at most  $\lfloor m/2 \rfloor - 1$  in the range  $n - 1 \geq m - 3$ . Denoting this polynomial by  $P_m(n)$ , we have

$$A_{m,0,0}(n) = A_{m,0,0}(n-1) + P_m(n)$$

for all  $n \geq m - 2$ . Iterating this, we obtain

$$A_{m,0,0}(n) = P_m(n) + P_m(n-1) + \dots + P_m(m-2) + A_{m,0,0}(m-3),$$

for all  $n \geq m - 2$ . The right-hand side is a polynomial in  $n$  of degree at most  $\lceil m/2 \rceil$ . This proves that (a) holds with  $\ell = m$ , thus completing the induction.

□

Proposition A.1 allows us to find explicit formulae for  $A_m(n)$  for every fixed  $m$  and all  $n \geq m - 2$ . This also yields the moments of the lacunary sums  $S_n$  because, as shown in Lemma 3.5, these are given by

$$\mathbb{E}[S_n^m] = \frac{A_m(n)}{2^m}, \quad m, n \in \mathbb{N}.$$

To compute  $A_m(n)$ , we can proceed as follows. Let some  $m \in \mathbb{N}$  be given. Using computer algebra, calculate the values  $A_m(n)$  for  $n = m - 2, \dots, m - 2 + \lceil m/2 \rceil$ . For example, one may just expand the Laurent polynomial

$$\left( \sum_{k=1}^n \left( x^{+q^k} + x^{-q^k} \right) \right)^m$$

and observe that  $A_m(n)$  is the coefficient of  $x^0$  there. Then, compute the unique interpolating polynomial of degree  $\lceil m/2 \rceil$  taking the same values as  $A_m(n)$  for  $n = m - 2, \dots, m - 2 + \lceil m/2 \rceil$ . By Proposition A.1, this yields a formula for  $A_m(n)$  for all  $n \geq m - 2$ . For example, for  $q = 2$  we obtained the following formula

$A_1(n) = 0$	for all $n \in \mathbb{N}$ ,
$A_2(n) = 2n$	for all $n \in \mathbb{N}$ ,
$A_3(n) = 6n - 6$	for all $n \in \mathbb{N}$ ,
$A_4(n) = 12n^2 + 18n - 48$	for all $n \geq 2$ ,
$A_5(n) = 120n^2 - 130n - 240$	for all $n \geq 3$ ,
$A_6(n) = 120n^3 + 900n^2 - 3310n + 870$	for all $n \geq 4$ ,
$A_7(n) = 2520n^3 + 840n^2 - 40446n + 48552$	for all $n \geq 5$ ,

and so on. By computing more values of  $A_m(n)$  than necessary, it is also possible to check the correctness of these formulas. Since the  $m$ -th cumulant  $\kappa_m(S_n)$  of  $S_n$  can be expressed as a polynomial of the first  $m$  moments  $\mathbb{E}[S_n], \dots, \mathbb{E}[S_n^m]$ , we obtain that  $\kappa_m(S_n)$  is a polynomial in  $n$  of degree at most  $\lceil m/2 \rceil$  for all  $n \geq m - 2$ . In fact, it is even a polynomial of degree 1. To see this, recall that the convergence of analytic functions in (2.10) is uniform on some disk around 0. Differentiating (2.10)  $m \in \mathbb{N}$  times, we get

$$\lim_{n \rightarrow \infty} \frac{1}{n} \kappa_m(S_n) = \Lambda_q^{(m)}(0),$$

which implies that  $\kappa_m(S_n)$  must be of degree 1. For example, in the case when  $q = 2$ , we obtained

$$\begin{aligned} \frac{\kappa_1(S_n)}{1!} &= 0 && \text{for all } n \in \mathbb{N}, \\ \frac{\kappa_2(S_n)}{2!} &= \frac{n}{4} && \text{for all } n \in \mathbb{N}, \\ \frac{\kappa_3(S_n)}{3!} &= \frac{n-1}{8} && \text{for all } n \in \mathbb{N}, \\ \frac{\kappa_4(S_n)}{4!} &= \frac{3n-8}{64} && \text{for all } n \geq 2, \\ \frac{\kappa_5(S_n)}{5!} &= \frac{-n-24}{384} && \text{for all } n \geq 3, \\ \frac{\kappa_6(S_n)}{6!} &= \frac{-115n-51}{4608} && \text{for all } n \geq 4, \\ \frac{\kappa_7(S_n)}{7!} &= \frac{916-393n}{15360} && \text{for all } n \geq 5, \end{aligned}$$

and so on. This yields the first few terms in the Taylor expansion of  $\Lambda_2$ . Since  $I_2'$  is the inverse function of  $\Lambda_2'$ , this easily yields the Taylor expansion of  $I_2$  stated in Proposition 2.4.

#### REFERENCES

- [1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*, 1964.
- [2] Dylan Airey and Bill Mance, *Normality of different orders for Cantor series expansions*, *Nonlinearity* **30** (2017), no. 10, 3719–3742, DOI 10.1088/1361-6544/aa82f0. MR3705133
- [3] Dylan Airey, Bill Mance, and Joseph Vandehey, *Normal number constructions for Cantor series with slowly growing bases*, *Czechoslovak Math. J.* **66(141)** (2016), no. 2, 465–480, DOI 10.1007/s10587-016-0269-7. MR3519615
- [4] Christoph Aistleitner and István Berkes, *On the central limit theorem for  $f(n_k x)$* , *Probab. Theory Related Fields* **146** (2010), no. 1-2, 267–289, DOI 10.1007/s00440-008-0190-6. MR2550364
- [5] Sergio Alberverio, Iryna Garko, Muslem Ibragim, and Grygoriy Torbin, *Non-normal numbers: full Hausdorff dimensionality vs zero dimensionality*, *Bull. Sci. Math.* **141** (2017), no. 2, 1–19, DOI 10.1016/j.bulsci.2016.04.001. MR3614114
- [6] Viviane Baladi, *Positive transfer operators and decay of correlations*, *Advanced Series in Nonlinear Dynamics*, vol. 16, World Scientific Publishing Co., Inc., River Edge, NJ, 2000, DOI 10.1142/9789812813633. MR1793194
- [7] É. Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, *Rend. Circ. Mat. Palermo (2)* **27** (1909), no. 1, 247–271.
- [8] Anne Broise, *Transformations dilatantes de l'intervalle et théorèmes limites* (French, with English and French summaries), *Astérisque* **238** (1996), 1–109. Études spectrales d'opérateurs de transfert et applications. MR1634271
- [9] Włodzimierz Bryc, *On the large deviation principle for stationary weakly dependent random fields*, *Ann. Probab.* **20** (1992), no. 2, 1004–1030. MR1159583
- [10] G. Cantor, *Ueber die einfachen zahlensysteme*, *Z. Math. Phys.*, **14** (1869), 121–128.
- [11] J.-R. Chazottes and P. Collet, *Almost-sure central limit theorems and the Erdős-Rényi law for expanding maps of the interval*, *Ergodic Theory Dynam. Systems* **25** (2005), no. 2, 419–441, DOI 10.1017/S0143385704000550. MR2129104
- [12] C. M. Colebrook and J. H. B. Kemperman, *On non-normal numbers*, *Nederl. Akad. Wetensch. Proc. Ser. A* 71=Indag. Math. **30** (1968), 1–11. MR0227128

- [13] Pierre Collet, *Some ergodic properties of maps of the interval*, Dynamical systems (Temuco, 1991/1992), Travaux en Cours, vol. 52, Hermann, Paris, 1996, pp. 55–91. MR1600931
- [14] Jean-Pierre Conze and Stéphane Le Borgne, *Limit law for some modified ergodic sums*, Stoch. Dyn. **11** (2011), no. 1, 107–133, DOI 10.1142/S021949371100319X. MR2771345
- [15] H. Cramér, *Sur un nouveau théorème-limite de la théorie des probabilités*, Actualités Scientifiques Industrielles **736** (1938), 5–23.
- [16] H. Cramér and H. Touchette, *On a new limit theorem in probability theory* (Translation of ‘Sur un nouveau théorème-limite de la théorie des probabilités’), arXiv e-prints, 2018.
- [17] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition, DOI 10.1007/978-3-642-03311-7. MR2571413
- [18] Manfred Denker and Marc Kesseböhmer, *Thermodynamic formalism, large deviation, and multifractals*, Stochastic climate models (Chorin, 1999), Progr. Probab., vol. 49, Birkhäuser, Basel, 2001, pp. 159–169. MR1948295
- [19] Manfred Denker and Matthew Nicol, *Erdős-Rényi laws for dynamical systems*, J. Lond. Math. Soc. (2) **87** (2013), no. 2, 497–508, DOI 10.1112/jlms/jds060. MR3046282
- [20] Michael Drmota and Robert F. Tichy, *Sequences, discrepancies and applications*, Lecture Notes in Mathematics, vol. 1651, Springer-Verlag, Berlin, 1997, DOI 10.1007/BFb0093404. MR1470456
- [21] Tanja Eisner, Bálint Farkas, Markus Haase, and Rainer Nagel, *Operator theoretic aspects of ergodic theory*, Graduate Texts in Mathematics, vol. 272, Springer, Cham, 2015, DOI 10.1007/978-3-319-16898-2. MR3410920
- [22] P. Erdős and A. Rényi, *On Cantor’s series with convergent  $\sum 1/q_n$* , Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **2** (1959), 93–109. MR126414
- [23] P. Erdős and I. S. Gál, *On the law of the iterated logarithm. I, II*, Nederl. Akad. Wetensch. Proc. Ser. A. **58** = Indag. Math. **17** (1955), 65–76, 77–84. MR0069309
- [24] Valentin Féray, Pierre-Loïc Méliot, and Ashkan Nikeghbali, *Mod- $\phi$  convergence*, Springer-Briefs in Probability and Mathematical Statistics, Springer, Cham, 2016. Normality zones and precise deviations, DOI 10.1007/978-3-319-46822-8. MR3585777
- [25] Ferdinánd Filip and Jan Šustek, *Normal numbers and Cantor expansions*, Unif. Distrib. Theory **9** (2014), no. 2, 93–101. MR3430812
- [26] Lorenz Frühwirth, Michael Juhos, and Joscha Prochno, *The large deviation behavior of lacunary sums*, Monatsh. Math. **199** (2022), no. 1, 113–133, DOI 10.1007/s00605-022-01733-x. MR4469810
- [27] K. Fukuyama, *The law of the iterated logarithm for discrepancies of  $\{\theta^n x\}$* , Acta Math. Hungar. **118** (2008), no. 1-2, 155–170, DOI 10.1007/s10474-007-6201-8. MR2378547
- [28] Katusi Fukuyama and Sho Miyamoto, *Metric discrepancy results for Erdős-Fortet sequence*, Studia Sci. Math. Hungar. **49** (2012), no. 1, 52–78, DOI 10.1556/SScMath.2011.1186. MR3059788
- [29] V. F. Gapoškin, *The central limit theorem for certain weakly dependent sequences* (Russian), Teor. Veroyatnost. i Primenen. **15** (1970), 666–684. MR0282394
- [30] J. Grigull, *Große Abweichungen und Fluktuationen für Gleichgewichtsmaße rationaler Abbildungen*, Dissertation, Georg-August-Universität Göttingen, 1993.
- [31] M. Kac, *On the distribution of values of sums of the type  $\sum f(2^k t)$* , Ann. of Math. (2) **47** (1946), 33–49, DOI 10.2307/1969033. MR15548
- [32] M. Kac, *Probability methods in some problems of analysis and number theory*, Bull. Amer. Math. Soc. **55** (1949), 641–665, DOI 10.1090/S0002-9904-1949-09242-X. MR31504
- [33] Tosio Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition. MR1335452
- [34] M. Kesseböhmer, *Multifraktale und Asymptotiken grosser Deviationen*, Dissertation, Georg-August-Universität Göttingen, 1999.
- [35] A. Khintchine and A. Kolmogorov, *Über Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden*, Rec. Math. Moscou, **32** (1925), 668–677.
- [36] Yuri Kifer, *Large deviations in dynamical systems and stochastic processes*, Trans. Amer. Math. Soc. **321** (1990), no. 2, 505–524, DOI 10.2307/2001571. MR1025756
- [37] A. Kolmogoroff, *Une contribution à l’étude de la convergence des séries de Fourier*, Fund. Math. **5** (1924), no. 1, 96–97.

- [38] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Pure and Applied Mathematics, Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974. MR0419394
- [39] Gunther Leobacher and Joscha Prochno, *Statistical independence in mathematics—the key to a Gaussian law*, Math. Semesterber. **68** (2021), no. 1, 69–104, DOI 10.1007/s00591-020-00287-z. MR4227856
- [40] Artur O. Lopes, *Entropy and large deviation*, Nonlinearity **3** (1990), no. 2, 527–546. MR1054587
- [41] Greg Martin, *Absolutely abnormal numbers*, Amer. Math. Monthly **108** (2001), no. 8, 746–754, DOI 10.2307/2695618. MR1865662
- [42] L. Olsen, *Extremely non-normal numbers*, Math. Proc. Cambridge Philos. Soc. **137** (2004), no. 1, 43–53, DOI 10.1017/S0305004104007601. MR2075041
- [43] Steven Orey and Stephan Pelikan, *Large deviation principles for stationary processes*, Ann. Probab. **16** (1988), no. 4, 1481–1495. MR958198
- [44] P. Erdős and A. Rényi, *Some further statistical properties of the digits in Cantor’s series* (English, with Russian summary), Acta Math. Acad. Sci. Hungar. **10** (1959), 21–29 (unbound insert), DOI 10.1007/BF02063287. MR107631
- [45] C.-E. Pfister and W. G. Sullivan, *Large deviations estimates for dynamical systems without the specification property. Applications to the  $\beta$ -shifts*, Nonlinearity **18** (2005), no. 1, 237–261, DOI 10.1088/0951-7715/18/1/013. MR2109476
- [46] Walter Philipp, *Limit theorems for lacunary series and uniform distribution mod 1*, Acta Arith. **26** (1974/75), no. 3, 241–251, DOI 10.4064/aa-26-3-241-251. MR379420
- [47] A. D. Pollington, *The Hausdorff dimension of a set of normal numbers*, Pacific J. Math. **95** (1981), no. 1, 193–204. MR631669
- [48] Hans Rademacher, *Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen* (German), Math. Ann. **87** (1922), no. 1-2, 112–138, DOI 10.1007/BF01458040. MR1512104
- [49] Alfréd Rényi, *On the distribution of the digits in Cantor’s series* (Hungarian, with English and Russian summaries), Mat. Lapok **7** (1956), 77–100. MR99968
- [50] R. Tyrrell Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970. MR0274683
- [51] Marek Rychlik, *Bounded variation and invariant measures*, Studia Math. **76** (1983), no. 1, 69–80, DOI 10.4064/sm-76-1-69-80. MR728198
- [52] R. Salem and A. Zygmund, *On lacunary trigonometric series*, Proc. Nat. Acad. Sci. U.S.A. **33** (1947), 333–338, DOI 10.1073/pnas.33.11.333. MR22263
- [53] R. Salem and A. Zygmund, *La loi du logarithme itéré pour les séries trigonométriques lacunaires* (French), Bull. Sci. Math. (2) **74** (1950), 209–224. MR39828
- [54] Daniel Schnellmann, *Law of iterated logarithm and invariance principle for one-parameter families of interval maps*, Probab. Theory Related Fields **162** (2015), no. 1-2, 365–409, DOI 10.1007/s00440-014-0575-7. MR3350048
- [55] Pál Turán, *On the distribution of “digits” in Cantor-systems* (Hungarian, with English and Russian summaries), Mat. Lapok **7** (1956), 71–76. MR99967
- [56] S. R. S. Varadhan, *Large deviations*, Ann. Probab. **36** (2008), no. 2, 397–419, DOI 10.1214/07-AOP348. MR2393987
- [57] Mary Weiss, *The law of the iterated logarithm for lacunary trigonometric series*, Trans. Amer. Math. Soc. **91** (1959), 444–469, DOI 10.2307/1993258. MR108681
- [58] Lai-Sang Young, *Large deviations in dynamical systems*, Trans. Amer. Math. Soc. **318** (1990), no. 2, 525–543, DOI 10.2307/2001318. MR975689
- [59] Michel Zinsmeister, *Thermodynamic formalism and holomorphic dynamical systems*, SMF/AMS Texts and Monographs, vol. 2, American Mathematical Society, Providence, RI; Société Mathématique de France, Paris, 2000. Translated from the 1996 French original by C. Greg Anderson. MR1724307
- [60] A. Zygmund, *On the convergence of lacunary trigonometric series*, Fund. Math. **16** (1930), no. 3, 90–107.

INSTITUTE OF ANALYSIS AND NUMBER THEORY, GRAZ UNIVERSITY OF TECHNOLOGY, STEYR-  
ERGASSE 30/II, 8010 GRAZ, AUSTRIA

*Email address:* [aistleitner@math.tugraz.at](mailto:aistleitner@math.tugraz.at)

DEPARTMENT OF MATHEMATICS, TECHNICAL UNIVERSITY OF MUNICH, PARKRING 11, 85748  
GARCHING, GERMANY

*Email address:* [nina.gantert@tum.de](mailto:nina.gantert@tum.de)

INSTITUTE FOR MATHEMATICAL STOCHASTICS, UNIVERSITY OF MÜNSTER, ORLÉANS-RING 10,  
48149 MÜNSTER, GERMANY

*Email address:* [zakhar.kabluchko@uni-muenster.de](mailto:zakhar.kabluchko@uni-muenster.de)

FACULTY OF COMPUTER SCIENCE AND MATHEMATICS, UNIVERSITY OF PASSAU, INNSTRASSE 33,  
94032 PASSAU, GERMANY

*Email address:* [joscha.prochno@uni-passau.de](mailto:joscha.prochno@uni-passau.de)

DIVISION OF APPLIED MATHEMATICS, BROWN UNIVERSITY, 182 GEORGE STREET, PROVIDENCE,  
RHODE ISLAND 02912

*Email address:* [kavita\\_ramanan@brown.edu](mailto:kavita_ramanan@brown.edu)