MOMENTS OF RANDOM MULTIPLICATIVE FUNCTIONS, I: LOW MOMENTS, BETTER THAN SQUAREROOT CANCELLATION, AND CRITICAL MULTIPLICATIVE CHAOS

ADAM J. HARPER

Mathematics Institute, Zeeman Building, University of Warwick, Coventry CV4 7AL,
England, UK;
email: A.Harper@warwick.ac.uk

Abstract

We determine the order of magnitude of $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q}$, where $f(n)$ is a Steinhaus or Rademacher random multiplicative function, and $0 \leqslant q \leqslant 1$. In the Steinhaus case, this is equivalent to determining the order of $\lim_{T \to \infty} \frac{1}{T} \int_0^T |\sum_{n \leqslant x} n^{-it}|^{2q} \, dt$.

In particular, we find that $\mathbb{E}|\sum_{n \leqslant x} f(n)| \asymp \sqrt{x}/(\log\log x)^{1/4}$. This proves a conjecture of Helson that one should have better than squareroot cancellation in the first moment and disproves counter-conjectures of various other authors. We deduce some consequences for the distribution and large deviations of $\sum_{n \leqslant x} f(n)$.

The proofs develop a connection between $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q}$ and the qth moment of a critical, approximately Gaussian, multiplicative chaos and then establish the required estimates for that. We include some general introductory discussion about critical multiplicative chaos to help readers unfamiliar with that area.

2010 Mathematics Subject Classification: 11N37 (primary); 11L40, 11K65, 60G15, 60G57 (secondary)

1. Introduction

Preliminary remark on notation. We will say a number n is y-smooth if all prime factors of n are $\leqslant y$. We will generally use the letter p to denote primes. Unless mentioned otherwise, the letters c, C will be used to denote positive constants, c usually being a small constant and C a large one. We write $f(x) = O(g(x))$ and $f(x) \ll g(x)$, both of which mean that there exists C such that $|f(x)| \leqslant Cg(x)$,

for all x. In a few places, this notation will be adorned with a subscript parameter (for example, $O_\epsilon(\cdot)$ and $\ll_\delta$), meaning that the implied constant C is allowed to depend on that parameter. We write $f(x) \asymp g(x)$ to mean that $g(x) \ll f(x) \ll g(x)$, in other words, that $cg(x) \leqslant |f(x)| \leqslant Cg(x)$ for some c, C, for all x.

Let $(f(p))_{p \text{ prime}}$ be a sequence of independent Steinhaus random variables, that is, independent random variables distributed uniformly on the unit circle $\{|z| = 1\}$. We define a *Steinhaus random multiplicative function* f by setting $f(n) := \prod_{p^a \| n} f(p)^a$ for all natural numbers n (where $p^a \| n$ means that p^a is the highest power of the prime p that divides n, so $n = \prod_{p^a \| n} p^a$). Thus, f is a random function taking values in the complex unit circle, that is, totally multiplicative. An alternative model is to let $(f(p))_{p \text{ prime}}$ be independent Rademacher random variables, taking values ± 1 with probability $1/2$ each. Then we define a *Rademacher random multiplicative function* f, supported on squarefree numbers n (that is, numbers n not divisible by any squares of primes) only, by $f(n) := \prod_{p|n} f(p)$.

Random multiplicative functions have attracted quite a lot of attention as models for functions of number theoretic interest: for example, Rademacher random multiplicative functions were introduced by Wintner [31] as a model for the Möbius function $\mu(n)$. In the Steinhaus case, for any real $q \geqslant 0$ and any given x, we have

$$\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{2q} = \lim_{T \to \infty} \frac{1}{T} \int_0^T \left|\sum_{n \leqslant x} n^{-it}\right|^{2q} dt$$

$$= \lim_{\text{prime } p \to \infty} \frac{1}{p-1} \sum_{\chi \bmod p} \left|\sum_{n \leqslant x} \chi(n)\right|^{2q},$$

where the final sum is over the $p - 1$ Dirichlet characters $\chi \bmod p$. Thus, questions about moments of a Steinhaus random multiplicative function are equivalent to questions about the limiting behaviour of the so-called zeta sums or of character sums.

Since we always have $\mathbb{E}f(p) = 0$ and since numbers have unique prime factorizations, it is easy to check that the random variables $f(n)$ are *orthogonal* (that is, $\mathbb{E}f(n)\overline{f(m)} = \mathbf{1}_{n=m}|f(n)|^2$ for all n, m, where $\mathbf{1}$ denotes the indicator function), and so

$$\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^2 = \begin{cases} \displaystyle\sum_{n \leqslant x} 1 = x + O(1) & \text{in the Steinhaus case,} \\ \displaystyle\sum_{n \leqslant x, \, n \text{ squarefree}} 1 = \frac{6}{\pi^2}x + O(\sqrt{x}) & \text{in the Rademacher case.} \end{cases}$$

With this orthogonality in mind as well as the fact that the $f(n)$ are built from 'quite a lot' of independent random variables $(f(p))_{p \text{ prime}}$, there are various natural questions one might investigate that compare the behaviour of $f(n)$ with a sequence of completely independent random variables:

(1) Does one have a central limit theorem for $\dfrac{\sum_{n \leqslant x} f(n)}{\sqrt{\mathbb{E}|\sum_{n \leqslant x} f(n)|^2}}$?

(2) What is the size of the moments $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q}$ for general $q \geqslant 0$?

(3) How large are the tail probabilities $\mathbb{P}(|\sum_{n \leqslant x} f(n)| \geqslant \lambda \sqrt{x})$ for λ large? In particular, do they decay rapidly (for example, exponentially) with λ?

(4) Does one have 'law of the iterated logarithm' type almost sure bounds for $\sum_{n \leqslant x} f(n)$?

Since the values $f(n)$ have a rather intricate dependence structure, these problems are probabilistically interesting. There are also number theoretic and analytic motivations for them. For example, Helson [15] conjectured in the Steinhaus case that one should have $\mathbb{E}|\sum_{n \leqslant x} f(n)| = o(\sqrt{x})$ and observed that if this is true, then a certain generalization of Nehari's theorem from harmonic analysis is false. See Saksman and Seip's open problems paper [26] for a functional analysis perspective on Helson's conjecture and related questions. On the number theoretic side, it is well known that the Riemann hypothesis is true if and only if $\sum_{n \leqslant x} \mu(n) = O_\epsilon(x^{1/2+\epsilon})$ for all positive ϵ. In fact, a conjecture of Gonek based on the distribution of zeros of the Riemann zeta function (see Ng's paper [21]) asserts that this sum should be $O(\sqrt{x}(\log \log \log x)^{5/4})$ and infinitely often as large as that. It would be very interesting to have sharp almost sure bounds for $\sum_{n \leqslant x} f(n)$ in the Rademacher case, to compare with Gonek's conjecture.

In this paper, we will answer the second question for $0 \leqslant q \leqslant 1$ and derive some consequences for the other questions. In a companion paper [12], we also answer the second question for $q > 1$ up to factors of size $e^{O(q^2)}$. Before explaining our theorems, we briefly summarize some of the previous literature on these problems.

Regarding the moments $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q}$, when q is a fixed natural number, one can expand the $2q$th power and reduce the expectation calculation to a number theoretic counting problem. This leads to an asymptotic for the moment as $x \to \infty$ in both the Steinhaus and Rademacher cases: see the papers of Harper, Nikeghbali and Radziwiłł [13] and Heap and Lindqvist [14] for such calculations and further references. We simply note here that the moments are of the form $x^q (\log x)^{\Theta(q^2)}$ and so grow rapidly with q in a way very unlike a sum

of independent random variables. When q is not a natural number, there are no obvious methods available, so much less is known, but Helson [15] conjectured (in the Steinhaus case) that the first moment should satisfy $\mathbb{E}|\sum_{n\leqslant x} f(n)| = o(\sqrt{x})$ as $x \to \infty$. This seems surprising from a number theoretic perspective since one rarely expects to achieve better than squareroot cancellation. Exploring the conjecture, Weber [30] established various results; Bondarenko and Seip [4] proved a lower bound $\mathbb{E}|\sum_{n\leqslant x} f(n)| \gg \sqrt{x}/(\log x)^{\delta}$ for a certain small explicit $\delta > 0$; Harper, Nikeghbali and Radziwiłł [13] proved a stronger lower bound $\mathbb{E}|\sum_{n\leqslant x} f(n)| \gg \sqrt{x}/(\log\log x)^{3+o(1)}$. They also conjectured, in opposition to Helson's conjecture, that $\mathbb{E}|\sum_{n\leqslant x} f(n)|^{2q} \sim C(q)x^q$ as $x \to \infty$ for each fixed $0 \leqslant q \leqslant 1$.

Turning to the almost sure behaviour of $\sum_{n\leqslant x} f(n)$, it is known in the Rademacher case that the sum is almost surely $O(\sqrt{x}(\log\log x)^{2+\epsilon})$ for each $\epsilon > 0$ due to the work of Lau, Tenenbaum and Wu [18]. It is also known that the sum is almost surely *not* $O(\sqrt{x}/(\log\log x)^{5/2+\epsilon})$ due to the work of Harper [10]. This builds on previous works of many people, most notably Halász [9]. These results are described by Lau, Tenenbaum and Wu as 'qualitatively matching the law of iterated logarithm', so one might expect the proofs to involve Gaussian-type decay estimates for the tail probabilities $\mathbb{P}(|\sum_{n\leqslant x} f(n)| \geqslant \lambda\sqrt{x})$. However, as far as the author is aware, no bound is known for these probabilities for moderately sized λ that improves on the Chebychev upper bound $\ll 1/\lambda^2$. The Lau–Tenenbaum–Wu almost sure upper bound instead exploits the special fact that $(1/\sqrt{x})\sum_{n\leqslant x} f(n)$ changes size extremely slowly.

Finally looking at distributional questions, in the Rademacher case, it is a natural default conjecture that $\frac{\sum_{n\leqslant x} f(n)}{\sqrt{\mathbb{E}|\sum_{n\leqslant x} f(n)|^2}} \xrightarrow{d} N(0,1)$ as $x \to \infty$, but Chatterjee (see Section 6 of Hough [16]) conjectured that this should not hold. Chatterjee's conjecture was proved by the author [11] using a special conditioning argument. On the other hand, if one restricts to many natural subsums, one does have a central limit theorem: see the papers of Chatterjee and Soundararajan [5], Harper [11] and Hough [16] for examples of such theorems, proved using Stein's method, a martingale decomposition and the method of moments, respectively. It has remained an open question whether $\frac{\sum_{n\leqslant x} f(n)}{\sqrt{\mathbb{E}|\sum_{n\leqslant x} f(n)|^2}}$ has a limit distribution, and if so, what this distribution is.

1.1. Statement of results. We shall prove the following theorems, which determine the order of magnitude of $\mathbb{E}|\sum_{n\leqslant x} f(n)|^{2q}$ for all $0 \leqslant q \leqslant 1$, in both the Steinhaus and Rademacher cases.

THEOREM 1. *If $f(n)$ is a Steinhaus random multiplicative function, then uniformly for all large x and $0 \leqslant q \leqslant 1$, we have*

$$\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{2q} \asymp \left(\frac{x}{1 + (1-q)\sqrt{\log \log x}}\right)^q.$$

THEOREM 2. *If $f(n)$ is a Rademacher random multiplicative function, then uniformly for all large x and $0 \leqslant q \leqslant 1$, we have*

$$\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{2q} \asymp \left(\frac{x}{1 + (1-q)\sqrt{\log \log x}}\right)^q.$$

In particular, we find that

$$\mathbb{E}|\sum_{n \leqslant x} f(n)| \asymp \frac{\sqrt{x}}{(\log \log x)^{1/4}},$$

which proves Helson's [15] somewhat surprising conjecture that the first moment should be $o(\sqrt{x})$ and disproves the counter-conjecture of Harper, Nikeghbali and Radziwiłł [13] (see also Conjecture 1 of Heap and Lindqvist [14]).

Theorem 1 also implies a negative answer to the so-called embedding problem for Dirichlet polynomials (see Question 2 of [25] or Problem 2.1 of [26]) for all exponents $0 < 2q < 2$. It is easy to check using Riemann–Stieltjes integration that

$$\int_0^1 \left|\sum_{n \leqslant x} \frac{1}{n^{1/2+it}}\right|^{2q} dt = \int_0^1 \left|\frac{x^{1/2-it}}{1/2 - it} + O(1)\right|^{2q} dt \asymp x^q,$$

and Theorem 1 implies that

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^T \left|\sum_{n \leqslant x} n^{-it}\right|^{2q} dt = \mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{2q} = o(x^q)$$

for any fixed $0 < q < 1$, as $x \to \infty$. Thus, there cannot exist any universal constant C_{2q} such that

$$\int_0^1 |P(1/2 + it)|^{2q} \, dt \leqslant C_{2q} \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^T |P(it)|^{2q} \, dt$$

for all Dirichlet polynomials $P(s)$, as the Dirichlet polynomials $P(s) = \sum_{n \leqslant x} \frac{1}{n^s}$ for growing x provide a sequence of counterexamples.

The proofs of the theorems divide into two parts, which we try to explain now. There are technical differences between the Steinhaus and Rademacher cases, but for $0 \leqslant q \leqslant 1$, most of the behaviour and key proof ideas are identical; so we confine this introductory discussion to the Steinhaus case.

The key number theoretic tool for studying multiplicative functions is complex analysis of the corresponding Euler products and Dirichlet series,

$$F(s) = \prod_{p \leqslant x} \left(1 - \frac{f(p)}{p^s} \right)^{-1} = \sum_{\substack{n=1, \\ p|n \Rightarrow p \leqslant x}}^{\infty} \frac{f(n)}{n^s}.$$

This tool is particularly appealing for random multiplicative $f(n)$ because in the Euler product, the different factors $(1 - \frac{f(p)}{p^s})^{-1}$ are independent, whereas in the sum $\sum_{n \leqslant x} f(n)$, the contributions from the underlying independent $f(p)$ are entangled with one another in a highly nontrivial way. Thus, the work of Halász [9] and Harper [10] on almost sure lower bounds for $\sum_{n \leqslant x} f(n)$ relied on a connection with lower bounds for the Euler product, and the work of Halász and of Lau, Tenenbaum and Wu [18] on upper bounds can also partially be understood in that way, though it is not presented like that. But passing directly from $\sum_{n \leqslant x} f(n)$ to $F(s)$, for example, using Perron's formula, wastes logarithmic factors that would be fatal when trying to prove our theorems. Thus, the first stage of our proofs is to pass from $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q}$ to a corresponding expectation involving $F(s)$ in an efficient way. Having done this, the second stage is to analyse the expectation involving $F(s)$.

More precisely, we first show (very roughly speaking) that

$$\mathbb{E} \left| \sum_{n \leqslant x} f(n) \right|^{2q} \approx x^q \mathbb{E} \left(\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2 \, dt \right)^q. \tag{1}$$

This is done in a few steps. We show that it suffices to prove a comparable statement for quantities like $\mathbb{E}|\sum_{n \leqslant x, P(n) > \sqrt{x}} f(n)|^{2q}$, where $P(n)$ denotes the largest prime factor of n. (For the lower bound, this is literally true since if $\sum_{n \leqslant x, P(n) > \sqrt{x}} f(n)$ is large, then with positive conditional probability, the complete sum will also be large. For the upper bound, one splits the sum into several pieces depending on the size of $P(n)$, on various ranges, and gives a separate but similarly shaped upper bound for the expectation of each piece.) The advantage of these sums is that, by multiplicativity, we can write

$$\sum_{n \leqslant x, P(n) > \sqrt{x}} f(n) = \sum_{\sqrt{x} < p \leqslant x \text{ prime}} f(p) \sum_{m \leqslant x/p} f(m),$$

and importantly since $x/p \leqslant \sqrt{x}$, the inner sums are independent of the outer random variables $(f(p))_{\sqrt{x} < p \leqslant x}$. Now we might expect a sum

of independent $f(p)$ weighted by the 'coefficients' $\sum_{m \leqslant x/p} f(m)$ to behave, very roughly speaking, like a Gaussian with mean 0 and variance $\sum_{\sqrt{x} < p \leqslant x \text{ prime}} |\sum_{m \leqslant x/p} f(m)|^2$. If this were the case, we would have

$$\mathbb{E} \left| \sum_{n \leqslant x, P(n) > \sqrt{x}} f(n) \right|^{2q} \approx \mathbb{E} \left(\sum_{\sqrt{x} < p \leqslant x \text{ prime}} \left| \sum_{m \leqslant x/p} f(m) \right|^2 \right)^q,$$

and this statement follows rigorously from Khintchine's inequality or, for the upper bound on our range of q, a suitable application of Hölder's inequality. Now we have passed from examining $\mathbb{E}| \sum_{n \leqslant x} f(n)|^{2q}$ at a single point x to examining the average behaviour of sums of $f(n)$ up to various points x/p. Since primes p are quite well distributed, one can work to replace the sum over primes by a corresponding integral, and after changing variables, reduce to studying

$$\frac{x^q}{\log^q x} \mathbb{E} \left(\int_1^{\sqrt{x}} \left| \sum_{m \leqslant z} f(m) \right|^2 \frac{dz}{z^2} \right)^q.$$

Finally, a version of Parseval's identity applied to this integral brings us to (1).

As noted above, the foregoing argument has various antecedents that we should mention. In the work of Harper, Nikeghbali and Radziwiłł [**13**] on lower bounds for $\mathbb{E}| \sum_{n \leqslant x} f(n)|$, one follows the same steps as far as the application of Khintchine's inequality (although only for the lower bound and for $q = 1/2$) but then establishes a connection with $F(s)$ in a different and less efficient way. In the work of Halász [**9**] and of Lau, Tenenbaum and Wu [**18**] on almost sure upper bounds, one also pulls out the value of f on large primes and ends up dealing with integral averages of $|\sum_{n \leqslant x} f(n)|^2$, although these arise in a different way because from the beginning of their problem, one is averaging over a sequence of x values. Those authors never apply Parseval's identity, which could easily be done for the integral expressions they arrive at but would not in itself improve their results.

Now we try to explain our approach to analysing

$$x^q \mathbb{E} \left(\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2 \, dt \right)^q.$$

To set the scene, we note that, by Hölder's inequality,

$$x^q \mathbb{E} \left(\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2 \, dt \right)^q$$
$$\leqslant x^q \left(\frac{1}{\log x} \int_{-1/2}^{1/2} \mathbb{E}|F(1/2 + it)|^2 \, dt \right)^q \quad \forall \, 0 \leqslant q \leqslant 1.$$

A standard calculation shows $\mathbb{E}|F(1/2+it)|^2 \asymp \log x$ (see (2)), and inserting this yields the trivial upper bound $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q} \ll x^q$. The major contribution to this expected size of $|F(1/2 + it)|^2$ comes from the fairly rare event that $|\log|F(1/2 + it)| - \log\log x| \leqslant \sqrt{\log\log x}$, but if integrating over $[-1/2, 1/2]$ roughly corresponded to taking $\log x$ independent samples of $|F(1/2 + it)|$ (because $F(s)$ varies with s on a scale of $1/\log x$), one might indeed typically find a few such values of $\log|F(1/2 + it)|$ with $|t| \leqslant 1/2$. (This discussion of $\mathbb{E}|F(1/2 + it)|^2$ is similar to the analysis of the second moment of the Riemann zeta function and the values of zeta that make the major contribution to it. See the introduction to Soundararajan's paper [29], for example.) So the essence of Theorems 1 and 2 is that, when looking at $\mathbb{E}(\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2 \, dt)^q$ with q a little smaller than 1, integrating over $[-1/2, 1/2]$ *does not correspond* to taking $\log x$ independent samples of $|F(1/2 + it)|$, so the above application of Hölder's inequality is wasteful.

It turns out that $\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2 \, dt$ is fairly close to (the total mass of a truncation of) a probabilistic object called *critical multiplicative chaos*, and our analysis of it draws on ideas from that field. We shall sketch these ideas now without assuming familiarity with the area and then remark on connections with the literature.

There is some dependence between the random products $(|F(1/2+it)|^2)_{|t| \leqslant 1/2}$ since if t changes slightly, then $p^{it} = e^{it \log p}$ only changes slightly. The largest primes involved in $F(1/2 + it)$ have size $p \approx x$, which is why we expect the product to change little when t varies by less than $1/\log x$. But many primes in the product are much smaller than x, so at least the subproduct over smaller primes will remain unchanged over wider t intervals. In fact, to understand $(|F(1/2 + it)|^2)_{|t| \leqslant 1/2}$ properly one should think of the product consisting of $\log\log x$ blocks or 'scales' of primes of comparable logarithmic size, each of which remains constant on a different t scale.

In view of these nontrivial dependencies amongst the $(|F(1/2 + it)|^2)_{|t| \leqslant 1/2}$, there are certain events involving the size of different subproducts of $|F(1/2 + it)|^2$ that occur with probability close to 1 but would not do so if the products behaved independently. Now if $\mathcal{G}$ is some event and if we let $q' = (1 + q)/2$, we have that $\mathbb{E}\left(\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2 \, dt\right)^q$ is

$$= \mathbb{E}\left(\mathbf{1}_{\mathcal{G}} \frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2\right)^q$$

$$+ \mathbb{E}\left(\mathbf{1}_{\mathcal{G} \text{ fails}} \frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2\right)^q$$

$$\leqslant \mathbb{E}\left(\mathbf{1}_{\mathcal{G}}\frac{1}{\log x}\int_{-1/2}^{1/2}\left|F\left(\frac{1}{2}+it\right)\right|^2\right)^q$$

$$+ \mathbb{P}(\mathcal{G} \text{ fails})^{\frac{q'-q}{q'}}\left(\mathbb{E}\left(\frac{1}{\log x}\int_{-1/2}^{1/2}\left|F\left(\frac{1}{2}+it\right)\right|^2\right)^{q'}\right)^{\frac{q}{q'}},$$

the second line following by Hölder's inequality. We may now apply Hölder's inequality to the first term, as above, but hope to obtain additional savings because of the indicator function $\mathbf{1}_{\mathcal{G}}$. In the second term, the prefactor $\mathbb{P}(\mathcal{G} \text{ fails})^{\frac{q'-q}{q'}} = \mathbb{P}(\mathcal{G} \text{ fails})^{\frac{1-q}{2q'}}$ can supply a saving if the event $\mathcal{G}$ is sufficiently probable, in particular, if it occurs with probability larger than $1 - e^{-K/(1-q)}$, for some large K.

It turns out that if $\mathcal{G}$ is the event, not just that $\log|F(1/2 + it)| \leqslant \log\log x + K/(1 - q)$, but that all subproducts of $F(1/2 + it)$ obey a comparable bound (with $\log\log x$ replaced by the number of 'scales' involved in the subproduct), then the indicator function $\mathbf{1}_{\mathcal{G}}$ produces a saving factor of the shape $\frac{K}{(1-q)\sqrt{\log\log x}}$. Meanwhile, $\mathcal{G}$ does occur with probability larger than $1 - e^{-K/(1-q)}$. This is essentially the argument that leads to the upper bounds in Theorems 1 and 2. To make things work, one needs to develop results that allow the estimation of $\mathbb{E}\mathbf{1}_{\mathcal{G}}|F(1/2 + it)|^2$ (an analogue of Girsanov's theorem from the theory of Gaussian random variables). Some technical work is also required to allow the estimation of $\mathbb{P}(\mathcal{G} \text{ fails})$ (a discretization argument) and to handle the term $\mathbb{E}(\frac{1}{\log x}\int_{-1/2}^{1/2}|F(\frac{1}{2}+it)|^2)^{q'}$ that emerged from Hölder's inequality (we use an iterative procedure, an alternative would be to consider a sequence of different K values).

The lower bounds in Theorems 1 and 2 are proved by comparing $\mathbb{E}\frac{1}{\log x}\int_{\mathcal{L}}|F(1/2 + it)|^2\,dt$ with $\mathbb{E}(\frac{1}{\log x}\int_{\mathcal{L}}|F(1/2 + it)|^2\,dt)^2$, where $\mathcal{L} \subseteq [-1/2, 1/2]$ is a certain random subset chosen so that this second moment remains roughly the same size as the square of the first moment. In fact, we essentially choose $\mathcal{L}$ as the set of points at which $\log|F(1/2+it)| \leqslant \log\log x + 1/(1 - q)$, and all subproducts of $F(1/2 + it)$ obey a comparable bound. The idea here is that when one expands out the second moment, one needs to estimate terms $\mathbb{E}|F(1/2 + it_1)|^2|F(1/2 + it_2)|^2$, and, ideally, one wants the answer to be approximately the product $\mathbb{E}|F(1/2 + it_1)|^2\mathbb{E}|F(1/2 + it_2)|^2$. One cannot achieve exactly this when $t_1 - t_2$ is small because of dependencies between terms in the two products, but by restricting t_1, t_2 to the set $\mathcal{L}$, one can ensure $\mathbb{E}|F(1/2+it_1)|^2|F(1/2+it_2)|^2$ only blows up slowly as $t_1 - t_2$ becomes small, so one retains control when integrating over t_1, t_2. The reader might think of this calculation as the underlying motivation for our choice of $\mathcal{G}$ in our upper

bound proof, as well. The lower bound proofs are conceptually easier than the upper bounds since we do not need any discretization argument and do not need to control the contribution from points outside the nice set $\mathcal{L}$. However, on the probabilistic side, we require a two-dimensional Girsanov-type theorem allowing us to estimate $\mathbb{E}\mathbf{1}_A|F(1/2 + it_1)|^2|F(1/2 + it_2)|^2$ for certain events A related to the definition of $\mathcal{L}$, and it requires quite a lot of work to set this up carefully.

Now we make a few remarks about multiplicative chaos: see the survey of Rhodes and Vargas [23] for more about the field. The theory begins with a random function $h(\cdot)$ on some space, say the interval $[-1/2, 1/2]$. We construct a family of random measures on that space, depending on a parameter $\gamma > 0$, by defining the measure of a subset to be the integral of $e^{\gamma h(\cdot)}$ over the subset. In particular, the total measure $\int_{-1/2}^{1/2} e^{\gamma h(t)}\, dt$ is a random variable. This description is inaccurate in several respects, notably that $h(\cdot)$ is actually taken as a random *generalized* function, and one 'truncates' or 'regularizes' h on different scales (depending on a parameter ϵ, say) to form a genuine random function h_ϵ and then studies the limit behaviour of the random measure as $\epsilon \to 0$. Significant probabilistic attention is devoted to showing that, for suitable $h(\cdot)$, different regularization procedures result in the same limit behaviour. The most studied situation is Gaussian multiplicative chaos, where $h(t)$ is a mean zero Gaussian generalized function, with (regularized) variance approximately the same for each t. To obtain a nontrivial behaviour, one needs to rescale $e^{\gamma h(t)}$ at each t by its expected value $e^{(1/2)\gamma^2 \mathbb{E}h(t)^2}$. One also assumes that the collection of $h(t)$, or in fact their regularizations, has a certain logarithmic kind of covariance structure.

Let us compare with our situation, where we are interested in $\int_{-1/2}^{1/2} |F(1/2 + it)|^2\, dt = \int_{-1/2}^{1/2} e^{2\log|F(1/2+it)|}\, dt$. We can think of the length x of the random product $F(1/2 + it)$ as a regularization parameter, but we are not here interested in constructing or analysing limit objects, so much as obtaining moment information about $\int_{-1/2}^{1/2} e^{2\log|F(1/2+it)|}\, dt$ uniformly in x. Our random function $\log|F(1/2+it)| = \sum_{p \leqslant x} \log|1 - \frac{f(p)}{p^{1/2+it}}|^{-1}$ is not Gaussian but is approximately Gaussian as it is a sum of many independent components. When we rescale $\int_{-1/2}^{1/2} |F(1/2 + it)|^2\, dt$ by dividing by $\log x \asymp \mathbb{E}|F(1/2 + it)|^2$, this directly corresponds to rescaling $e^{\gamma h(t)}$ by $e^{(1/2)\gamma^2 \mathbb{E}h(t)^2}$. It also turns out that the random variables $\log|F(1/2 + it)|$ have a logarithmic covariance structure. This is why ideas from the theory of multiplicative chaos are relevant to $\int_{-1/2}^{1/2} |F(1/2 + it)|^2\, dt$.

Now the behaviour of Gaussian multiplicative chaos changes as γ increases. There is a *critical value* γ_c (depending, for example, on the dimension of the space one is working in and on the variance of $h(t)$) such that

$\int e^{\gamma h_\epsilon(t) - (1/2)\gamma^2 \mathbb{E}h_\epsilon(t)^2}$ converges to a nontrivial limit measure as $\epsilon \to 0$ when $\gamma < \gamma_c$, but converges to the zero measure when $\gamma = \gamma_c$. It is natural to expect such a transition because as γ increases, the dominant contribution to $\int e^{\gamma h_\epsilon(t) - (1/2)\gamma^2 \mathbb{E}h_\epsilon(t)^2}$ comes from larger values of $h_\epsilon(t)$, so when γ is large enough, one expects with high probability never to find such values on the whole range of integration. Number theory readers may again recognize a parallel with the analysis of moments of the Riemann zeta function. Critical multiplicative chaos, where $\gamma = \gamma_c$, is more difficult to analyse than the subcritical case, but recently Duplantier, Rhodes, Sheffield and Vargas [6] showed in some generality that if one replaces $\int e^{\gamma_c h_\epsilon(t) - (1/2)\gamma_c^2 \mathbb{E}h_\epsilon(t)^2}$ by $\int \sqrt{\mathbb{E}h_\epsilon(t)^2} e^{\gamma_c h_\epsilon(t) - (1/2)\gamma_c^2 \mathbb{E}h_\epsilon(t)^2}$, this converges to a nonzero limit measure. In our case, the exponent 2 of $|F(1/2+it)|$ corresponds to critical γ (see our earlier discussion about the main contribution to our integral coming when $\log|F(1/2 + it)| \approx \log\log x$), and we have $\mathbb{E}(\log|F(1/2 + it)|)^2 \asymp \log\log x$, so the factor $\sqrt{\mathbb{E}h_\epsilon(t)^2}$ directly suggests the factor $\sqrt{\log\log x}$ in the denominator in Theorems 1 and 2, for q away from 1.

We end with specific connections between our problem and the multiplicative chaos literature. The most relevant work is a preprint of Saksman and Webb [27], see also [28], showing a random model for $\log\zeta(1/2 + it)$ is well approximated by a perturbation of a Gaussian random field, so one can apply many results about Gaussian multiplicative chaos to the multiplicative chaos arising from that random model as well. The model for $\log\zeta(1/2 + it)$ is very close to $\log|F(1/2 + it)|$ in the Steinhaus case, so it is possible that combining (the rigorous version of) (1), Saksman and Webb's approximation and the results of Duplantier, Rhodes, Sheffield and Vargas [6] about moments of the total measure of critical Gaussian chaos (which ultimately stem from Kahane's convexity inequality and results of Hu and Shi [17] for branching random walk), one could get another proof of Theorem 1 for q bounded away from 1. It is not clear whether one could prove the full uniform version of Theorem 1, where $q = q(x)$ may be close to 1, in this way. In any case, our proofs here are self-contained. One inspiration for our arguments is a beautiful short paper of Berestycki [3], proving convergence results for subcritical Gaussian chaos using Girsanov's theorem and restricting attention to high probability good events. We are in the more delicate critical case, in a non-Gaussian setting, and trying to prove different kinds of results, but several features carry over.

We should also mention the connection between multiplicative chaos and the maxima of logarithmically correlated random processes, corresponding to $h(t)$ in our earlier discussion. This interaction is well known to workers in the area (see Section 6.4 of Rhodes and Vargas [23], for example), and the arguments of Berestycki [3] build on approaches to analysing the maxima of such processes. Roughly speaking, one could perhaps say that multiplicative chaos is a bit easier

to analyse (because exponentiating boosts the role of very large values and somewhat dilutes the interaction at nearby t), but, on the other hand, one is interested in proving sharper results about multiplicative chaos (for example, bounds that are sharp up to constants). See the paper of Arguin, Belius and Harper [1] for results on the maximum of a random model for $\log \zeta(1/2 + it)$, which again is close to $\log |F(1/2 + it)|$ in the Steinhaus case.

1.2. Some corollaries.
The following is an immediate corollary of our theorems.

COROLLARY 1. *If $f(n)$ is a Steinhaus or Rademacher random multiplicative function, then*

$$\frac{\sum_{n \leqslant x} f(n)}{\sqrt{\mathbb{E}| \sum_{n \leqslant x} f(n)|^2}} \xrightarrow{p} 0 \quad as \ x \to \infty,$$

where $\xrightarrow{p}$ denotes convergence in probability.

Proof of Corollary 1. As noted earlier, we have $\sqrt{\mathbb{E}| \sum_{n \leqslant x} f(n)|^2} \asymp \sqrt{x}$, whereas we have $\mathbb{E}| \sum_{n \leqslant x} f(n)| \asymp \sqrt{x}/(\log \log x)^{1/4}$ by Theorem 1 in the Steinhaus case or Theorem 2 in the Rademacher case. Thus, $\frac{\sum_{n \leqslant x} f(n)}{\sqrt{\mathbb{E}| \sum_{n \leqslant x} f(n)|^2}}$ converges to 0 in L^1, which is a stronger statement than convergence to 0 in probability. $\qquad\square$

Corollary 1 resolves the question of the limiting distribution of $\frac{\sum_{n \leqslant x} f(n)}{\sqrt{\mathbb{E}| \sum_{n \leqslant x} f(n)|^2}}$, which previously has generated quite a lot of discussion, exploration of the behaviour when one conditions on $f(p)$ on small primes p, and numerical simulations (see the papers of Chatterjee and Soundararajan [5] and Hough [16], for example), as well as Harper's [11] proof that the distribution is not $N(0, 1)$ in the Rademacher case. One now has an obvious and interesting follow-up question, namely what is the distribution of the properly renormalized sum $\frac{\sum_{n \leqslant x} f(n)}{\sqrt{x}/(\log \log x)^{1/4}}$? Our arguments here may be a reasonable starting point for investigating this since they show that matters substantially reduce to understanding the distribution of $\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2 + it)|^2 \, dt$, which we can try to access by understanding the distribution of the total mass of critical multiplicative chaos.

By Chebychev's inequality, we immediately have

$$\mathbb{P}\left(\left|\sum_{n\leqslant x} f(n)\right| \geqslant \lambda\sqrt{x}\right) \leqslant \frac{\mathbb{E}|\sum_{n\leqslant x} f(n)|^2}{(\lambda\sqrt{x})^2} \leqslant \frac{1}{\lambda^2} \quad \forall\, \lambda > 0.$$

It is natural to want to improve this, and one might even hope to obtain exponential decay for large λ by analogy with large deviation estimates for sums of independent random variables. However, as far as the author is aware, no improved estimate whatsoever is known for fixed x and moderately large λ (for λ larger than a suitable power of $\log x$, one gets a better bound by looking, for example, at the fourth moment $\mathbb{E}|\sum_{n\leqslant x} f(n)|^4$). In the following corollary, we obtain a small improvement of the Chebychev upper bound. We also show, perhaps surprisingly, that this is close to best possible on a wide range of λ.

COROLLARY 2. *Let x be large and let $f(n)$ be a Steinhaus or Rademacher random multiplicative function. For all $\lambda \geqslant 2$, we have*

$$\mathbb{P}\left(\left|\sum_{n\leqslant x} f(n)\right| \geqslant \lambda\frac{\sqrt{x}}{(\log\log x)^{1/4}}\right) \ll \frac{\min\{\log\lambda, \sqrt{\log\log x}\}}{\lambda^2}.$$

In addition, for all $2 \leqslant \lambda \leqslant e^{\sqrt{\log\log x}}$, we have

$$\mathbb{P}\left(\left|\sum_{n\leqslant x} f(n)\right| \geqslant \lambda\frac{\sqrt{x}}{(\log\log x)^{1/4}}\right) \gg \frac{1}{\lambda^2(\log\log x)^{O(1)}}.$$

The upper bound here is a rather direct corollary of Theorems 1 and 2, and we shall prove it immediately.

Proof of Corollary 2, upper bound. If $\lambda \geqslant e^{\sqrt{\log\log x}}$, then the result follows from Chebychev's inequality applied to $\mathbb{E}|\sum_{n\leqslant x} f(n)|^2$.

For smaller λ, for any $q \leqslant 1$, we have

$$\mathbb{P}\left(\left|\sum_{n\leqslant x} f(n)\right| \geqslant \lambda\frac{\sqrt{x}}{(\log\log x)^{1/4}}\right) \leqslant \frac{\mathbb{E}|\sum_{n\leqslant x} f(n)|^{2q}}{(\lambda\frac{\sqrt{x}}{(\log\log x)^{1/4}})^{2q}}$$

$$\ll \frac{1}{\lambda^{2q}}\frac{(\log\log x)^{q/2}}{(1+(1-q)\sqrt{\log\log x})^q}.$$

If we set $q = 1 - \delta$ with $0 \leqslant \delta \leqslant 1$, then the right-hand side is

$$\frac{1}{\lambda^2}\frac{\lambda^{2\delta}(\log\log x)^{q/2}}{(1+\delta\sqrt{\log\log x})^q} \leqslant \frac{1}{\lambda^2}\frac{\lambda^{2\delta}}{\delta} = \frac{\log\lambda}{\lambda^2}\frac{e^{2\delta\log\lambda}}{\delta\log\lambda}.$$

Choosing $\delta = \frac{1}{2\log\lambda}$ yields the claimed upper bound. $\qquad\square$

Proving the lower bound in Corollary 2 requires some additional ideas, so its proof is deferred to Section 6. Very roughly speaking, since the value of $|F(1/2+it)|^2$ does not change much on t intervals of length $1/\log x$ and since we have something like $|\sum_{n \leqslant x} f(n)| \approx \sqrt{x}(\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2+it)|^2 \, dt)^{1/2}$, it will suffice to prove that

$$\mathbb{P}\left(\max_{|t| \leqslant 1/2} |F(1/2+it)|^2 \geqslant \lambda^2 \log^2 x\right)$$
$$\gg \frac{1}{\lambda^2 (\log \log x)^{O(1)}} \quad \forall \, 2 \leqslant \lambda \leqslant e^{\sqrt{\log \log x}}.$$

This can be done by approximating the random variables $(|F(1/2+it)|^2)_{|t| \leqslant 1/2}$ by the exponentials of certain correlated Gaussian random variables and applying known results about the maximum of such Gaussians.

It seems another interesting open question to determine the exact magnitude of the probabilities in Corollary 2. They may be $\asymp 1/\lambda^2$ for any fixed λ since Barral, Kupiainen, Nikula, Saksman and Webb [2] have shown that the limiting total measure of critical multiplicative chaos has upper tails of the shape $\asymp 1/\lambda$, which would suggest tails $\asymp 1/\lambda^2$ here because of the power $1/2$ in the approximation of $|\sum_{n \leqslant x} f(n)|$ above. On a wide range of $\lambda \geqslant e^{\sqrt{\log \log x}}$, the proof of the lower bound in Corollary 2 still yields a result, of the form $e^{-(\log \lambda)^2/\log \log x}/(\lambda^2 (\log \log x)^{O(1)})$. This order of magnitude may be essentially the correct answer for $\lambda \geqslant e^{\sqrt{\log \log x}}$. It is unclear what one should expect as the correct answer between these ranges.

1.3. Organizational remarks.

In Section 2, we make rigorous the statement that $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q} \approx x^q \mathbb{E}(\frac{1}{\log x} \int_{-1/2}^{1/2} |F(1/2+it)|^2 \, dt)^q$. In Section 3, which is the longest section, we obtain various probabilistic estimates for $\mathbb{E}|F(1/2+it)|^2$ and for $\mathbb{E}\mathbf{1}_A |F(1/2+it)|^2$, for certain kinds of events A. Section 4 contains the fairly quick deduction of the upper bound parts of Theorems 1 and 2. Section 5 includes some further probabilistic estimates, generalizing the work from Section 3 to two dimensions, and then the deduction of the lower bound parts of Theorems 1 and 2. Finally, Section 6 proves the lower bound from Corollary 2, and the appendix gives proofs of two probability results on Gaussian random walks, deferred from Section 3.

Note that it will suffice to prove Theorems 1 and 2 for $2/3 \leqslant q \leqslant 1$. For the upper bounds, if we know that $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{4/3} \ll \frac{x^{2/3}}{(\log \log x)^{1/3}}$ (the result

with $q = 2/3$), then Hölder's inequality yields

$$\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{2q} \leqslant \left(\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{4/3}\right)^{3q/2} \ll \frac{x^q}{(\log \log x)^{q/2}} \quad \forall \, 0 \leqslant q \leqslant 2/3,$$

as desired. For the lower bounds, if we know that $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{4/3} \asymp \frac{x^{2/3}}{(\log \log x)^{1/3}}$ and that $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{3/2} \asymp \frac{x^{3/4}}{(\log \log x)^{3/8}}$, then, by Hölder's inequality,

$$\frac{x^{2/3}}{(\log \log x)^{1/3}} \asymp \mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{4/3}$$

$$= \mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{\frac{q/3}{3/2-2q}}\left|\sum_{n \leqslant x} f(n)\right|^{\frac{2-3q}{3/2-2q}}$$

$$\leqslant \left(\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{2q}\right)^{\frac{1}{6(3/2-2q)}}\left(\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{3/2}\right)^{\frac{4-6q}{3(3/2-2q)}}$$

$$\ll \left(\mathbb{E}\left|\sum_{n \leqslant x} f(n)\right|^{2q}\right)^{\frac{1}{6(3/2-2q)}}\left(\frac{x}{\sqrt{\log \log x}}\right)^{\frac{2-3q}{3-4q}}$$

for all $0 \leqslant q \leqslant 2/3$, which implies that $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q} \gg (\frac{x}{\sqrt{\log \log x}})^q$. Restricting to $2/3 \leqslant q \leqslant 1$ will be a useful simplification in our main arguments, making various series involving q converge and giving us access to Minkowski's inequality in certain places.

We finish with a remark on references. Less standard results that we use are stated explicitly as results in the text, and we also try to give references in line for most of the standard material in number theory and probability to help readers who are less familiar with one area or the other. The books of Gut [8] and of Montgomery and Vaughan [20] may be consulted as excellent general references for probabilistic and number theoretic background.

2. The reduction to Euler products

In this section, we shall prove four propositions that make precise and rigorous the assertion in (1), that $\mathbb{E}|\sum_{n \leqslant x} f(n)|^{2q}$ may be bounded by studying integrals of Euler products.

2.1. Some lemmas.
We begin by recording two fairly straightforward lemmas we need.

NUMBER THEORY RESULT 1 (see Lemma 2.1 of Lau, Tenenbaum and Wu [18]). *Let $0 < \delta < 1$, let $m \geqslant 1$ and suppose that $\max\{3, 2m\} \leqslant y \leqslant z \leqslant y^{10}$ and that u, v satisfy $1 < u \leqslant v(1 - y^{-\delta})$. As usual, let $\Omega(d)$ denote the total number of prime factors of d (counted with multiplicity). Then*

$$\sum_{\substack{u \leqslant d \leqslant v, \\ p|d \Rightarrow y \leqslant p \leqslant z}} m^{\Omega(d)} \ll_\delta \frac{(v-u)m}{\log y} \prod_{y \leqslant p \leqslant z} \left(1 - \frac{m}{p}\right)^{-1}.$$

In particular, for any large x and any $k \leqslant \log\log x - 5$, we have

$$\sum_{\substack{d \leqslant x, \\ p|d \Rightarrow x^{e^{-(k+1)}} \leqslant p \leqslant x^{e^{-k}}}} 1 \ll 2^{-e^k} \frac{x}{\log x}.$$

The first statement here is a slight generalization of Lemma 2.1 of Lau, Tenenbaum and Wu [18] (see also Lemma 3 of Halász [9]), in which the sum over d was restricted to squarefree numbers. The proof is short, so we give it in full.

Proof of Number Theory Result 1. Since $u > 1$, each d in the sum has at least one prime factor, say q, and this must lie in the interval $[y, z]$. So we have

$$\sum_{\substack{u \leqslant d \leqslant v, \\ p|d \Rightarrow y \leqslant p \leqslant z}} m^{\Omega(d)} \leqslant m \sum_{\substack{y \leqslant q \leqslant z, \\ q \text{ prime}}} \sum_{\substack{u/q \leqslant d \leqslant v/q, \\ p|d \Rightarrow y \leqslant p \leqslant z}} m^{\Omega(d)} \leqslant m \sum_{\substack{u/z \leqslant d \leqslant v/y, \\ p|d \Rightarrow y \leqslant p \leqslant z}} m^{\Omega(d)} \sum_{\substack{u/d \leqslant q \leqslant v/d, \\ q \text{ prime}}} 1.$$

Now $(v-u)/d \geqslant y(v-u)/v \geqslant y^{1-\delta}$ here, whilst $u/d \leqslant z \leqslant y^{10}$, so the Brun–Titchmarsh upper bound for primes in intervals is available (see, for example, Corollary 3.4 of Montgomery and Vaughan [20] or Theorem 3.9 of Montgomery and Vaughan [20] for a more general statement) and implies that $\sum_{\substack{u/d \leqslant q \leqslant v/d, \\ q \text{ prime}}} 1 \ll_\delta (v-u)/(d \log y)$. The first part of the result now follows on inserting this estimate and noting that
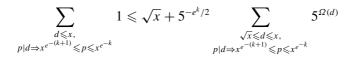
$$\sum_{\substack{u/z \leqslant d \leqslant v/y, \\ p|d \Rightarrow y \leqslant p \leqslant z}} \frac{m^{\Omega(d)}}{d} \leqslant \sum_{\substack{d: \\ p|d \Rightarrow y \leqslant p \leqslant z}} \frac{m^{\Omega(d)}}{d} = \prod_{y \leqslant p \leqslant z} \left(1 - \frac{m}{p}\right)^{-1},$$

the product certainly being well defined in view of our assumption that $y \geqslant 2m$.

To deduce the second part of the result, we note that

$$\sum_{\substack{d \leqslant x, \\ p|d \Rightarrow x^{e^{-(k+1)}} \leqslant p \leqslant x^{e^{-k}}}} 1 \leqslant \sqrt{x} + 5^{-e^k/2} \sum_{\substack{\sqrt{x} \leqslant d \leqslant x, \\ p|d \Rightarrow x^{e^{-(k+1)}} \leqslant p \leqslant x^{e^{-k}}}} 5^{\Omega(d)}$$

since a number $d \geqslant \sqrt{x}$ with all its prime factors smaller than $x^{e^{-k}}$ must have at least $e^k/2$ such factors (counted with multiplicity). We can apply the first part of the result with $\delta = 1/2$, say, and deduce that

$$5^{-e^k/2} \sum_{\substack{\sqrt{x} \leqslant d \leqslant x, \\ p|d \Rightarrow x^{e^{-(k+1)}} \leqslant p \leqslant x^{e^{-k}}}} 5^{\Omega(d)} \ll 5^{-e^k/2} \frac{x}{\log(x^{e^{-(k+1)}})} \prod_{x^{e^{-(k+1)}} \leqslant p \leqslant x^{e^{-k}}} \left(1 - \frac{5}{p}\right)^{-1}$$

$$\ll 5^{-e^k/2} e^k \frac{x}{\log x} \ll 2^{-e^k} \frac{x}{\log x}.$$

Here the Mertens estimate (see, for example, Theorem 2.7 of Montgomery and Vaughan [20]) was used to obtain that the product over primes is $\ll 1$. Since $2^{-e^k} \frac{x}{\log x}$ is always larger than $\sqrt{x}$ on our range of k, the result follows. $\qquad\square$

We will need the following version of Parseval's identity for Dirichlet series.

HARMONIC ANALYSIS RESULT 1 (see (5.26) in Section 5.1 of Montgomery and Vaughan [20]). *Let $(a_n)_{n=1}^{\infty}$ be any sequence of complex numbers, let $A(s) := \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ denote the corresponding Dirichlet series and let σ_c denote its abscissa of convergence. Then for any $\sigma > \max\{0, \sigma_c\}$, we have*

$$\int_0^{\infty} \frac{|\sum_{n \leqslant x} a_n|^2}{x^{1+2\sigma}} \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left|\frac{A(\sigma + it)}{\sigma + it}\right|^2 \, dt.$$

2.2. Upper bounds: statement of the propositions. We will need a little notation. Given a random multiplicative function $f(n)$ (either Steinhaus or Rademacher, depending on the context) and an integer $0 \leqslant k \leqslant \log\log x$, let F_k denote the partial Euler product of $f(n)$ over $x^{e^{-(k+1)}}$-smooth numbers. Thus, for all complex s with $\Re(s) > 0$, we have

$$F_k(s) = \prod_{p \leqslant x^{e^{-(k+1)}}} \left(1 - \frac{f(p)}{p^s}\right)^{-1} = \sum_{\substack{n=1, \\ n \text{ is } x^{e^{-(k+1)}} \text{ smooth}}}^{\infty} \frac{f(n)}{n^s}$$

in the Steinhaus case, and

$$F_k(s) = \prod_{p \leqslant x^{e^{-(k+1)}}} \left(1 + \frac{f(p)}{p^s}\right) = \sum_{\substack{n=1, \\ n \text{ is } x^{e^{-(k+1)}} \text{ smooth}}}^{\infty} \frac{f(n)}{n^s}$$

in the Rademacher case (the product taking a different form because $f(n)$ is only supported on squarefree numbers in that case).

PROPOSITION 1. *Let $f(n)$ be a Steinhaus random multiplicative function, let x be large and set $\mathcal{K} := \lfloor \log \log \log x \rfloor$. Uniformly for all $2/3 \leqslant q \leqslant 1$, we have*

$$\left\| \sum_{n \leqslant x} f(n) \right\|_{2q} \ll \sqrt{\frac{x}{\log x}} \sum_{0 \leqslant k \leqslant \mathcal{K}} \left\| \int_{-1/2}^{1/2} \left| F_k\left(1/2 - \frac{k}{\log x} + it\right) \right|^2 dt \right\|_q^{1/2}$$
$$+ \sqrt{\frac{x}{\log x}}.$$

It is perhaps worth emphasizing that $\| \cdot \|_r := (\mathbb{E}| \cdot |^r)^{1/r}$ is a genuine norm when $r \geqslant 1$ but not for smaller r. Thus, $\| \cdot \|_{2q}$ is a genuine norm on our range of q (so we may apply Minkowski's inequality to it as we shall in the proof), but $\| \cdot \|_q$ is not and Minkowski's inequality is not applicable.

PROPOSITION 2. *Let $f(n)$ be a Rademacher random multiplicative function, let x be large and set $\mathcal{K} := \lfloor \log \log \log x \rfloor$. Uniformly for all $2/3 \leqslant q \leqslant 1$, we have*

$$\left\| \sum_{n \leqslant x} f(n) \right\|_{2q} \ll \sqrt{\frac{x}{\log x}} \sum_{0 \leqslant k \leqslant \mathcal{K}} \max_{N \in \mathbb{Z}} \frac{1}{(|N|+1)^{1/8}}$$
$$\times \left\| \int_{N-1/2}^{N+1/2} \left| F_k\left(1/2 - \frac{k}{\log x} + it\right) \right|^2 dt \right\|_q^{1/2}$$
$$+ \sqrt{\frac{x}{\log x}}.$$

In these bounds we expect, and it will turn out to be the case, that the main contribution comes from small k and small N. In the Steinhaus case, for any fixed $t \in \mathbb{R}$, the distribution of the sequence $(f(n)n^{it})$ is the same as the distribution of $(f(n))$, which is why one does not need to deal with translates by N in the Steinhaus case. This is one of a few differences between Rademacher and Steinhaus random multiplicative functions that will recur a number of times in our analysis.

2.3. Lower bounds: statement of the propositions. For our work on lower bounds, we again connect the size of $\|\sum_{n \leqslant x} f(n)\|_{2q}$ with a certain integral average and thence with random Euler products. Let F denote the partial Euler product of $f(n)$, either Steinhaus or Rademacher, over x-smooth numbers. (Thus, $F = F_{-1}$ if we slightly abuse our earlier notation).

PROPOSITION 3. *There exists a large absolute constant $C > 0$ such that the following is true. If $f(n)$ is a Steinhaus random multiplicative function and x is large, then uniformly for all $2/3 \leqslant q \leqslant 1$, we have*

$$\left\| \sum_{n \leqslant x} f(n) \right\|_{2q} \gg \sqrt{\frac{x}{\log x}} \left\| \int_{1}^{\sqrt{x}} \left| \sum_{m \leqslant z} f(m) \right|^2 \frac{dz}{z^2} \right\|_{q}^{1/2} - C \sqrt{\frac{x}{\log x}}.$$

In particular, for any large quantity V, we have that $\| \sum_{n \leqslant x} f(n) \|_{2q}$ is

$$\gg \sqrt{\frac{x}{\log x}} \left(\left\| \int_{-1/2}^{1/2} \left| F\left(1/2 + \frac{4V}{\log x} + it \right) \right|^2 dt \right\|_{q}^{1/2} \right.$$

$$\left. - \frac{C}{e^V} \left\| \int_{-1/2}^{1/2} \left| F\left(1/2 + \frac{2V}{\log x} + it \right) \right|^2 dt \right\|_{q}^{1/2} - C \right).$$

PROPOSITION 4. *If $f(n)$ is a Rademacher random multiplicative function, the first bound in Proposition 3 continues to hold and the second bound may be replaced by the statement that*

$$\left\| \sum_{n \leqslant x} f(n) \right\|_{2q}$$

$$\gg \sqrt{\frac{x}{\log x}} \left(\left\| \int_{-1/2}^{1/2} \left| F\left(1/2 + \frac{4V}{\log x} + it \right) \right|^2 dt \right\|_{q}^{1/2} \right.$$

$$\left. - \frac{C}{e^V} \max_{N \in \mathbb{Z}} \frac{1}{(|N|+1)^{1/8}} \left\| \int_{N-1/2}^{N+1/2} \left| F\left(1/2 + \frac{2V}{\log x} + it \right) \right|^2 dt \right\|_{q}^{1/2} - C \right).$$

When we come to apply these propositions, we will choose V to be a sufficiently large fixed constant that, because of the factor C/e^V, the second subtracted Euler product integral is negligible compared with the first.

2.4. Proof of Propositions 1 and 2. We begin with Proposition 1. Let $P(n)$ denote the largest prime factor of n and let $\Psi(x, y) := \#\{n \leqslant x : n \text{ is } y \text{ smooth}\} = \#\{n \leqslant x : P(n) \leqslant y\}$. Then by Minkowski's inequality, for all $2/3 \leqslant q \leqslant 1$, we have

$$\left\| \sum_{n \leqslant x} f(n) \right\|_{2q} \leqslant \sum_{0 \leqslant k \leqslant \mathcal{K}} \left\| \sum_{\substack{n \leqslant x, \\ x^{e^{-(k+1)}} < P(n) \leqslant x^{e^{-k}}}} f(n) \right\|_{2q} + \left\| \sum_{\substack{n \leqslant x, \\ P(n) \leqslant x^{e^{-(\mathcal{K}+1)}}}} f(n) \right\|_{2q}.$$

Furthermore, by Hölder's inequality and a mean square calculation, we have

$$\left\| \sum_{\substack{n \leqslant x, \\ P(n) \leqslant x^{e^{-(\mathcal{K}+1)}}}} f(n) \right\|_{2q} \leqslant \left\| \sum_{\substack{n \leqslant x, \\ P(n) \leqslant x^{e^{-(\mathcal{K}+1)}}}} f(n) \right\|_{2} = \Psi(x, x^{e^{-(\mathcal{K}+1)}})^{1/2},$$

and recalling that $\mathcal{K} := \lfloor \log \log \log x \rfloor$ as well as standard estimates for smooth numbers (see Theorem 7.6 of Montgomery and Vaughan [20], for example), the above is $\leqslant \Psi(x, x^{1/\log\log x})^{1/2} \ll \sqrt{x}(\log x)^{-c \log\log\log x}$. This contribution is more than acceptable.

If we let $\mathbb{E}^{(k)}$ denote expectation conditional on the values $(f(p))_{p \leqslant x^{e^{-(k+1)}}}$ and then use Hölder's inequality and a conditional mean square calculation, we see that $\sum_{0 \leqslant k \leqslant \mathcal{K}} \| \sum_{\substack{n \leqslant x, \\ x^{e^{-(k+1)}} < P(n) \leqslant x^{e^{-k}}}} f(n)\|_{2q}$ is

$$= \sum_{0 \leqslant k \leqslant \mathcal{K}} \left\| \sum_{\substack{1 < m \leqslant x, \\ p | m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} f(m) \sum_{\substack{n \leqslant x/m, \\ n \text{ is } x^{e^{-(k+1)}}\text{-smooth}}} f(n) \right\|_{2q}$$

$$= \sum_{0 \leqslant k \leqslant \mathcal{K}} \left(\mathbb{E}\mathbb{E}^{(k)} \left| \sum_{\substack{1 < m \leqslant x, \\ p | m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} f(m) \sum_{\substack{n \leqslant x/m, \\ n \text{ is } x^{e^{-(k+1)}}\text{-smooth}}} f(n) \right|^{2q} \right)^{1/2q}$$

$$\leqslant \sum_{0 \leqslant k \leqslant \mathcal{K}} \left(\mathbb{E}\left(\mathbb{E}^{(k)} \left| \sum_{\substack{1 < m \leqslant x, \\ p | m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} f(m) \sum_{\substack{n \leqslant x/m, \\ n \text{ is } x^{e^{-(k+1)}}\text{-smooth}}} f(n) \right|^{2} \right)^{q} \right)^{1/2q}$$

$$= \sum_{0 \leqslant k \leqslant \mathcal{K}} \left\| \sum_{\substack{1 < m \leqslant x, \\ p | m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} \left| \sum_{\substack{n \leqslant x/m, \\ n \text{ is } x^{e^{-(k+1)}}\text{-smooth}}} f(n) \right|^{2} \right\|_{q}^{1/2}.$$

To proceed further, we want to replace $|\sum_{\substack{n \leqslant x/m, \\ n \text{ is } x^{e^{-(k+1)}}\text{-smooth}}} f(n)|^2$ in the above by a smoothed version. Set $X = e^{\sqrt{\log x}}$, say, and note that (uniformly for any $2/3 \leqslant q \leqslant 1$ and any $0 \leqslant k \leqslant \mathcal{K}$) we have

$$
\mathbb{E}\left(\sum_{\substack{1 < m \leqslant x, \\ p|m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} \left| \sum_{\substack{n \leqslant x/m, \\ n \text{ is } x^{e^{-(k+1)}}\text{-smooth}}} f(n) \right|^2 \right)^q
$$

$$
\ll \mathbb{E}\left(\sum_{\substack{1 < m \leqslant x, \\ p|m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} \frac{X}{m} \int_m^{m(1+1/X)} \left| \sum_{\substack{n \leqslant x/t, \\ x^{e^{-(k+1)}}\text{-smooth}}} f(n) \right|^2 dt \right)^q +
$$

$$
+ \mathbb{E}\left(\sum_{\substack{1 < m \leqslant x, \\ p|m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} \frac{X}{m} \int_m^{m(1+1/X)} \left| \sum_{\substack{x/t < n \leqslant x/m, \\ x^{e^{-(k+1)}}\text{-smooth}}} f(n) \right|^2 dt \right)^q.
$$

Using Hölder's inequality and a mean square calculation, the second term is at most

$$
\left(\sum_{\substack{1 < m \leqslant x, \\ p|m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} \frac{X}{m} \int_m^{m(1+1/X)} \mathbb{E}\left| \sum_{\substack{x/t < n \leqslant x/m, \\ x^{e^{-(k+1)}}\text{-smooth}}} f(n) \right|^2 dt \right)^q
$$

$$
\leqslant \left(\sum_{\substack{1 < m \leqslant x, \\ p|m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} \left(1 + \frac{x}{mX} \right) \right)^q.
$$

In particular, using the second part of Number Theory Result 1, we obtain that the sum

$$
\sum_{\substack{1 < m \leqslant x, \\ p|m \Rightarrow x^{e^{-(k+1)}} < p \leqslant x^{e^{-k}}}} 1 \ll 2^{-e^k} \frac{x}{\log x},
$$

and we have

$$
\sum_{m \leqslant x} \frac{x}{mX} \ll \frac{x \log x}{X} \ll 2^{-e^k} \frac{x}{\log x}
$$

on our range $0 \leqslant k \leqslant \mathcal{K}$ as well. So taking $2q$th roots and summing over $0 \leqslant k \leqslant \mathcal{K}$ leads to an acceptable overall contribution $\ll \sqrt{x/\log x}$ in Proposition 1.

Meanwhile, by swapping the order of the sum and integral, we see the first term in the above is

$$\mathbb{E}\left(\int_{xe^{-(k+1)}}^{x}\left|\sum_{\substack{n\leqslant x/t,\\ xe^{-(k+1)}\text{-smooth}}}f(n)\right|^2\sum_{\substack{t/(1+1/X)\leqslant m\leqslant t,\\ p|m\Rightarrow xe^{-(k+1)}<p\leqslant xe^{-k}}}\frac{X}{m}\,dt\right)^q,$$

and a standard sieve estimate applied to the sum over m (sieving out by all primes in $[2,t^{1/10}]\setminus(xe^{-(k+1)},xe^{-k}]$, say; see, for example, Theorem 3.6 of Montgomery and Vaughan [20] for a suitable sieve result) shows that this has order at most

$$\mathbb{E}\left(\int_{xe^{-(k+1)}}^{x}\left|\sum_{\substack{n\leqslant x/t,\\ xe^{-(k+1)}\text{-smooth}}}f(n)\right|^2\frac{dt}{\log t}\right)^q$$

$$=x^q\mathbb{E}\left(\int_{1}^{x^{1-e^{-(k+1)}}}\left|\sum_{\substack{n\leqslant z,\\ xe^{-(k+1)}\text{-smooth}}}f(n)\right|^2\frac{dz}{z^2\log(\frac{x}{z})}\right)^q.$$

To obtain the second expression here, we made a substitution $z=x/t$ in the integral.

To obtain a satisfactory dependence on k in our final estimations, we now note that if $z\leqslant\sqrt{x}$, we have $\log(x/z)\gg\log x$, whereas if $\sqrt{x}<z\leqslant x^{1-e^{-(k+1)}}$, we have $\log(x/z)\gg e^{-k}\log x$. Thus, in any case, we have $\log(x/z)\gg z^{-2k/\log x}\log x$, so the above is

$$\ll\frac{x^q}{\log^q x}\mathbb{E}\left(\int_{1}^{x^{1-e^{-(k+1)}}}\left|\sum_{\substack{n\leqslant z,\\ xe^{-(k+1)}\text{-smooth}}}f(n)\right|^2\frac{dz}{z^{2-2k/\log x}}\right)^q.$$

Finally, recalling that F_k denotes the partial Euler product of $f(n)$ over $xe^{-(k+1)}$-smooth numbers and that $2/3\leqslant q\leqslant 1$, we can apply Harmonic Analysis Result 1 to deduce that the expectation above is

$$\leqslant\mathbb{E}\left(\int_{-\infty}^{\infty}\frac{|F_k(\frac{1}{2}-\frac{k}{\log x}+it)|^2}{|1/2-\frac{k}{\log x}+it|^2}\,dt\right)^q$$

$$\leqslant\sum_{n\in\mathbb{Z}}\mathbb{E}\left(\int_{n-1/2}^{n+1/2}\frac{|F_k(\frac{1}{2}-\frac{k}{\log x}+it)|^2}{|1/2-\frac{k}{\log x}+it|^2}\,dt\right)^q$$

$$\ll\sum_{n\in\mathbb{Z}}\frac{1}{|n|^{2q}+1}\mathbb{E}\left(\int_{n-1/2}^{n+1/2}\left|F_k\left(\frac{1}{2}-\frac{k}{\log x}+it\right)\right|^2\,dt\right)^q.$$

In the Steinhaus case, since the law of the random function $f(n)$ is the same as the law of $f(n)n^{it}$ for any fixed $t \in \mathbb{R}$, we have

$$\mathbb{E}\left(\int_{n-1/2}^{n+1/2}\left|F_k\left(1/2 - \frac{k}{\log x} + it\right)\right|^2 dt\right)^q$$

$$= \mathbb{E}\left(\int_{-1/2}^{1/2}\left|F_k\left(1/2 - \frac{k}{\log x} + it\right)\right|^2 dt\right)^q \quad \forall\, n.$$

Proposition 1 now follows on putting everything together. □

The proof of Proposition 2, covering the Rademacher case, is extremely similar to the Steinhaus case. Indeed, the only nontrivial change comes at the very end, where (since it is no longer the case that the law of the random function $f(n)$ is the same as the law of $f(n)n^{it}$) we apply the bound

$$\sum_{n \in \mathbb{Z}} \frac{1}{|n|^{2q} + 1} \mathbb{E}\left(\int_{n-1/2}^{n+1/2} |F_k(1/2 - \frac{k}{\log x} + it)|^2 dt\right)^q$$

$$\ll \max_{N \in \mathbb{Z}} \frac{1}{(|N| + 1)^{1/4}} \mathbb{E}\left(\int_{N-1/2}^{N+1/2} |F_k(1/2 - \frac{k}{\log x} + it)|^2 dt\right)^q.$$

Proposition 2 follows on inserting this into the argument and putting everything together and taking $2q$th roots. □

2.5. Proof of Propositions 3 and 4. Again we let $P(n)$ denote the largest prime factor of n, and we introduce an auxiliary Rademacher random variable ϵ (independent of everything else). Proceeding similarly as in Section 2.2 of Harper, Nikeghbali and Radziwiłł [13], we find that if $2/3 \leqslant q \leqslant 1$, then

$$\mathbb{E}\left|\sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n)\right|^{2q}$$

$$= \frac{1}{2^{2q}} \mathbb{E}\left|\sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) + \sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n) + \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) - \sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n)\right|^{2q}$$

$$\leqslant \mathbb{E}\left|\sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) + \sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n)\right|^{2q} + \mathbb{E}\left|\sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) - \sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n)\right|^{2q}$$

$$= 2\mathbb{E}\left|\epsilon \sum_{\substack{n\leqslant x,\\P(n)>\sqrt{x}}} f(n) + \sum_{\substack{n\leqslant x,\\P(n)\leqslant\sqrt{x}}} f(n)\right|^{2q} = 2\mathbb{E}\left|\sum_{n\leqslant x} f(n)\right|^{2q}$$

since the law of

$$\epsilon \sum_{\substack{n\leqslant x,\\P(n)>\sqrt{x}}} f(n) = \epsilon \sum_{\sqrt{x}<p\leqslant x} f(p) \sum_{m\leqslant x/p} f(m)$$

conditional on the values $(f(p))_{p\leqslant\sqrt{x}}$ is the same as the law of $\sum_{\substack{n\leqslant x,\\P(n)>\sqrt{x}}} f(n)$.
We can rewrite this as

$$\left\|\sum_{n\leqslant x} f(n)\right\|_{2q} \gg \left\|\sum_{\substack{n\leqslant x,\\P(n)>\sqrt{x}}} f(n)\right\|_{2q}.$$

Now in the decomposition $\sum_{\substack{n\leqslant x,\\P(n)>\sqrt{x}}} f(n) = \sum_{\sqrt{x}<p\leqslant x} f(p) \sum_{m\leqslant x/p} f(m)$,
the inner sums are determined by the values $(f(p))_{p\leqslant\sqrt{x}}$, which are independent of the outer random variables $(f(p))_{\sqrt{x}<p\leqslant x}$. So conditioning on the values $(f(p))_{p\leqslant\sqrt{x}}$ determining the inner sums and applying the lower bound part of Khintchine's inequality (see, for example, Lemma 3.8.1 of Gut [8] for the Rademacher case of this, the Steinhaus case may be proved similarly), it follows that

$$\mathbb{E}\left|\sum_{\substack{n\leqslant x,\\P(n)>\sqrt{x}}} f(n)\right|^{2q} \gg \mathbb{E}\left(\sum_{\sqrt{x}<p\leqslant x}\left|\sum_{m\leqslant x/p} f(m)\right|^{2}\right)^{q}$$

$$\geqslant \frac{1}{\log^q x}\mathbb{E}\left(\sum_{\sqrt{x}<p\leqslant x}\log p\left|\sum_{m\leqslant x/p} f(m)\right|^{2}\right)^{q}.$$

Next, there comes a smoothing step, similarly as in our work on upper bounds. Recall that we let $X = e^{\sqrt{\log x}}$. It is easy to check that we always have $|a+b|^2 \geqslant (1/4)|a|^2 - \min\{|b|^2, |a/2|^2\} \geqslant 0$, say, and therefore

$$\sum_{\sqrt{x}<p\leqslant x}\log p\left|\sum_{m\leqslant x/p} f(m)\right|^{2} = \sum_{\sqrt{x}<p\leqslant x}\log p\frac{X}{p}\int_{p}^{p(1+1/X)}\left|\sum_{m\leqslant x/p} f(m)\right|^{2} dt$$

$$\geqslant \frac{1}{4}\sum_{\sqrt{x}<p\leqslant x}\log p\frac{X}{p}\int_{p}^{p(1+1/X)}\left|\sum_{m\leqslant x/t} f(m)\right|^{2} dt$$

$$- \sum_{\sqrt{x} < p \leqslant x} \log p \frac{X}{p}$$

$$\times \int_{p}^{p(1+1/X)} \min \left\{ \left| \sum_{x/t < m \leqslant x/p} f(m) \right|^2, \frac{1}{4} \left| \sum_{m \leqslant x/t} f(m) \right|^2 \right\} dt.$$

Thus, when $2/3 \leqslant q \leqslant 1$, we get

$$\mathbb{E} \left| \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \right|^{2q}$$

$$\gg \frac{1}{\log^q x} \mathbb{E} \left(\frac{1}{4} \sum_{\sqrt{x} < p \leqslant x} \log p \frac{X}{p} \int_{p}^{p(1+1/X)} \left| \sum_{m \leqslant x/t} f(m) \right|^2 dt \right)^q$$

$$- \frac{1}{\log^q x} \mathbb{E} \left(\sum_{\sqrt{x} < p \leqslant x} \log p \frac{X}{p} \int_{p}^{p(1+1/X)} \left| \sum_{x/t < m \leqslant x/p} f(m) \right|^2 dt \right)^q.$$

To complete the proof of the first part of Proposition 3, we apply Hölder's inequality and obtain that

$$\mathbb{E} \left(\sum_{\sqrt{x} < p \leqslant x} \log p \frac{X}{p} \int_{p}^{p(1+1/X)} \left| \sum_{x/t < m \leqslant x/p} f(m) \right|^2 dt \right)^q$$

is

$$\leqslant \left(\sum_{\sqrt{x} < p \leqslant x} \log p \frac{X}{p} \int_{p}^{p(1+1/X)} \mathbb{E} \left| \sum_{x/t < m \leqslant x/p} f(m) \right|^2 dt \right)^q$$

$$\ll \left(\sum_{\sqrt{x} < p \leqslant x} \log p \left(\frac{x}{pX} + 1 \right) \right)^q \ll \left(\frac{x \log x}{X} + x \right)^q \ll x^q.$$

Here, the penultimate inequality again uses the standard Chebychev prime number estimates, namely

$$\sum_{p \leqslant x} \frac{\log p}{p} \ll \log x \quad \text{and} \quad \sum_{p \leqslant x} \log p \ll x.$$

We also find that the main term

$$
\mathbb{E}\left(\sum_{\sqrt{x}<p\leqslant x}\log p\,\frac{X}{p}\int_{p}^{p(1+1/X)}\left|\sum_{m\leqslant x/t}f(m)\right|^{2}dt\right)^{q}
$$

is

$$
=\mathbb{E}\left(\int_{\sqrt{x}}^{x}\sum_{t/(1+1/X)<p\leqslant t}\log p\,\frac{X}{p}\left|\sum_{m\leqslant x/t}f(m)\right|^{2}dt\right)^{q}
$$

$$
\gg\mathbb{E}\left(\int_{\sqrt{x}}^{x}\left|\sum_{m\leqslant x/t}f(m)\right|^{2}dt\right)^{q}=x^{q}\mathbb{E}\left(\int_{1}^{\sqrt{x}}\left|\sum_{m\leqslant z}f(m)\right|^{2}\frac{dz}{z^{2}}\right)^{q}.
$$

The first part of Proposition 3 follows on putting all our calculations together.

To deduce the second part of Proposition 3, we simply note that for any large V and any $2/3\leqslant q\leqslant 1$, we have

$$
\mathbb{E}\left(\int_{1}^{\sqrt{x}}\left|\sum_{m\leqslant z}f(m)\right|^{2}\frac{dz}{z^{2}}\right)^{q}\geqslant\mathbb{E}\left(\int_{1}^{\sqrt{x}}\left|\sum_{\substack{m\leqslant z,\\x\text{-smooth}}}f(m)\right|^{2}\frac{dz}{z^{2+8V/\log x}}\right)^{q}
$$

$$
\geqslant\mathbb{E}\left(\int_{1}^{\infty}\left|\sum_{\substack{m\leqslant z,\\x\text{-smooth}}}f(m)\right|^{2}\frac{dz}{z^{2+8V/\log x}}\right)^{q}
$$

$$
-\mathbb{E}\left(\int_{\sqrt{x}}^{\infty}\left|\sum_{\substack{m\leqslant z,\\x\text{-smooth}}}f(m)\right|^{2}\frac{dz}{z^{2+8V/\log x}}\right)^{q}
$$

$$
\geqslant\mathbb{E}\left(\int_{1}^{\infty}\left|\sum_{\substack{m\leqslant z,\\x\text{-smooth}}}f(m)\right|^{2}\frac{dz}{z^{2+8V/\log x}}\right)^{q}
$$

$$
-\frac{1}{e^{2Vq}}\mathbb{E}\left(\int_{1}^{\infty}\left|\sum_{\substack{m\leqslant z,\\x\text{-smooth}}}f(m)\right|^{2}\frac{dz}{z^{2+4V/\log x}}\right)^{q}.
$$

By Harmonic Analysis Result 1, the first term here is

$$\gg \mathbb{E}\left(\int_{-1/2}^{1/2}\left|F\left(1/2+\frac{4V}{\log x}+it\right)\right|^2 dt\right)^q$$

and the subtracted second term is

$$\ll e^{-2Vq}\mathbb{E}\left(\int_{-\infty}^{\infty}\frac{|F(1/2+\frac{2V}{\log x}+it)|^2}{|1/2+\frac{2V}{\log x}+it|^2}\,dt\right)^q,$$

which in the Steinhaus case is

$$\ll e^{-2Vq}\mathbb{E}\left(\int_{-1/2}^{1/2}\left|F\left(1/2+\frac{2V}{\log x}+it\right)\right|^2 dt\right)^q$$

by 'translation invariance in law' (as at the end of the proof of Proposition 1). Putting everything together, this finishes the proof of Proposition 3. □

The arguments in the Rademacher case are, once again, exactly the same until the final line, where we do not have 'translation invariance', so we must upper bound

$$\mathbb{E}\left(\int_{-\infty}^{\infty}\frac{|F(1/2+\frac{2V}{\log x}+it)|^2}{|1/2+\frac{2V}{\log x}+it|^2}\,dt\right)^q$$

by

$$\max_{N\in\mathbb{Z}}\frac{1}{(|N|+1)^{1/4}}\mathbb{E}\left(\int_{N-1/2}^{N+1/2}\left|F\left(1/2+\frac{2V}{\log x}+it\right)\right|^2 dt\right)^q. \qquad \square$$

3. Probabilistic calculations

In this section, we collect together various probabilistic calculations that we wish to isolate in advance of our main proofs. These are of two basic kinds: first, the estimation of the mean square, and related quantities, for some random Euler

products; and second, the estimation of the probability of certain events where the measure is weighted by the mean square of random Euler products. The latter calculations provide analogues of Girsanov's theorem in our setting, where the logarithms of our random products are not actually Gaussian random variables but only approximately so.

3.1. The mean square of random Euler products.

LEMMA 1. *If f is a Steinhaus random multiplicative function, then for any real t and u, any real $400(1 + u^2) \leqslant x \leqslant y$ and any real $\sigma \geqslant -1/\log y$, we have*

$$
\mathbb{E} \prod_{x < p \leqslant y} \left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-2} \left| 1 - \frac{f(p)}{p^{1/2+\sigma+it}} \right|^{-iu}
$$
$$
= \exp \left\{ \sum_{x < p \leqslant y} \frac{1 + iu \cos(t \log p) - u^2/4}{p^{1+2\sigma}} + T(u) \right\},
$$

where $T(u) = T_{x,y,\sigma,t}(u)$ satisfies $|T(u)| \ll \frac{1+|u|^3}{\sqrt{x} \log x}$, and its derivative satisfies $|T'(u)| \ll \frac{1}{\sqrt{x} \log x}$ when $|u| \leqslant 1$.

Proof of Lemma 1. To simplify the writing of the proof, we temporarily set $R_p(t) := -\Re \log(1 - \frac{f(p)}{p^{1/2+\sigma+it}})$. Thus, we may rewrite

$$
\left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-2} \left| 1 - \frac{f(p)}{p^{1/2+\sigma+it}} \right|^{-iu} = \exp\{2R_p(0) + iuR_p(t)\}
$$
$$
= 1 + \sum_{j=1}^{\infty} \frac{(2R_p(0) + iuR_p(t))^j}{j!}.
$$

Now using the Taylor expansion of the logarithm, we have

$$
R_p(t) = \sum_{k=1}^{\infty} \frac{\Re(f(p)p^{-it})^k}{kp^{k(1/2+\sigma)}} = \frac{\Re f(p)p^{-it}}{p^{1/2+\sigma}} + O\left(\frac{1}{p^{1+2\sigma}} \right).
$$

In particular, by symmetry, we have $\mathbb{E}\Re(f(p)p^{-it})^k = \mathbb{E}\Re f(p)^k = 0$ for all $k \geqslant 1$, and, therefore, $\mathbb{E}R_p(t) = 0$. A simple trigonometric calculation

also shows that $\mathbb{E}(\Re f(p))^2 = 1/2$ and, similarly, $\mathbb{E}\Re f(p)\Re f(p)p^{-it} = \cos(t \log p)/2$. So we have

$$\mathbb{E}R_p(t)^2 = \mathbb{E}\frac{(\Re f(p)p^{-it})^2}{p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right) = \frac{1}{2p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right)$$

as well as

$$\mathbb{E}R_p(0)R_p(t) = \mathbb{E}\frac{(\Re f(p))(\Re f(p)p^{-it})}{p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right)$$
$$= \frac{\cos(t \log p)}{2p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right).$$

For $j \geqslant 3$, we can use the trivial bound

$$|R_p(t)^j| \leqslant \left(\sum_{k=1}^{\infty}\frac{1}{p^{k(1/2+\sigma)}}\right)^j = \frac{1}{(p^{1/2+\sigma}-1)^j}.$$

Next, we note that for primes $y \geqslant p > x \geqslant 400(1+u^2)$ as in our product, and for $\sigma \geqslant -\frac{1}{\log y}$, we have $\frac{1}{p^{1/2+\sigma}} = \frac{e^{-\sigma \log p}}{p^{1/2}} \leqslant \frac{e}{p^{1/2}}$, and, therefore, $(2+|u|)/p^{1/2+\sigma} \leqslant e/5$. So putting things together, for such primes, we have

$$\mathbb{E}\left|1 - \frac{f(p)}{p^{1/2+\sigma}}\right|^{-2}\left|1 - \frac{f(p)}{p^{1/2+\sigma+it}}\right|^{-iu}$$
$$= 1 + \frac{(4\mathbb{E}R_p(0)^2 + 4iu\mathbb{E}R_p(0)R_p(t) - u^2\mathbb{E}R_p(t)^2)}{2}$$
$$\quad + \mathbb{E}\sum_{j=3}^{\infty}\frac{(2R_p(0)+iuR_p(t))^j}{j!}$$
$$= 1 + \frac{(1+iu\cos(t \log p)-u^2/4)}{p^{1+2\sigma}} + O\left(\sum_{j=3}^{\infty}\frac{(2+|u|)^j}{j!(p^{1/2+\sigma}-1)^j}\right)$$
$$= 1 + \frac{(1+iu\cos(t \log p)-u^2/4)}{p^{1+2\sigma}} + D_p(u),$$

where $D_p(u)$ satisfies $|D_p(u)| \ll \frac{1+|u|^3}{p^{3/2+3\sigma}} \ll \frac{1+|u|^3}{p^{3/2}}$ and its derivative satisfies $|D_p'(u)| \ll \frac{1}{p^{3/2+3\sigma}} \ll \frac{1}{p^{3/2}}$ for $|u| \leqslant 1$.

Finally, note that we may rewrite our conclusions as

$$\mathbb{E}\left|1 - \frac{f(p)}{p^{1/2+\sigma}}\right|^{-2}\left|1 - \frac{f(p)}{p^{1/2+\sigma+it}}\right|^{-iu}$$

$$= 1 + \frac{(1 + iu\cos(t\log p) - u^2/4)}{p^{1+2\sigma}} + D_p(u)$$

$$= \exp\left\{\frac{(1 + iu\cos(t\log p) - u^2/4)}{p^{1+2\sigma}} + T_p(u)\right\},$$

where $T_p(u)$ again satisfies $|T_p(u)| \ll \frac{1+|u|^3}{p^{3/2}}$ and also $|T_p'(u)| \ll \frac{1}{p^{3/2}}$ for $|u| \leqslant 1$. Since f is independent on distinct primes, we then deduce

$$\mathbb{E}\prod_{x < p \leqslant y}\left|1 - \frac{f(p)}{p^{1/2+\sigma}}\right|^{-2}\left|1 - \frac{f(p)}{p^{1/2+\sigma+it}}\right|^{-iu}$$

$$= \exp\left\{\sum_{x < p \leqslant y}\frac{1 + iu\cos(t\log p) - u^2/4}{p^{1+2\sigma}} + \sum_{x < p \leqslant y}T_p(u)\right\},$$

which implies Lemma 1 in view of the standard Chebychev-type estimate $\sum_{p > x} 1/p^{3/2} \ll 1/(\sqrt{x}\log x)$. □

We note in particular that, by Lemma 1, for any $400 \leqslant x \leqslant y$ and any $\sigma \geqslant -1/\log y$ and for Steinhaus random multiplicative f, we have

$$\mathbb{E}\prod_{x < p \leqslant y}\left|1 - \frac{f(p)}{p^{1/2+\sigma}}\right|^{-2} = \exp\left\{\sum_{x < p \leqslant y}\frac{1}{p^{1+2\sigma}} + O\left(\frac{1}{\sqrt{x}\log x}\right)\right\}. \quad (2)$$

Moreover, under the same conditions, the same proof shows that

$$\mathbb{E}\prod_{x < p \leqslant y}\left|1 - \frac{f(p)}{p^{1/2+\sigma}}\right|^2$$

satisfies this estimate too (note the change of the exponent from -2 to 2).

We will need an analogue of Lemma 1 for the Rademacher case. Both the formulation and the proof of this are slightly more complicated. This is because the distribution of a Rademacher random multiplicative function is not invariant under shifts by n^{it} and also because $f(p)^2 \equiv 1$ in the Rademacher case and the Euler products we deal with here have slightly different forms.

LEMMA 2. *If f is a Rademacher random multiplicative function, then for any real t_1, t_2 and u, any real $400(1 + u^2) \leqslant x \leqslant y$ and any real $\sigma \geqslant -1/\log y$, we have*

$$\mathbb{E} \prod_{x < p \leqslant y} \left| 1 + \frac{f(p)}{p^{1/2+\sigma+it_1}} \right|^2 \left| 1 + \frac{f(p)}{p^{1/2+\sigma+i(t_1+t_2)}} \right|^{iu}$$

$$= \exp \left\{ \sum_{x < p \leqslant y} \frac{1 + iuc(t_1, t_2, p) - (u^2/4)(1 + \cos(2(t_1 + t_2) \log p))}{p^{1+2\sigma}} \right.$$

$$\left. + T(u) \right\}.$$

Here, $c(t_1, t_2, p) = 2 \cos(t_1 \log p) \cos((t_1 + t_2) \log p) - (1/2) \cos(2(t_1 + t_2) \log p)$, and $T(u) = T_{x,y,\sigma,t_1,t_2}(u)$ satisfies $|T(u)| \ll \frac{1+|u|^3}{\sqrt{x} \log x}$, and $|T'(u)| \ll \frac{1}{\sqrt{x} \log x}$ when $|u| \leqslant 1$.

Proof of Lemma 2. Let us temporarily set $R_p(t) := \Re \log(1 + \frac{f(p)}{p^{1/2+\sigma+it}})$, so we may rewrite

$$\left| 1 + \frac{f(p)}{p^{1/2+\sigma+it_1}} \right|^2 \left| 1 + \frac{f(p)}{p^{1/2+\sigma+i(t_1+t_2)}} \right|^{iu} = \exp\{2R_p(t_1) + iuR_p(t_1 + t_2)\}$$

$$= 1 + \sum_{j=1}^{\infty} \frac{(2R_p(t_1) + iuR_p(t_1 + t_2))^j}{j!}.$$

Using Taylor expansion (and the fact that $f(p) \in \{\pm 1\}$), we obtain that

$$R_p(t) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{\Re(f(p)p^{-it})^k}{kp^{k(1/2+\sigma)}}$$

$$= \frac{f(p) \cos(t \log p)}{p^{1/2+\sigma}} - \frac{\cos(2t \log p)}{2p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right).$$

In particular, this implies that

$$\mathbb{E}R_p(t) = -\frac{\cos(2t \log p)}{2p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right).$$

We also have

$$\mathbb{E}R_p(t)^2 = \frac{\cos^2(t \log p)}{p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right) = \frac{(1 + \cos(2t \log p))}{2p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right)$$

as well as

$$\mathbb{E} R_p(t_1) R_p(t_1 + t_2) = \frac{\cos(t_1 \log p) \cos((t_1 + t_2) \log p)}{p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right).$$

For $j \geqslant 3$, we again have $|R_p(t)^j| \leqslant (\sum_{k=1}^{\infty} \frac{1}{p^{k(1/2+\sigma)}})^j = \frac{1}{(p^{1/2+\sigma}-1)^j}$.

As in Lemma 1, for primes $y \geqslant p > x \geqslant 400(1 + u^2)$, we have $(2 + |u|)/p^{1/2+\sigma} \leqslant e/5$, so for such primes, we get

$$\mathbb{E}\left|1 + \frac{f(p)}{p^{1/2+\sigma+it_1}}\right|^2 \left|1 + \frac{f(p)}{p^{1/2+\sigma+i(t_1+t_2)}}\right|^{iu}$$

$$= 1 - \frac{\cos(2t_1 \log p)}{p^{1+2\sigma}} - \frac{(iu/2)\cos(2(t_1 + t_2)\log p)}{p^{1+2\sigma}} + O\left(\frac{1 + |u|}{p^{3/2+3\sigma}}\right)$$

$$+ \frac{(1 + \cos(2t_1 \log p)) + 2iu \cos(t_1 \log p)\cos((t_1 + t_2)\log p)}{p^{1+2\sigma}}$$

$$- \frac{\frac{u^2}{4}(1 + \cos(2(t_1 + t_2)\log p))}{p^{1+2\sigma}} + O\left(\frac{1 + u^2}{p^{3/2+3\sigma}}\right)$$

$$+ \mathbb{E}\sum_{j=3}^{\infty} \frac{(2R_p(t_1) + iu R_p(t_1 + t_2))^j}{j!}$$

$$= 1 + \frac{1 + iuc(t_1, t_2, p) - (u^2/4)(1 + \cos(2(t_1 + t_2)\log p))}{p^{1+2\sigma}} + D_p(u),$$

where $D_p(u)$ satisfies $|D_p(u)| \ll \frac{1+|u|^3}{p^{3/2+3\sigma}} \ll \frac{1+|u|^3}{p^{3/2}}$ and its derivative satisfies $|D_p'(u)| \ll \frac{1}{p^{3/2+3\sigma}} \ll \frac{1}{p^{3/2}}$ for $|u| \leqslant 1$.

From this point, the proof concludes exactly as the proof of Lemma 1.   □

Again, Lemma 2 implies that for any $400 \leqslant x \leqslant y$ and $\sigma \geqslant -1/\log y$ and real t_1, and for Rademacher random multiplicative f, we have

$$\mathbb{E} \prod_{x < p \leqslant y} \left|1 + \frac{f(p)}{p^{1/2+\sigma+it_1}}\right|^2 = \exp\left\{\sum_{x < p \leqslant y} \frac{1}{p^{1+2\sigma}} + O\left(\frac{1}{\sqrt{x}\log x}\right)\right\}. \tag{3}$$

Under these conditions, we can use the same proof to estimate

$$\mathbb{E} \prod_{x < p \leqslant y} \left|1 + \frac{f(p)}{p^{1/2+\sigma+it_1}}\right|^{-2}$$

(note the change of the exponent from 2 to -2), and in the Rademacher case, this slightly changes the answer: we find

$$\mathbb{E} \prod_{x < p \leqslant y} \left| 1 + \frac{f(p)}{p^{1/2+\sigma+it_1}} \right|^{-2}$$

$$= \exp \left\{ \sum_{x < p \leqslant y} \frac{1 + 2\cos(2t_1 \log p)}{p^{1+2\sigma}} + O\left(\frac{1}{\sqrt{x} \log x} \right) \right\}. \quad (4)$$

3.2. Girsanov-type calculations in the Steinhaus case.

Let $f(n)$ be a Steinhaus random multiplicative function and let x be large and $-1/100 \leqslant \sigma \leqslant 1/100$, say. Later we will impose some further restrictions on σ. Let us introduce a new 'tilted' probability measure $\tilde{\mathbb{P}} = \tilde{\mathbb{P}}_{x,\sigma}$ by setting

$$\tilde{\mathbb{P}}(A) := \frac{\mathbb{E} \mathbf{1}_A \prod_{p \leqslant x^{1/e}} \left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-2}}{\mathbb{E} \prod_{p \leqslant x^{1/e}} \left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-2}}$$

for each event A, where $\mathbf{1}$ denotes the indicator function. We will also sometimes write $\tilde{\mathbb{E}}$ to denote expectation (that is, integration) with respect to the measure $\tilde{\mathbb{P}}$. Thus, if G is some random variable, we have

$$\tilde{\mathbb{E}} G = \frac{\mathbb{E} G \prod_{p \leqslant x^{1/e}} \left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-2}}{\mathbb{E} \prod_{p \leqslant x^{1/e}} \left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-2}}.$$

The exact choice of the range of p in the definition of $\tilde{\mathbb{P}}$ is not too important since the independence of the $f(p)$ means that if the event A does not involve a particular prime, the expectation of that term will factor out and cancel between the numerator and denominator.

Furthermore, for each $l \in \mathbb{N} \cup \{0\}$, set

$$I_l(s) := \prod_{x^{e^{-(l+2)}} < p \leqslant x^{e^{-(l+1)}}} \left(1 - \frac{f(p)}{p^s} \right)^{-1},$$

the lth 'increment' of the Euler product corresponding to Steinhaus f.

Our goal here is to estimate $\tilde{\mathbb{P}}(A)$ for certain events A corresponding to restrictions on partial Euler products. We wish to show that these probabilities are essentially the same as they would be if the increments $\log |I_l(s)|$ were Gaussian. In that case, the very useful Girsanov theorem shows that the tilted probabilities are themselves Gaussian probabilities for certain shifted

Gaussians and so can be understood and estimated fairly straightforwardly. We will achieve all this by combining Lemma 1 (which corresponds to a characteristic function calculation for the tilted measure $\tilde{\mathbb{P}}$) with the Berry–Esseen theorem on distributional approximation and then develop the necessary Gaussian estimates. Since we must apply the Berry–Esseen theorem multiple times and it only supplies an absolute rather than a relative error, we must be careful to control the sizes of things to make this all work.

Let $(l_j)_{j=1}^n$ denote a *strictly decreasing* sequence of nonnegative integers, with $l_1 \leqslant \log\log x - 2$, and define a corresponding increasing sequence of real numbers $(x_j)_{j=1}^n$ by setting $x_j := x^{e^{-(l_j+1)}}$.

LEMMA 3. *Let the situation be as above and suppose that x_1 is sufficiently large (that is, larger than a suitable absolute constant) and that $|\sigma| \leqslant 1/\log x_n$. Suppose further that $(v_j)_{j=1}^n$ is any sequence of real numbers satisfying*

$$|v_j| \leqslant (1/40)\sqrt{\log x_j} + 2 \quad \forall 1 \leqslant j \leqslant n$$

and $(t_j)_{j=1}^n$ is any sequence of real numbers.
  Then we have

$$\tilde{\mathbb{P}}(v_j \leqslant \log|I_{l_j}(1/2 + \sigma + it_j)| \leqslant v_j + 1/j^2 \; \forall 1 \leqslant j \leqslant n)$$
$$= \left(1 + O\left(\frac{1}{x_1^{1/100}}\right)\right)\mathbb{P}(v_j \leqslant N_j \leqslant v_j + 1/j^2 \; \forall 1 \leqslant j \leqslant n),$$

where N_j are independent Gaussian random variables with mean

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t_j \log p)}{p^{1+2\sigma}}$$

and variance

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}}.$$

Proof of Lemma 3. By independence, both the probability on the left and the one on the right factor as a product over j, so it will suffice to prove that

$$\tilde{\mathbb{P}}(v_j \leqslant \log|I_{l_j}(1/2 + \sigma + it_j)| \leqslant v_j + 1/j^2)$$
$$= \left(1 + O\left(\frac{1}{x_j^{1/100}}\right)\right)\mathbb{P}(v_j \leqslant N_j \leqslant v_j + 1/j^2) \quad \forall j \leqslant n.$$

We also note at the outset that, because of our assumption that $|\sigma| \leqslant \frac{1}{\log x_n} \leqslant \frac{1}{\log x_j}$ as well as the standard Mertens estimate for sums over primes (see, for example, Theorem 2.7 of Montgomery and Vaughan [20]), we have

$$e^{-2}\left(1 + O\left(\frac{1}{\log x_j}\right)\right) \leqslant \sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{p^{1+2\sigma}} \leqslant e^2 \sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{p}$$

$$= e^2\left(1 + O\left(\frac{1}{\log x_j}\right)\right) \quad \forall 1 \leqslant j \leqslant n.$$

For Gaussian random variables, it is a standard calculation that the characteristic function

$$\mathbb{E}e^{iuN_j} = \mathbb{E}\exp\left\{iu \sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t_j \log p)}{p^{1+2\sigma}} + iu\sqrt{\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}}} N(0,1)\right\}$$

$$= \exp\left\{\sum_{x_j^{1/e} < p \leqslant x_j} \frac{iu\cos(t_j \log p) - u^2/4}{p^{1+2\sigma}}\right\}.$$

Now for any $|u| \leqslant x_j^{1/20}$, Lemma 1 (applied with x, y replaced by $x_j^{1/e}$ and x_j) implies that

$$\tilde{\mathbb{E}}e^{iu\log|I_{l_j}(1/2+\sigma+it_j)|} = \exp\left\{\sum_{x_j^{1/e} < p \leqslant x_j} \frac{iu\cos(t_j \log p) - u^2/4}{p^{1+2\sigma}} + T(u) - T(0)\right\}.$$

Combining this with the above calculation and with the estimates for $T(u), T'(u)$ in Lemma 1, we get

$$\left|\tilde{\mathbb{E}}e^{iu\log|I_{l_j}(1/2+\sigma+it_j)|} - \mathbb{E}e^{iuN_j}\right| = \exp\left\{-\frac{u^2}{4}\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{p^{1+2\sigma}}\right\}\left|e^{T(u)-T(0)} - 1\right|$$

$$\ll \exp\left\{-\frac{u^2}{4}\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{p^{1+2\sigma}}\right\}\frac{|u| + |u|^3}{\sqrt{x_j^{1/e}}\log x_j}.$$

Then by the Berry–Esseen theorem (see, for example, Lemma 7.6.1 of Gut [8]), we have

$$\left|\tilde{\mathbb{P}}(v_j \leqslant \log|I_{l_j}(1/2 + \sigma + it_j)| \leqslant v_j + 1/j^2) - \mathbb{P}(v_j \leqslant N_j \leqslant v_j + 1/j^2)\right|$$

$$\ll \int_{-x_j^{1/20}}^{x_j^{1/20}} \left| \frac{\tilde{\mathbb{E}} e^{iu \log |I_{l_j}(1/2+\sigma+it_j)|} - \mathbb{E} e^{iuN_j}}{u} \right| du + \frac{1}{x_j^{1/20}}$$

$$\ll \int_{-x_j^{1/20}}^{x_j^{1/20}} \frac{1+u^2}{\sqrt{x_j^{1/e} \log x_j}} \, du + \frac{1}{x_j^{1/20}} \ll \frac{1}{x_j^{1/60}}.$$

Finally, since the Gaussian N_j has mean $O(1)$ and (as we observed at the beginning) variance

$$\frac{1}{2} \sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{p^{1+2\sigma}} \geqslant \frac{e^{-2} + O(1/\log x_j)}{2}$$

and since $|v_j| \leqslant (1/40)\sqrt{\log x_j} + 2$, we have

$$\mathbb{P}(v_j \leqslant N_j \leqslant v_j + 1/j^2) \gg (1/j^2) e^{-(|v_j|+O(1))^2/(e^{-2}+O(1/\log x_j))}$$

$$\gg \frac{1}{j^2 x_j^{9/1600}} \gg \frac{1}{x_j^{1/150}}.$$

So we may rewrite the error $1/x_j^{1/60}$ from the Berry–Esseen theorem as $O(\frac{\mathbb{P}(v_j \leqslant N_j \leqslant v_j+1/j^2)}{x_j^{1/60-1/150}})$, which is $O(\frac{\mathbb{P}(v_j \leqslant N_j \leqslant v_j+1/j^2)}{x_j^{1/100}})$. $\qquad\square$

Using Lemma 3 and a suitable slicing argument, we can obtain a similar approximation result for the probabilities of slightly more complicated events.

LEMMA 4. *Suppose $(u_j)_{j=1}^n$ and $(v_j)_{j=1}^n$ are sequences of real numbers satisfying*

$$-(1/80)\sqrt{\log x_j} \leqslant u_j \leqslant v_j \leqslant (1/80)\sqrt{\log x_j} \quad \forall 1 \leqslant j \leqslant n,$$

and otherwise let the situation be as in Lemma 3. Then we have

$$\mathbb{P}\left(u_j + 2 \leqslant \sum_{m=1}^j N_m \leqslant v_j - 2 \, \forall j \leqslant n \right)$$

$$\ll \tilde{\mathbb{P}}\left(u_j \leqslant \sum_{m=1}^j \log \left| I_{l_m}\left(\frac{1}{2}+\sigma+it_m\right) \right| \leqslant v_j \, \forall j \leqslant n \right)$$

$$\ll \mathbb{P}\left(u_j - 2 \leqslant \sum_{m=1}^j N_m \leqslant v_j + 2 \, \forall j \leqslant n \right),$$

where the Gaussian random variables N_m are also as in Lemma 3.

In addition, if the numbers $(t_j)_{j=1}^n$ satisfy $|t_j| \leqslant \frac{1}{j^{2/3} \log x_j}$, then we have

$$\mathbb{P}\left((u_j - j) + O(1) \leqslant \sum_{m=1}^{j} G_m \leqslant (v_j - j) - O(1) \; \forall j \leqslant n\right)$$

$$\ll \tilde{\mathbb{P}}\left(u_j \leqslant \sum_{m=1}^{j} \log |I_{l_m}(1/2 + \sigma + it_m)| \leqslant v_j \; \forall j \leqslant n\right)$$

$$\ll \mathbb{P}\left((u_j - j) - O(1) \leqslant \sum_{m=1}^{j} G_m \leqslant (v_j - j) + O(1) \; \forall j \leqslant n\right),$$

where G_m are independent Gaussian random variables, each having mean 0 and variance $\sum_{x_m^{1/e} < p \leqslant x_m} \frac{1}{2p^{1+2\sigma}}$.

In the second part of Lemma 4, the assumption that $|t_m| \leqslant \frac{1}{m^{2/3} \log x_m}$ for all m together with our standing assumption $|\sigma| \leqslant \frac{1}{\log x_n} \leqslant \frac{1}{\log x_m}$ implies that

$$\mathbb{E} N_m = \sum_{x_m^{1/e} < p \leqslant x_m} \frac{\cos(t_m \log p)}{p^{1+2\sigma}} \approx \sum_{x_m^{1/e} < p \leqslant x_m} \frac{1}{p} \approx 1.$$

Subtracting these means for each m yields the mean zero random variables G_m and produces the subtracted term $-j$ in the upper and lower bounds of our events. As we shall see, this subtracted term in the upper bound (coming from $\mathbb{E} N_m$, which ultimately comes from the 'tilting' of our measures) will be at the heart of everything. It can reduce a sequence of fairly large terms $v_j \approx j$ that one would expect to impose a very little constraint on our partial sums to a smaller sequence that does impose a nontrivial constraint.

Proof of Lemma 4. As remarked previously, to prove the first part (involving the N_m), we shall simply approximate the probability we are interested in by a sum of probabilities of the form treated in Lemma 3.

More precisely, note that if we have

$$-(1/80)\sqrt{\log x_j} \leqslant \sum_{m=1}^{j} \log |I_{l_m}(1/2 + \sigma + it_m)| \leqslant (1/80)\sqrt{\log x_j}$$

and the analogous bounds for the sum up to $j - 1$, then we must have

$$|\log |I_{l_j}(1/2 + \sigma + it_j)\| \leqslant (1/80)\sqrt{\log x_{j-1}} + (1/80)\sqrt{\log x_j} \leqslant (1/40)\sqrt{\log x_j}.$$

So if we let $\mathcal{R}_1 := \{r \in \mathbb{Z} : |r| \leqslant (1/40)\sqrt{\log x_1} + 2\}$ (say) and, more generally, let $\mathcal{R}_j := \{r \in (1/j^2)\mathbb{Z} : |r| \leqslant (1/40)\sqrt{\log x_j} + 2\}$, then in order to have $u_j \leqslant \sum_{m=1}^{j} \log |I_{l_m}(1/2 + \sigma + it_m)| \leqslant v_j$ for all $1 \leqslant j \leqslant n$, we must have

$$r_j \leqslant \log |I_{l_j}(1/2 + \sigma + it_j)| \leqslant r_j + 1/j^2 \quad \forall 1 \leqslant j \leqslant n,$$

for some $r_1 \in \mathcal{R}_1, \ldots, r_n \in \mathcal{R}_n$ satisfying

$$u_j - \sum_{m=1}^{j} 1/m^2 \leqslant \sum_{m=1}^{j} r_m \leqslant v_j \quad \text{for all } 1 \leqslant j \leqslant n.$$

Thus, summing over all possible choices of the r_j and applying Lemma 3, we obtain

$$\tilde{\mathbb{P}}(u_j \leqslant \sum_{m=1}^{j} \log |I_{l_m}(1/2 + \sigma + it_m)| \leqslant v_j \ \forall 1 \leqslant j \leqslant n)$$

$$\ll \sum_{\substack{r_1 \in \mathcal{R}_1, \ldots, r_n \in \mathcal{R}_n, \\ u_j - \sum_{m=1}^{j} 1/m^2 \leqslant \sum_{m=1}^{j} r_m \leqslant v_j \quad \forall 1 \leqslant j \leqslant n}} \mathbb{P}(r_j \leqslant N_j \leqslant r_j + 1/j^2 \ \forall 1 \leqslant j \leqslant n)$$

$$\leqslant \mathbb{P}\left(u_j - \sum_{m=1}^{j} \frac{1}{m^2} \leqslant \sum_{m=1}^{j} N_m \leqslant v_j + \sum_{m=1}^{j} \frac{1}{m^2} \ \forall j\right)$$

$$\leqslant \mathbb{P}\left(u_j - 2 \leqslant \sum_{m=1}^{j} N_m \leqslant v_j + 2 \ \forall j \leqslant n\right).$$

This is the desired upper bound, and the proof of the corresponding lower bound is exactly similar.

To deduce the second part of the lemma, we will use the facts that $x_m \geqslant x_1^{e^{m-1}}$ and, more generally, $x_n \geqslant x_m^{e^{n-m}}$, both of which follow from the definition of the sequence $(x_j)_{j=1}^{n}$. The second part of the lemma follows by setting

$$G_m = N_m - \mathbb{E}N_m = N_m - \sum_{x_m^{1/e} < p \leqslant x_m} \frac{\cos(t_m \log p)}{p^{1+2\sigma}} \quad \text{for all } m,$$

and noting that

$$\sum_{x_m^{1/e} < p \leqslant x_m} \frac{\cos(t_m \log p)}{p^{1+2\sigma}} = \sum_{x_m^{1/e} < p \leqslant x_m} \frac{1}{p^{1+2\sigma}} + O\left(\sum_{x_m^{1/e} < p \leqslant x_m} \frac{(|t_m| \log p)^2}{p^{1+2\sigma}}\right)$$

$$= \sum_{x_m^{1/e} < p \leqslant x_m} \frac{1}{p} + O\left(\sum_{x_m^{1/e} < p \leqslant x_m} \frac{|\sigma| \log p + (|t_m| \log p)^2}{p} \right)$$

$$= 1 + O\left(\frac{1}{\log x_m} + \frac{\log x_m}{\log x_n} + \frac{1}{m^{4/3}} \right)$$

$$= 1 + O\left(\frac{1}{e^{n-m}} + \frac{1}{m^{4/3}} \right),$$

under our conditions that $|\sigma| \leqslant 1/\log x_n$ and $|t_m| \leqslant \frac{1}{m^{2/3} \log x_m}$. Here we used the standard Mertens and Chebychev estimates, namely

$$\sum_{x_m^{1/e} < p \leqslant x_m} \frac{1}{p} = 1 + O\left(\frac{1}{\log x_m} \right) \quad \text{and} \quad \sum_{x_m^{1/e} < p \leqslant x_m} \frac{\log p}{p} \ll \log x_m. \qquad \square$$

At this point, it will be helpful to examine the kinds of events we shall actually be interested in for our application and their probabilities in the Gaussian case.

PROBABILITY RESULT 1. *Let a and n be large (that is, larger than a suitable absolute constant) and let $G_1, \ldots, G_n$ be independent Gaussian random variables, each having mean zero and variance between $1/20$ and 20 (say). Then uniformly for any function $h(j)$ satisfying $|h(j)| \leqslant 10 \log j$, we have*

$$\mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + h(j) \; \forall 1 \leqslant j \leqslant n \right) \asymp \min\left\{ 1, \frac{a}{\sqrt{n}} \right\}.$$

Since Probability Result 1 is a purely probabilistic statement about Gaussian random walks, we postpone its proof to the appendix. We remark that such a statement with $h(j) \equiv 0$ is standard. It requires a little more work to obtain the more general estimate, but it is natural to think such a result should hold because one typically thinks of random variables fluctuating on the order of their standard deviation (here $\asymp \sqrt{j}$ up to step j), so perturbing an event by a term much smaller than this should not alter its probability very much. Having a more general estimate will be very useful later (see, for example, the proof of Key Proposition 2, in Section 4) because, on taking exponentials, the function $h(j)$ can supply extra savings of powers of j in various places.

We will also need a small variant of Probability Result 1, where the event involves a nontrivial but relaxed lower barrier and a slightly tightened upper barrier as well.

PROBABILITY RESULT 2. *There is a large absolute constant B such that the following is true. Let a and n be large and let $G_1, \ldots, G_n$ be independent Gaussian random variables, each having mean zero and variance between $1/20$ and 20. Then uniformly for any function $h(j)$ satisfying $|h(j)| \leqslant 10 \log j$ and any function $g(j)$ satisfying $g(j) \leqslant -Bj$, we have*

$$\mathbb{P}\left(g(j) \leqslant \sum_{m=1}^{j} G_m \leqslant \min\{a, Bj\} + h(j) \ \forall 1 \leqslant j \leqslant n\right) \asymp \min\left\{1, \frac{a}{\sqrt{n}}\right\}.$$

Again, it is natural to think such a result should hold because, if the sum up to j typically fluctuates on the scale of $\sqrt{j}$, then the sums should be rather insensitive to a much more relaxed barrier of the shape Bj or $-Bj$. We postpone the proof to the appendix.

By combining Lemma 4 with Probability Results 1 and 2, we can finally prove the $\tilde{\mathbb{P}}$ probability estimate we shall need to obtain Theorem 1. Again, this requires a little care to ensure the size restrictions on u_j, v_j in Lemma 4 are respected when it is applied.

PROPOSITION 5. *There is a large natural number B such that the following is true.*

Let $n \leqslant \log \log x - (B+1)$ be large and define the decreasing sequence $(l_j)_{j=1}^{n}$ of nonnegative integers by $l_j := \lfloor \log \log x \rfloor - (B+1) - j$. Suppose that $|\sigma| \leqslant \frac{1}{e^{B+n+1}}$ and that $(t_j)_{j=1}^{n}$ is a sequence of real numbers satisfying $|t_j| \leqslant \frac{1}{j^{2/3} e^{B+j+1}}$ for all j.

Then uniformly for any large a and any function $h(j)$ satisfying $|h(j)| \leqslant 10 \log j$, and with $I_l(s)$ denoting the increments of the Euler product corresponding to a Steinhaus random multiplicative function (as before), we have

$$\tilde{\mathbb{P}}\left(-a - Bj \leqslant \sum_{m=1}^{j} \log |I_{l_m}(1/2 + \sigma + it_m)| \leqslant a + j + h(j) \ \forall j \leqslant n\right)$$
$$\asymp \min\left\{1, \frac{a}{\sqrt{n}}\right\}.$$

Proof of Proposition 5. We note first that $\log x_j = e^{-(l_j+1)} \log x = e^{B+j} \frac{\log x}{e^{\lfloor \log \log x \rfloor}}$ here, so our assumptions imply that $|\sigma| \leqslant 1/\log x_n$ and that $|t_j| \leqslant 1/(j^{2/3} \log x_j)$ for all $1 \leqslant j \leqslant n$. Furthermore, provided B is fixed large enough, then $x_1 = x^{e^{-(l_1+1)}}$ will be sufficiently large that Lemmas 3 and 4 may legitimately be applied.

We can lower bound the probability we are interested in by

$$\tilde{\mathbb{P}}\left(-Bj \leqslant \sum_{m=1}^{j} \log|I_{l_m}(1/2 + \sigma + it_m)| \leqslant \min\{a, Bj\} + j + h(j) \,\forall j \leqslant n\right),$$

and here since $\log x_j \geqslant e^{B+j}$, we have $Bj \leqslant (1/80)\sqrt{\log x_j}$ and $\min\{a, Bj\} + j + h(j) \leqslant (1/80)\sqrt{\log x_j}$ for all $1 \leqslant j \leqslant n$, provided B is fixed large enough. Thus, Lemma 4 is applicable and yields that our probability is

$$\gg \mathbb{P}\left(-(B+1)j + O(1) \leqslant \sum_{m=1}^{j} G_m \leqslant \min\{a, Bj\} + h(j) - O(1) \,\forall j \leqslant n\right),$$

where G_j are independent Gaussians with mean 0 and variance $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}}$. In particular, the variance here is $\leqslant \frac{e^{(2\log x_j)/\log x_n}}{2} \sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{p} \leqslant 5$, and similarly the variance is $\geqslant 1/20$, so we can use Probability Result 2 to deduce the desired lower bound

$$\mathbb{P}\left(-(B+1)j + O(1) \leqslant \sum_{m=1}^{j} G_m \leqslant \min\{a, Bj\} + h(j) - O(1) \,\forall j \leqslant n\right)$$
$$\gg \min\left\{1, \frac{a}{\sqrt{n}}\right\}.$$

Obtaining a matching upper bound for our probability will be slightly more involved, though not too much so. If $a \geqslant \sqrt{n}$, then the upper bound is trivial, so we may assume instead that $a < \sqrt{n}$. If our event occurs, we must have

$$-(B+1)a \leqslant \sum_{m \leqslant a} \log|I_{l_m}(1/2 + \sigma + it_m)| \leqslant 2a + h(\lfloor a \rfloor),$$

and this will be independent of the behaviour of the summands with $m > a$. So we may upper bound the probability we are interested in by

$$\tilde{\mathbb{P}}\left(-3a - h(\lfloor a \rfloor) - Bj \leqslant \sum_{a < m \leqslant j} \log|I_{l_m}(1/2 + \sigma + it_m)|\right.$$
$$\left. \leqslant (B+2)a + j + h(j) \,\forall a < j \leqslant n\right).$$

Now the point is that (since $\log x_j \geqslant e^{B+j}$) we have $3a + h(\lfloor a \rfloor) + Bj \leqslant (1/80)\sqrt{\log x_j}$ and $(B+2)a + j + h(j) \leqslant (1/80)\sqrt{\log x_j}$ for $j > a$, so Lemma 4 is now applicable and we may upper bound our probability by

$$\mathbb{P}\left(-2a - (B+1)j - h(\lfloor a \rfloor) - O(1) \leqslant \sum_{a < m \leqslant j} G_m \right.$$

$$\left. \leqslant (B+2)a + a + h(j) + O(1) \;\forall a < j \leqslant n \right).$$

Note here that Lemma 4 is applied with j replaced by $j - \lfloor a \rfloor$, hence the added a terms in the upper and lower limits. Finally, Probability Result 1 is applicable, and since B is an absolute constant and $n - a \geqslant n - \sqrt{n} \gg n$, we obtain our desired upper bound $\ll \min\{1, \frac{Ba}{\sqrt{n}}\} \ll \min\{1, \frac{a}{\sqrt{n}}\}$ for the probability. $\qquad\square$

3.3. Girsanov-type calculations in the Rademacher case. In this section, we develop an analogue of Proposition 5 to cover Rademacher random multiplicative functions. The set-up and arguments are very similar as in the Steinhaus case, broadly speaking, but there is the usual complication that $f(n)n^{it}$ no longer has the same law as $f(n)$, and both the formulation and proofs must be adjusted to address this. One already sees this issue when comparing the Euler product calculations in Lemma 2 with those in Lemma 1.

Thus, for each $t \in \mathbb{R}$, and for large x and $-1/100 \leqslant \sigma \leqslant 1/100$, we define a tilted probability measure $\tilde{\mathbb{P}}_t^{\mathrm{Rad}} = \tilde{\mathbb{P}}_{x,\sigma,t}^{\mathrm{Rad}}$ by setting

$$\tilde{\mathbb{P}}_t^{\mathrm{Rad}}(A) := \frac{\mathbb{E}\mathbf{1}_A \prod_{p \leqslant x^{1/e}} \left|1 + \frac{f(p)}{p^{1/2+\sigma+it}}\right|^2}{\mathbb{E} \prod_{p \leqslant x^{1/e}} \left|1 + \frac{f(p)}{p^{1/2+\sigma+it}}\right|^2}$$

for each event A, where f is a Rademacher random multiplicative function. Again, we will sometimes write $\tilde{\mathbb{E}}_t^{\mathrm{Rad}}$ to denote expectation (that is, integration) with respect to the measure $\tilde{\mathbb{P}}_t^{\mathrm{Rad}}$.

Now for each $l \in \mathbb{N} \cup \{0\}$, we set $I_l(s) := \prod_{x^{e^{-(l+2)}} < p \leqslant x^{e^{-(l+1)}}} (1 + \frac{f(p)}{p^s})$, the lth 'increment' of the Euler product corresponding to f. There should be no confusion with the corresponding notation from the Steinhaus case since one simply takes increments of the appropriate Euler product for the kind of random multiplicative function one is working with. As before, we let $(l_j)_{j=1}^n$ denote a *strictly decreasing* sequence of nonnegative integers, with $l_1 \leqslant \log\log x - 2$, and define a corresponding increasing sequence of real numbers $(x_j)_{j=1}^n$ by setting $x_j := x^{e^{-(l_j+1)}}$.

The following result is the Rademacher analogue of Lemma 4. Note that in Lemma 4, we had $t = 0$ (although, because of translation invariance in law, this was not actually a restriction), so the condition on $|t_j - t|$ below is analogous to the condition on $|t_j|$ that we had there.

LEMMA 5. *Let $t \in \mathbb{R} \backslash \{0\}$, and let the situation be as above.*

Suppose that $x_1 \geqslant \max\{e^{C/|t|}, e^{C\log^2|t|}\}$ is large, that $|\sigma| \leqslant 1/\log x_n$ and that the numbers $(t_j)_{j=1}^n$ satisfy $|t_j - t| \leqslant \frac{1}{j^{2/3}\log x_j}$. Suppose $(u_j)_{j=1}^n$ and $(v_j)_{j=1}^n$ are sequences of real numbers satisfying

$$-(1/80)\sqrt{\log x_j} \leqslant u_j \leqslant v_j \leqslant (1/80)\sqrt{\log x_j} \quad \forall 1 \leqslant j \leqslant n.$$

Then we have

$$\mathbb{P}\left((u_j - j) + O(1) \leqslant \sum_{m=1}^j G_m \leqslant (v_j - j) - O(1) \; \forall j \leqslant n\right)$$

$$\ll \tilde{\mathbb{P}}_t^{\mathrm{Rad}}\left(u_j \leqslant \sum_{m=1}^j \log|I_{l_m}(1/2 + \sigma + it_m)| \leqslant v_j \; \forall j \leqslant n\right)$$

$$\ll \mathbb{P}\left((u_j - j) - O(1) \leqslant \sum_{m=1}^j G_m \leqslant (v_j - j) + O(1) \; \forall j \leqslant n\right),$$

where G_m are independent Gaussian random variables, each having mean 0 and variance $\sum_{x_m^{1/e} < p \leqslant x_m} \frac{1 + \cos(2t_m \log p)}{2p^{1+2\sigma}}$.

Proof of Lemma 5. Lemma 2 implies that for any $|u| \leqslant x_j^{1/20}$, say, the characteristic function $\tilde{\mathbb{E}}_t^{\mathrm{Rad}} e^{iu \log|I_{l_j}(1/2+\sigma+it_j)|}$ is

$$= \exp\left\{\sum_{x_j^{1/e} < p \leqslant x_j} \frac{iuc(t, t_j - t, p) - (u^2/4)(1 + \cos(2t_j \log p))}{p^{1+2\sigma}} + T(u) - T(0)\right\},$$

where $c(t, t_j - t, p) = 2\cos(t \log p)\cos(t_j \log p) - (1/2)\cos(2t_j \log p)$. Using a standard trigonometric identity, together with our assumption on $|t_j - t|$, we have

$$2\cos(t \log p)\cos(t_j \log p) = \cos((t + t_j)\log p) + \cos((t - t_j)\log p)$$
$$= \cos((t + t_j)\log p) + 1 + O(1/j^{4/3}).$$

As noted in the proof of Lemma 4, we have

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{p^{1+2\sigma}} = 1 + O\left(\frac{1}{\log x_j} + \frac{\log x_j}{\log x_n}\right).$$

Furthermore, recall we assume that $x_1 \geqslant \max\{e^{C/|t|}, e^{C \log^2 |t|}\}$ is large, which implies that $x_j^{1/e} \geqslant x_1^{1/e} \geqslant \max\{e^{C/e|t|}, e^{(C/e) \log^2 |t|}\}$. In particular, this means that $|t| \geqslant \frac{C}{\log x_j}$, and so $t + t_j \asymp t$ and $2t_j \asymp t$. Thus, a strong form of the prime number theorem implies that both

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos((t + t_j) \log p)}{p^{1+2\sigma}},$$

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(2t_j \log p)}{p^{1+2\sigma}} \ll \frac{1}{|t| \log x_j} \ll \frac{1}{e^j |t| \log x_1} \ll \frac{1}{Ce^j}.$$

See Section 6.1 of Harper [10] for details of such calculations.

Having made these preliminary observations, we can check the proofs of Lemmas 3 and 4 and see they carry over to the present case. Here the relevant Gaussian random variables N_j will have mean $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{c(t, t_j - t, p)}{p^{1+2\sigma}}$ and variance $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1 + \cos(2t_j \log p)}{2p^{1+2\sigma}}$ (rather than mean $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t_j \log p)}{p^{1+2\sigma}}$ and variance $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}}$, in the Steinhaus case), but our assumptions on t_j, x_j and the above calculations show we still have

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{c(t, t_j - t, p)}{p^{1+2\sigma}} = 1 + O\left(\frac{1}{\log x_j} + \frac{\log x_j}{\log x_n} + \frac{1}{j^{4/3}}\right)$$

$$= 1 + O\left(\frac{1}{e^{n-j}} + \frac{1}{j^{4/3}}\right)$$

as well as

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1 + \cos(2t_j \log p)}{2p^{1+2\sigma}} = \sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}} + O(1/Ce^j).$$

Thus, the arguments of Lemmas 3 and 4 go through to prove Lemma 5. $\qquad \square$

Combining Lemma 5 with Probability Results 1 and 2 leads to the following proposition, which is a Rademacher analogue of Proposition 5.

PROPOSITION 6. *There is a large natural number B such that the following is true.*

Let $t \in \mathbb{R} \backslash \{0\}$ and let $D \geqslant \max\{\log(1/|t|), 2 \log \log(1 + |t|)\} + (B + 1)$ be any natural number. Let $n \leqslant \log \log x - D$ be large, and define the decreasing sequence $(l_j)_{j=1}^n$ of nonnegative integers by $l_j := \lfloor \log \log x \rfloor - D - j$. Suppose that $|\sigma| \leqslant \frac{1}{e^{D+n}}$ and that $(t_j)_{j=1}^n$ is a real sequence satisfying $|t_j - t| \leqslant \frac{1}{j^{2/3} e^{D+j}}$ for all j.

Then uniformly for any large a and any function $h(j)$ satisfying $|h(j)| \leqslant 10 \log j$, and with $I_l(s)$ denoting the increments of the Euler product corresponding to a Rademacher random multiplicative function (as before), we have

$$\tilde{\mathbb{P}}_t^{\mathrm{Rad}} \left(-a - Bj \leqslant \sum_{m=1}^{j} \log |I_{l_m}(1/2 + \sigma + it_m)| \leqslant a + j + h(j) \; \forall j \leqslant n \right)$$
$$\asymp \min \left\{ 1, \frac{a}{\sqrt{n}} \right\}.$$

Proof of Proposition 6. The only nontrivial change from the proof of Proposition 5 is that we must verify the extra condition $x_1 \geqslant \max\{e^{C/|t|}, e^{C \log^2 |t|}\}$ in Lemma 5 (as compared with Lemma 4). However, since we have $\log x_j = e^{-(l_j+1)} \log x = e^{D+j-1} \frac{\log x}{e^{\lfloor \log \log x \rfloor}}$, in particular, we have $\log x_1 \geqslant e^D \geqslant e^{B+1} \max\{\frac{1}{|t|}, \log^2(1 + |t|)\}$, so the condition will be satisfied provided B is fixed large enough in terms of the absolute constant C in Lemma 5. $\square$

4. Proofs of the upper bounds in Theorems 1 and 2

4.1. The upper bound in the Steinhaus case. For each $|t| \leqslant 1/2$, set $t(-1) = t$, and then iteratively for each $0 \leqslant j \leqslant \log \log x - 2$, define

$$t(j) := \max \left\{ u \leqslant t(j-1) : \right.$$
$$\left. u = \frac{n}{((\log x)/e^{j+1}) \log((\log x)/e^{j+1})} \text{ for some } n \in \mathbb{Z} \right\}.$$

Thus, the points $t(j)$ form a sequence of approximations to t, in which for each j, we have

$$|t - t(j)| = t - t(j) = \sum_{l=0}^{j} t(l-1) - t(l)$$

$$\leqslant \sum_{l=0}^{j} \frac{1}{((\log x)/e^{l+1}) \log((\log x)/e^{l+1})}$$

$$\leqslant \frac{2}{((\log x)/e^{j+1}) \log((\log x)/e^{j+1})}. \qquad (5)$$

With our choice of labelling, these approximations become coarser (that is, worse) as j increases, but recall that in our labelling, the lengths of our Euler products $F_j(s)$ also go down as j increases, so a coarser approximation will turn out to be acceptable.

Furthermore, recall that we let $I_l(s) = \prod_{x^{e^{-(l+2)}} < p \leqslant x^{e^{-(l+1)}}} (1 - \frac{f(p)}{p^s})^{-1}$, the lth 'increment' of the Euler product corresponding to Steinhaus f.

Given this notation, let B be the large fixed natural number from Proposition 5 and let $\mathcal{G}(k)$ denote the event that for all $|t| \leqslant 1/2$ and all $k \leqslant j \leqslant \log \log x - B - 2$, we have

$$\left(\frac{\log x}{e^{j+1}} e^{g(x,j)} \right)^{-1} \leqslant \prod_{l=j}^{\lfloor \log \log x \rfloor - B - 2} \left| I_l \left(1/2 - \frac{k}{\log x} + it(l) \right) \right| \leqslant \frac{\log x}{e^{j+1}} e^{g(x,j)},$$

where $g(x, j) := C \min\{\sqrt{\log \log x}, \frac{1}{1-q}\} + 2 \log \log(\frac{\log x}{e^{j+1}})$ for a large constant C. Thus, $\mathcal{G}(k)$ is the event that our Euler product is not too large or too small on any 'scale' $k \leqslant j \leqslant \log \log x - B - 2$. The fact that one has a different point $t(l)$ in each part I_l of the Euler product is a little inelegant: it would be nice to keep t itself everywhere, in which case for each j, the product would essentially just be $|F_j(1/2 - k/\log x + it)|$. But this technical device of varying t slightly, *and thereby discretizing the set of t*, will make it easier to prove an important estimate we shall need (Key Proposition 2). Similarly, the lower bound condition in the definition of $\mathcal{G}(k)$ is really a technical device to make Proposition 5 applicable: it is the upper bound condition that imposes a significant constraint.

With the above preparations, we can state two key estimates, from which we will deduce the upper bound part of Theorem 1.

KEY PROPOSITION 1. *For all large x and uniformly for $0 \leqslant k \leqslant \mathcal{K} = \lfloor \log \log \log x \rfloor$ and $2/3 \leqslant q \leqslant 1$, we have*

$$\mathbb{E}\left(\mathbf{1}_{\mathcal{G}(k)}\int_{-1/2}^{1/2}\left|F_k\left(1/2-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q$$

$$\ll\left(\frac{\log x}{e^k}C\min\left\{1,\frac{1}{(1-q)\sqrt{\log\log x}}\right\}\right)^q,$$

*where **1** denotes the indicator function.*

KEY PROPOSITION 2. *For all large x, and uniformly for $0\leqslant k\leqslant\mathcal{K}=\lfloor\log\log\log x\rfloor$ and $2/3\leqslant q\leqslant 1$, we have*

$$\mathbb{P}(\mathcal{G}(k)\text{ fails})\ll e^{-2C\min\{\sqrt{\log\log x},\frac{1}{1-q}\}}.$$

Proof of the upper bound in Theorem 1, assuming Key Propositions 1 and 2. If q satisfies $1-\frac{1}{\sqrt{\log\log x}}\leqslant q\leqslant 1$, then the upper bound we need to prove is $\mathbb{E}|\sum_{n\leqslant x}f(n)|^{2q}\ll x^q$, which is trivial in view of Hölder's inequality and the estimate $\mathbb{E}|\sum_{n\leqslant x}f(n)|^2\leqslant x$. In view of this and of Proposition 1, the upper bound in Theorem 1 will follow if we can show that

$$\mathbb{E}\left(\frac{e^k(1-q)\sqrt{\log\log x}}{\log x}\int_{-1/2}^{1/2}\left|F_k\left(1/2-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q\ll 1,$$

uniformly for all $0\leqslant k\leqslant\mathcal{K}=\lfloor\log\log\log x\rfloor$ and $2/3\leqslant q\leqslant 1-\frac{1}{\sqrt{\log\log x}}$.

To prove this, for $1/\sqrt{\log\log x}\leqslant\delta\leqslant 1/6$ (say), we define

$$R(\delta)=R(\delta,k,x)$$

$$:=\sup_{1-2\delta\leqslant q\leqslant 1-\delta}\mathbb{E}\left(\frac{e^k(1-q)\sqrt{\log\log x}}{\log x}\right.$$

$$\left.\times\int_{-1/2}^{1/2}\left|F_k\left(1/2-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q.$$

We have

$$\mathbb{E}\left(\int_{-1/2}^{1/2}\left|F_k\left(1/2-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q$$

$$=\mathbb{E}\left(\mathbf{1}_{\mathcal{G}(k)}\int_{-1/2}^{1/2}\left|F_k\left(\frac{1}{2}-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q$$

$$+\mathbb{E}\left(\mathbf{1}_{\mathcal{G}(k)\text{ fails}}\int_{-1/2}^{1/2}\left|F_k\left(\frac{1}{2}-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q,$$

so applying Key Proposition 1 with C a large fixed constant, we obtain

$$
\begin{aligned}
R(\delta) &\ll C^{1-\delta} + \sup_{1-2\delta \leqslant q \leqslant 1-\delta} \mathbb{E}\left(\mathbf{1}_{\mathcal{G}(k)\text{ fails}} \frac{e^k(1-q)\sqrt{\log\log x}}{\log x}\right.\\
&\qquad\qquad \left.\times \int_{-1/2}^{1/2}\left|F_k\left(\frac{1}{2}-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q\\
&\leqslant C + \sup_{1-2\delta \leqslant q \leqslant 1-\delta} \mathbb{E}\left(\mathbf{1}_{\mathcal{G}(k)\text{ fails}} \frac{e^k(1-q)\sqrt{\log\log x}}{\log x}\right.\\
&\qquad\qquad \left.\times \int_{-1/2}^{1/2}\left|F_k\left(\frac{1}{2}-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q.
\end{aligned}
$$

Next, for each $1-2\delta \leqslant q \leqslant 1-\delta$, we set $q' = (1+q)/2$ so that $1-\delta \leqslant q' \leqslant 1-\delta/2$. Then by Hölder's inequality with exponents $q'/(q'-q)$ and q'/q, we have

$$
\begin{aligned}
&\mathbb{E}\left(\mathbf{1}_{\mathcal{G}(k)\text{ fails}} \frac{e^k(1-q)\sqrt{\log\log x}}{\log x} \int_{-1/2}^{1/2}\left|F_k\left(1/2-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q\\
&\leqslant (\mathbb{E}\mathbf{1}_{\mathcal{G}(k)\text{ fails}})^{(q'-q)/q'}\left(\mathbb{E}\left(\frac{e^k(1-q)\sqrt{\log\log x}}{\log x}\right.\right.\\
&\qquad \left.\left.\times \int_{-1/2}^{1/2}\left|F_k\left(1/2-\frac{k}{\log x}+it\right)\right|^2 dt\right)^{q'}\right)^{q/q'}\\
&\ll \mathbb{P}(\mathcal{G}(k)\text{ fails})^{\delta/2}\left(\mathbb{E}\left(\frac{e^k(1-q')\sqrt{\log\log x}}{\log x}\right.\right.\\
&\qquad \left.\left.\times \int_{-1/2}^{1/2}\left|F_k\left(1/2-\frac{k}{\log x}+it\right)\right|^2 dt\right)^{q'}\right)^{q/q'}.
\end{aligned}
$$

The point is that, by Key Proposition 2, we have $\mathbb{P}(\mathcal{G}(k)$ fails$) \ll e^{-2C\min\{\sqrt{\log\log x}, \frac{1}{1-q}\}} \leqslant e^{-C/\delta}$. So substituting back above, noting that always $q/q' \leqslant 1$, we deduce the recursive bound

$$
R(\delta) \ll C + e^{-C/2}(1 + R(\delta/2)) \ll C + e^{-C/2}R(\delta/2).
$$

Iterating the recursive bound (with C fixed sufficiently large to compensate for the implicit constant there) and replacing δ by $\delta/2$, $\delta/4$, $\delta/8$, and so on, we

see that uniformly for $1/\sqrt{\log \log x} \leqslant \delta \leqslant 1/6$, we have

$$R(\delta) \ll 1 + R(1/\sqrt{\log \log x}).$$

However, using Hölder's inequality (and using the fact that $f(n)n^{it}$ has the same law as $f(n)$ in the Steinhaus case and using (2)), we have trivially that

$$
\begin{aligned}
R(1/\sqrt{\log \log x}) &\ll \sup_{1-2/\sqrt{\log \log x} \leqslant q \leqslant 1-1/\sqrt{\log \log x}} \\
&\quad \times \mathbb{E} \left(\frac{e^k}{\log x} \int_{-1/2}^{1/2} \left| F_k \left(\frac{1}{2} - \frac{k}{\log x} + it \right) \right|^2 dt \right)^q \\
&\leqslant \sup_{1-2/\sqrt{\log \log x} \leqslant q \leqslant 1-1/\sqrt{\log \log x}} \\
&\quad \times \left(\frac{e^k}{\log x} \int_{-1/2}^{1/2} \mathbb{E} \left| F_k \left(\frac{1}{2} - \frac{k}{\log x} + it \right) \right|^2 dt \right)^q \\
&= \sup_{1-2/\sqrt{\log \log x} \leqslant q \leqslant 1-1/\sqrt{\log \log x}} \\
&\quad \times \left(\frac{e^k}{\log x} \mathbb{E} \left| F_k \left(\frac{1}{2} - \frac{k}{\log x} \right) \right|^2 \right)^q \\
&\ll 1.
\end{aligned}
$$

Inserting this above implies that $R(\delta) \ll 1$ for all $1/\sqrt{\log \log x} \leqslant \delta \leqslant 1/6$, which yields our upper bound theorem. $\qquad\square$

4.2. Proof of Key Proposition 1.

By Hölder's inequality, to prove Key Proposition 1, it will suffice to show that

$$
\begin{aligned}
\mathbb{E} \mathbf{1}_{\mathcal{G}(k)} \int_{-1/2}^{1/2} &\left| F_k \left(1/2 - \frac{k}{\log x} + it \right) \right|^2 dt \\
&\ll \frac{\log x}{e^k} C \min \left\{ 1, \frac{1}{(1-q)\sqrt{\log \log x}} \right\},
\end{aligned}
$$

uniformly for $0 \leqslant k \leqslant \mathcal{K} = \lfloor \log \log \log x \rfloor$ and $2/3 \leqslant q \leqslant 1$.

We can upper bound the left-hand side by

$$\int_{-1/2}^{1/2} \mathbb{E} \mathbf{1}_{\mathcal{G}(k,t)} \left| F_k \left(1/2 - \frac{k}{\log x} + it \right) \right|^2 dt,$$

where $\mathcal{G}(k, t)$ denotes the event that

$$\left(\frac{\log x}{e^{j+1}} e^{g(x,j)}\right)^{-1} \leqslant \prod_{l=j}^{\lfloor \log \log x \rfloor - B - 2} \left| I_l\left(1/2 - \frac{k}{\log x} + it(l)\right)\right| \leqslant \frac{\log x}{e^{j+1}} e^{g(x,j)}$$

for all $k \leqslant j \leqslant \log \log x - B - 2$. This is an upper bound because $\mathcal{G}(k)$ is the event that $\mathcal{G}(k, t)$ holds for all $|t| \leqslant 1/2$. Furthermore, since the law of $f(n)$ is the same as the law of $f(n)n^{-it}$, we have

$$\mathbb{E}\mathbf{1}_{\mathcal{G}(k,t)} \left| F_k\left(1/2 - \frac{k}{\log x} + it\right)\right|^2 = \mathbb{E}\mathbf{1}_{\mathcal{H}(k,t)} \left| F_k\left(1/2 - \frac{k}{\log x}\right)\right|^2,$$

where $\mathcal{H}(k, t)$ denotes the event that

$$\left(\frac{\log x}{e^{j+1}} e^{g(x,j)}\right)^{-1} \leqslant \prod_{l=j}^{\lfloor \log \log x \rfloor - B - 2} \left| I_l\left(\frac{1}{2} - \frac{k}{\log x} + i(t(l) - t)\right)\right| \leqslant \frac{\log x}{e^{j+1}} e^{g(x,j)}$$
$$\leqslant e^{\lfloor \log \log x \rfloor - (B+1) - j} e^{B+1+g(x,j)}$$

for all $k \leqslant j \leqslant \log \log x - B - 2$.

Now to understand the event $\mathcal{H}(k, t)$ further, we can take logarithms everywhere in the preceding display and reparametrize by replacing l by $\lfloor \log \log x \rfloor - (B + 1) - m$, where $1 \leqslant m \leqslant \lfloor \log \log x \rfloor - (B + 1) - j$. Let us also recall, using the inequality $\log(X + Y) \leqslant \log X + \log(1 + Y)$ that holds when $X, Y \geqslant 1$, that

$$g(x, j) = 2 \log\log\left(\frac{\log x}{e^{j+1}}\right) + C \min\left\{\sqrt{\log\log x}, \frac{1}{1-q}\right\}$$
$$\leqslant 2 \log(\lfloor \log\log x \rfloor - (B+1) - j) + 2\log(B+2)$$
$$+ C \min\left\{\sqrt{\log\log x}, \frac{1}{1-q}\right\}.$$

Combining these observations, we see that $\mathcal{H}(k, t)$ is an event of the form treated in Proposition 5, taking $n = \lfloor \log \log x \rfloor - (B + 1) - k$, $\sigma = -\frac{k}{\log x}$, $t_m = t(\lfloor \log \log x \rfloor - (B + 1) - m) - t$ for all m and

$$a = C \min\left\{\sqrt{\log\log x}, \frac{1}{1-q}\right\} + (B+1) + 2\log(B+2), \quad h(j) = 2 \log j.$$

We may check that these parameters do satisfy the condition $|\sigma| \leqslant \frac{1}{e^{B+n+1}}$ and (using the approximation bound (5)) that $|t_m| \leqslant \frac{1}{m^{2/3}e^{B+m+1}}$ for all m.

So Proposition 5 implies that

$$\frac{\mathbb{E}\mathbf{1}_{\mathcal{H}(k,t)}|F_k(1/2 - \frac{k}{\log x})|^2}{\mathbb{E}|F_k(1/2 - \frac{k}{\log x})|^2} = \tilde{\mathbb{P}}(\mathcal{H}(k,t)) \ll \min\left\{1, \frac{a}{\sqrt{n}}\right\}$$

$$\ll \min\left\{1, C, \frac{C}{(1-q)\sqrt{\log\log x}}\right\}.$$

(Here the lower bound we impose on our product corresponds to a lower bound $-(a + j + h(j))$ in Proposition 5. This is a more stringent bound than the barrier $-a - Bj$ allowed there, so the Proposition 5 upper bound is certainly still applicable.) Finally, Key Proposition 1 follows by combining the above display with the fact that

$$\mathbb{E}\left|F_k\left(1/2 - \frac{k}{\log x}\right)\right|^2 = \exp\left\{\sum_{p \leqslant x^{e^{-(k+1)}}} \frac{1}{p^{1-2k/\log x}} + O(1)\right\}$$

$$= \exp\left\{\sum_{p \leqslant x^{e^{-(k+1)}}} \frac{1}{p} + O\left(\sum_{p \leqslant x^{e^{-(k+1)}}} \frac{k\log p}{p\log x} + 1\right)\right\} \ll \frac{\log x}{e^k},$$

which follows from (2) and the Mertens and Chebychev estimates for sums over primes (see, for example, Theorem 2.7 of Montgomery and Vaughan [20]). $\qquad\square$

4.3. Proof of Key Proposition 2.

By the union bound, we see $\mathbb{P}(\mathcal{G}(k)$ fails) is at most

$$\sum_{k \leqslant j \leqslant \log\log x - B - 2} \mathbb{P}\left(\prod_{l=j}^{\lfloor\log\log x\rfloor - B - 2}\left|I_l\left(\frac{1}{2} - \frac{k}{\log x} + it(l)\right)\right|\right.$$

$$> \frac{\log x}{e^{j+1}}e^{g(x,j)} \text{ for some } |t| \leqslant \frac{1}{2}\right)$$

$$+ \sum_{k \leqslant j \leqslant \log\log x - B - 2} \mathbb{P}\left(\prod_{l=j}^{\lfloor\log\log x\rfloor - B - 2}\left|I_l\left(\frac{1}{2} - \frac{k}{\log x} + it(l)\right)\right|^{-1}\right.$$

$$> \frac{\log x}{e^{j+1}}e^{g(x,j)} \text{ for some } |t| \leqslant \frac{1}{2}\right).$$

Let us temporarily write Σ_1, Σ_2 to denote these two sums.

We concentrate on trying to bound Σ_1. Because of the definition of $t(l)$, we see the probability inside the sum is at most as large as with the set $|t| \leqslant 1/2$ replaced by the set

$$\mathcal{T}(x, j) := \left\{ \frac{n}{((\log x)/e^{j+1}) \log((\log x)/e^{j+1})} : \right.$$
$$\left. |n| \leqslant ((\log x)/e^{j+1}) \log((\log x)/e^{j+1}) \right\}.$$

Since we are now dealing with a discrete set of points, we can apply the union bound again to obtain that

$$\Sigma_1 \leqslant \sum_{k \leqslant j \leqslant \log\log x - B - 2} \sum_{t(j) \in \mathcal{T}(x,j)} \mathbb{P}\left(\prod_{l=j}^{\lfloor \log\log x \rfloor - B - 2} \left| I_l\left(1/2 - \frac{k}{\log x} + it(l) \right) \right| \right.$$
$$\left. > \frac{\log x}{e^{j+1}} e^{g(x,j)} \right),$$

and by Chebychev's inequality, this is all

$$\leqslant \sum_{k \leqslant j \leqslant \log\log x - B - 2} \sum_{t(j) \in \mathcal{T}(x,j)} \frac{1}{(\frac{\log x}{e^{j+1}} e^{g(x,j)})^2}$$
$$\times \mathbb{E} \prod_{l=j}^{\lfloor \log\log x \rfloor - B - 2} \left| I_l\left(1/2 - \frac{k}{\log x} + it(l) \right) \right|^2.$$

Finally, since f is independent on distinct primes (so the different increments $I_l(s)$ are independent) and (in the Steinhaus case) its law is the same as the law of $f(n) n^{it}$ for any fixed t, the above is

$$= \sum_{k \leqslant j \leqslant \log\log x - B - 2} \sum_{t(j) \in \mathcal{T}(x,j)} \frac{1}{(\frac{\log x}{e^{j+1}} e^{g(x,j)})^2} \prod_{l=j}^{\lfloor \log\log x \rfloor - B - 2} \mathbb{E} \left| I_l\left(1/2 - \frac{k}{\log x} \right) \right|^2$$
$$\ll \sum_{k \leqslant j \leqslant \log\log x - B - 2} e^{-2g(x,j)} \frac{\log((\log x)/e^{j+1})}{\frac{\log x}{e^{j+1}}}$$
$$\times \prod_{l=j}^{\lfloor \log\log x \rfloor - B - 2} \mathbb{E} \left| I_l\left(1/2 - \frac{k}{\log x} \right) \right|^2.$$

As noted in (2) (and in the proof of Key Proposition 1), Lemma 1 implies that the product of expectations is $\ll \exp\{\sum_{p \leqslant x^{e^{-(j+1)}}} \frac{1}{p^{1-2k/\log x}}\} \ll \frac{\log x}{e^{j+1}}$, and inserting this

and the fact that $g(x, j) := C \min\{\sqrt{\log\log x}, \frac{1}{1-q}\} + 2\log\log(\frac{\log x}{e^{j+1}})$, we obtain

$$\Sigma_1 \ll \sum_{k \leqslant j \leqslant \log\log x - B - 2} e^{-2C\min\{\sqrt{\log\log x}, \frac{1}{1-q}\} - 3\log\log(\frac{\log x}{e^{j+1}})} \ll e^{-2C\min\{\sqrt{\log\log x}, \frac{1}{1-q}\}}.$$

One can bound Σ_2 in exactly the same way since (as remarked following (2)) one has the same estimate for $\mathbb{E}|I_l(1/2 - \frac{k}{\log x})|^{-2}$ as for $\mathbb{E}|I_l(1/2 - \frac{k}{\log x})|^2$. This completes the proof of Key Proposition 2. $\qquad\square$

4.4. The upper bound in the Rademacher case. Let x be large and let $0 \leqslant k \leqslant \mathcal{K} = \lfloor \log\log\log x \rfloor$ and $2/3 \leqslant q \leqslant 1$. We note immediately that, by Hölder's inequality, we have

$$\left\| \int_{|t| \leqslant 1/\sqrt{\log\log x}} \left| F_k\left(1/2 - \frac{k}{\log x} + it\right) \right|^2 dt \right\|_q$$

$$\leqslant \int_{|t| \leqslant 1/\sqrt{\log\log x}} \mathbb{E}\left| F_k\left(1/2 - \frac{k}{\log x} + it\right) \right|^2 dt$$

$$\ll \frac{\log x}{e^k \sqrt{\log\log x}}.$$

Here, we used the estimate

$$\mathbb{E}\left| F_k\left(1/2 - \frac{k}{\log x} + it\right) \right|^2 \ll \exp\left\{ \sum_{p \leqslant x^{e^{-(k+1)}}} \frac{1}{p^{1-2k/\log x}} \right\} \ll (\log x)/e^k,$$

which follows from (3) and from standard estimates for sums over primes (as in the proofs of Key Propositions 1 and 2). Similarly, when $|N| \geqslant (\log\log x)^2$, we have

$$\frac{1}{|N|^{1/4}} \left\| \int_{N-1/2}^{N+1/2} \left| F_k\left(\frac{1}{2} - \frac{k}{\log x} + it\right) \right|^2 dt \right\|_q$$

$$\leqslant \frac{1}{|N|^{1/4}} \int_{N-1/2}^{N+1/2} \mathbb{E}\left| F_k\left(\frac{1}{2} - \frac{k}{\log x} + it\right) \right|^2 dt$$

$$\ll \frac{\log x}{e^k |N|^{1/4}} \ll \frac{\log x}{e^k \sqrt{\log\log x}}.$$

If we insert these bounds into Proposition 2, we obtain an acceptable contribution for the Theorem 2 upper bound. Therefore, to finish the proof of that upper bound using Proposition 2, it will suffice to show that

$$\frac{1}{(|N|+1)^{q/4}} \mathbb{E}\left(\int_{\substack{|t-N|\leqslant 1/2, \\ |t|>1/\sqrt{\log\log x}}} \left|F_k\left(\frac{1}{2}-\frac{k}{\log x}+it\right)\right|^2 dt\right)^q$$

$$\ll \left(\frac{\log x}{e^k}\min\left\{1,\frac{1}{(1-q)\sqrt{\log\log x}}\right\}\right)^q,$$

uniformly for $0 \leqslant k \leqslant \mathcal{K} = \lfloor\log\log\log x\rfloor$ and $2/3 \leqslant q \leqslant 1$ and $|N| \leqslant (\log\log x)^2$. The preceding reductions, dismissing very small and large t from consideration, will be convenient when we come to apply Proposition 6.

We will concentrate on the case $N = 0$ since this will reveal essentially all the important differences between the Rademacher and Steinhaus arguments. Thus, for each $1/\sqrt{\log\log x} < |t| \leqslant 1/2$, we can define the sequence of approximations $(t(j))_{0\leqslant j\leqslant\log\log x-2}$ as we did in the Steinhaus case in Section 4.1 and have the same bounds (5) on $|t - t(j)|$ as we did there. Let us further define $D(t) := \lceil\log(1/|t|)\rceil + (B + 1)$, where B is as in Proposition 6. Note that $D(t) \ll \log\log\log x$ for $1/\sqrt{\log\log x} < |t| \leqslant 1/2$. Then with $I_l(s) = \prod_{x^{e^{-(l+2)}} < p\leqslant x^{e^{-(l+1)}}}(1 + \frac{f(p)}{p^s})$ denoting the lth 'increment' of the Rademacher Euler product, we will let $\mathcal{G}^{\mathrm{Rad}}(k, t)$ denote the event that for all $k \leqslant j \leqslant \log\log x - D - 1$, we have

$$\left(\frac{\log x}{e^{j+1}}e^{g(x,j)}\right)^{-1} \leqslant \prod_{l=j}^{\lfloor\log\log x\rfloor-D-1}\left|I_l\left(1/2-\frac{k}{\log x}+it(l)\right)\right| \leqslant \frac{\log x}{e^{j+1}}e^{g(x,j)},$$

where $g(x, j) := C\min\{\sqrt{\log\log x}, \frac{1}{1-q}\} + 2\log\log(\frac{\log x}{e^{j+1}})$ for a large constant C. Furthermore, we let $\mathcal{G}^{\mathrm{Rad}}(k)$ denote the event that $\mathcal{G}^{\mathrm{Rad}}(k, t)$ holds for all $1/\sqrt{\log\log x} < |t| \leqslant 1/2$. Note that this set-up is as close as possible to what we did in the Steinhaus case, the only real change being the introduction of the term $D(t)$ to trim down the range of j that we examine, which ensures we will be able to apply Proposition 6.

Now we have the following two key estimates.

KEY PROPOSITION 3. *For all large x, and uniformly for $0 \leqslant k \leqslant \mathcal{K} = \lfloor\log\log\log x\rfloor$ and $2/3 \leqslant q \leqslant 1$, we have*

$$\mathbb{E}\left(\mathbf{1}_{\mathcal{G}^{\mathrm{Rad}}(k)}\int_{1/\sqrt{\log\log x}<|t|\leqslant 1/2}\left|F_k\left(\frac{1}{2}-\frac{k}{\log x}+it\right)\right|^2\,dt\right)^q$$

$$\ll\left(\frac{\log x}{e^k}C\min\left\{1,\frac{1}{(1-q)\sqrt{\log\log x}}\right\}\right)^q,$$

where $\mathbf{1}$ denotes the indicator function.

KEY PROPOSITION 4. *For all large x, and uniformly for $0\leqslant k\leqslant\mathcal{K}=\lfloor\log\log\log x\rfloor$ and $2/3\leqslant q\leqslant 1$, we have*

$$\mathbb{P}(\mathcal{G}^{\mathrm{Rad}}(k)\text{ fails})\ll e^{-2C\min\{\sqrt{\log\log x},\frac{1}{1-q}\}}.$$

Proof of Key Proposition 3. Almost all the details are the same as in the proof of Key Proposition 1 from the Steinhaus case. Now we cannot translate the event $\mathcal{G}^{\mathrm{Rad}}(k,t)$ by shifting t to 0, as we did in the Steinhaus case, but there is no need to do so because we formulated our tilted probability estimates for $\tilde{\mathbb{P}}_t^{\mathrm{Rad}}$ for general $t\in\mathbb{R}\backslash\{0\}$. When we apply Proposition 6 to estimate $\tilde{\mathbb{P}}_t^{\mathrm{Rad}}(\mathcal{G}^{\mathrm{Rad}}(k,t))$, we must take $a=C\min\{\sqrt{\log\log x},\frac{1}{1-q}\}+D+2\log(D+1)$, rather than $a=C\min\{\sqrt{\log\log x},\frac{1}{1-q}\}+(B+1)+2\log(B+2)$ (and take $n=\lfloor\log\log x\rfloor-D-k$ and $\sigma=-\frac{k}{\log x}$ and $t_m=t(\lfloor\log\log x\rfloor-D-m)$ for all m). Then Proposition 6 and (3) imply that

$$\mathbb{E}\mathbf{1}_{\mathcal{G}^{\mathrm{Rad}}(k,t)}|F_k\left(1/2-\frac{k}{\log x}+it\right)|^2$$

$$\ll\min\left\{1,\frac{a}{\sqrt{\log\log x}}\right\}\mathbb{E}\left|F_k\left(1/2-\frac{k}{\log x}+it\right)\right|^2$$

$$\ll CD(t)\min\left\{1,\frac{1}{(1-q)\sqrt{\log\log x}}\right\}\frac{\log x}{e^k}.$$

Since we have $\int_{|t|\leqslant 1/2}D(t)\,dt\ll\int_{|t|\leqslant 1/2}\log(1/|t|)\,dt\ll 1$, this bound works the same as the corresponding bound (without $D(t)$) from the proof of Key Proposition 1, and so Key Proposition 3 follows. □

Proof of Key Proposition 4. The proof of Key Proposition 4 closely follows that of Key Proposition 2 from the Steinhaus case, until the final line where Σ_2 must be bounded. (Strictly speaking, one needs to be careful when applying the union bound in the proof because the upper limit $\lfloor\log\log x\rfloor-D(t)-1$ in the product now depends on t. However, $D(t)$ is constant on t intervals of the form

$e^{-(r+1)} \leqslant |t| < e^{-r}$, so one can split up first according to which such interval t lies in.) There we must invoke (4) to estimate each term $\mathbb{E}|I_l(1/2 - \frac{k}{\log x} + it(l))|^{-2}$ for $k \leqslant l \leqslant \log \log x - D - 1$, and the estimate this supplies is

$$\exp\left\{\sum_{x^{e^{-(l+2)}} < p \leqslant x^{e^{-(l+1)}}} \frac{1 + 2\cos(2t(l)\log p)}{p^{1-2k/\log x}} + O\left(\frac{1}{\sqrt{x^{e^{-(l+2)}}} e^{-(l+2)} \log x}\right)\right\},$$

which is not identical to the corresponding estimate for $\mathbb{E}|I_l(1/2 - \frac{k}{\log x} + it(l))|^2$ (which would not include the $2\cos(2t(l)\log p)$ term). However, since (using the approximation bound (5) on $|t - t(l)|$ and the fact that $l \leqslant \log \log x - D - 1$) we always have $|t| \ll |t(l)| \ll 1$, standard estimates for sums over primes (as in the proof of Lemma 5) show that

$$\sum_{x^{e^{-(l+2)}} < p \leqslant x^{e^{-(l+1)}}} \frac{2\cos(2t(l)\log p)}{p^{1-2k/\log x}} \ll \frac{1}{|t|e^{-l}\log x}.$$

Recalling that $l \leqslant \log \log x - D - 1$ and that $D = \lceil \log(1/|t|) \rceil + (B+1)$, we see these terms involving $2\cos(2t(l)\log p)$ give a negligible contribution. $\qquad\square$

Given Key Propositions 3 and 4, exactly the same argument as in the Steinhaus case in Section 4.1 confirms that, as we wanted,

$$\mathbb{E}\left(\int_{\substack{|t| \leqslant 1/2, \\ |t| > 1/\sqrt{\log \log x}}} \left|F_k\left(\frac{1}{2} - \frac{k}{\log x} + it\right)\right|^2 dt\right)^q$$

$$\ll \left(\frac{\log x}{e^k} \min\left\{1, \frac{1}{(1-q)\sqrt{\log \log x}}\right\}\right)^q.$$

To handle the integral over $|t - N| \leqslant 1/2$ for general $1 \leqslant |N| \leqslant (\log \log x)^2$, one follows the same argument as above but with $\log \log x - D - 1$ replaced by $\log \log x - \lceil 2\log \log 10N \rceil - (B+1) - 1$, say. This produces an extra factor $\log \log 10N$ in the analogue of Key Proposition 3, which is more than cancelled out by the prefactor $1/(|N|+1)^{q/4}$ attached to the N integral. $\qquad\square$

5. Proofs of the lower bounds in Theorems 1 and 2

Recall that $F(s)$ denotes the Euler product of $f(n)$ over x-smooth numbers. In view of Propositions 3 and 4, the lower bound parts of Theorems 1 and 2 will essentially follow if we can prove suitable lower bounds for $\|\int_{-1/2}^{1/2} |F(1/2 +$

$\frac{4V}{\log x} + it)|^2 \, dt \|_q^{1/2}$, where V is a large constant. (We also need upper bounds for quantities like $\| \int_{-1/2}^{1/2} |F(1/2 + \frac{2V}{\log x} + it)|^2 \, dt \|_q^{1/2}$, but those will follow directly from our work in Section 4.) As described in Section 1, we shall actually seek a lower bound for $\| \int_{\mathcal{L}} |F(1/2 + \frac{4V}{\log x} + it)|^2 \, dt \|_q^{1/2}$, where $\mathcal{L} \subseteq [-1/2, 1/2]$ is a suitable random subset that makes things work nicely.

When we proved Key Proposition 1 (or the Rademacher version, Key Proposition 3) during our work on upper bounds, we used Hölder's inequality to replace a restricted qth moment by a restricted first moment that we could estimate. This straightforward procedure is not available for lower bounds. However, we can write

$$\mathbb{E} \int_{\mathcal{L}} \left| F\left(\frac{1}{2} + \frac{4V}{\log x} + it \right) \right|^2 \, dt$$

$$= \mathbb{E} \left(\int_{\mathcal{L}} \left| F\left(\frac{1}{2} + \frac{4V}{\log x} + it \right) \right|^2 \, dt \right)^{\frac{q}{2-q}}$$

$$\times \left(\int_{\mathcal{L}} \left| F\left(\frac{1}{2} + \frac{4V}{\log x} + it \right) \right|^2 \, dt \right)^{\frac{2(1-q)}{(2-q)}},$$

so applying Hölder's inequality with exponents $2 - q$ and $(2-q)/(1-q)$, we obtain that $\mathbb{E} \int_{\mathcal{L}} |F(\frac{1}{2} + \frac{4V}{\log x} + it)|^2 \, dt$ is

$$\leqslant \left(\mathbb{E} \left(\int_{\mathcal{L}} \left| F\left(\frac{1}{2} + \frac{4V}{\log x} + it \right) \right|^2 \, dt \right)^q \right)^{\frac{1}{2-q}}$$

$$\times \left(\mathbb{E} \left(\int_{\mathcal{L}} \left| F\left(\frac{1}{2} + \frac{4V}{\log x} + it \right) \right|^2 \, dt \right)^2 \right)^{\frac{1-q}{2-q}}.$$

We will be able to lower bound $\mathbb{E} \int_{\mathcal{L}} |F(\frac{1}{2} + \frac{4V}{\log x} + it)|^2 \, dt$ using the Girsanov-type estimates we already proved in Section 3. The extra ingredient we require now is an upper bound for the second moment of the integral over $\mathcal{L}$, and to prove that, we must perform some further probabilistic preparations.

5.1. Further probabilistic calculations. In Lemma 1, we computed the expectation of the second power of one Steinhaus Euler product times an imaginary power of a shifted Euler product. We now require a variant of this.

LEMMA 6. *If f is a Steinhaus random multiplicative function, then for any real t, u, v, any real $400(1 + u^2 + v^2) \leqslant x \leqslant y$ and any real $\sigma \geqslant -1/\log y$, we have*

$$\mathbb{E} \prod_{x < p \leqslant y} \left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-(2+iu)} \left| 1 - \frac{f(p)}{p^{1/2+\sigma+it}} \right|^{-(2+iv)}$$

$$= \exp \left\{ \sum_{x < p \leqslant y} \frac{(1 + iu/2)^2 + (1 + iv/2)^2}{p^{1+2\sigma}} \right.$$

$$\left. + \sum_{x < p \leqslant y} \frac{(2 + iu)(2 + iv)\cos(t \log p)}{2p^{1+2\sigma}} + T(u, v) \right\},$$

where $T(u, v) = T_{x,y,\sigma,t}(u, v)$ satisfies $|T(u, v)| \ll \frac{1+|u|^3+|v|^3}{\sqrt{x}\log x}$ and its partial derivatives satisfy $\left| \frac{\partial T(u,v)}{\partial u} \right| \ll \frac{1+u^2+v^2}{\sqrt{x}\log x}$, $\left| \frac{\partial T(u,v)}{\partial v} \right| \ll \frac{1+u^2+v^2}{\sqrt{x}\log x}$ and $\left| \frac{\partial T(u,v)}{\partial u \partial v} \right| \ll \frac{1+|u|+|v|}{\sqrt{x}\log x}$.

In particular, Lemma 6 implies that for any real t and any $400 \leqslant x \leqslant y$ and any $\sigma \geqslant -1/\log y$, and for Steinhaus random multiplicative f, we have

$$\mathbb{E} \prod_{x < p \leqslant y} \left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-2} \left| 1 - \frac{f(p)}{p^{1/2+\sigma+it}} \right|^{-2}$$

$$= \exp \left\{ \sum_{x < p \leqslant y} \frac{2 + 2\cos(t \log p)}{p^{1+2\sigma}} + O\left(\frac{1}{\sqrt{x}\log x} \right) \right\}. \tag{6}$$

Proof of Lemma 6. The proof is an extension of the proof of Lemma 1. As there, we temporarily set $R_p(t) := -\Re \log(1 - \frac{f(p)}{p^{1/2+\sigma+it}})$, and then we may rewrite

$$\left| 1 - \frac{f(p)}{p^{1/2+\sigma}} \right|^{-(2+iu)} \left| 1 - \frac{f(p)}{p^{1/2+\sigma+it}} \right|^{-(2+iv)}$$

$$= \exp\{(2 + iu)R_p(0) + (2 + iv)R_p(t)\}$$

$$= 1 + \sum_{j=1}^{\infty} \frac{((2 + iu)R_p(0) + (2 + iv)R_p(t))^j}{j!}.$$

In the proof of Lemma 1, we used the Taylor expansion of the logarithm to determine that $\mathbb{E}R_p(t) = 0$. We also obtained that

$$\mathbb{E}R_p(t)^2 = \frac{1}{2p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}} \right)$$

and

$$\mathbb{E} R_p(0) R_p(t) = \frac{\cos(t \log p)}{2 p^{1+2\sigma}} + O\left(\frac{1}{p^{3/2+3\sigma}}\right),$$

and for $j \geqslant 3$, we have the trivial bound $|R_p(t)^j| \leqslant (\sum_{k=1}^{\infty} \frac{1}{p^{k(1/2+\sigma)}})^j = \frac{1}{(p^{1/2+\sigma}-1)^j}$. Furthermore, we noted there that for primes $y \geqslant p > x \geqslant 400(1 + u^2 + v^2)$ and for $\sigma \geqslant -\frac{1}{\log y}$, we have $\frac{1}{p^{1/2+\sigma}} = \frac{e^{-\sigma \log p}}{p^{1/2}} \leqslant \frac{e}{p^{1/2}}$, which now implies that $(4 + |u| + |v|)/p^{1/2+\sigma} \leqslant 3e/10$.

So putting things together, for such primes, we have

$$\mathbb{E}\left|1 - \frac{f(p)}{p^{1/2+\sigma}}\right|^{-(2+iu)} \left|1 - \frac{f(p)}{p^{1/2+\sigma+it}}\right|^{-(2+iv)}$$

$$= 1 + \frac{(2+iu)^2 \mathbb{E} R_p(0)^2 + 2(2 + iu)(2+iv)\mathbb{E} R_p(0) R_p(t) + (2 + iv)^2 \mathbb{E} R_p(t)^2}{2}$$

$$+ \mathbb{E} \sum_{j=3}^{\infty} \frac{((2 + iu) R_p(0) + (2 + iv) R_p(t))^j}{j!}$$

$$= 1 + \frac{(1 + \frac{iu}{2})^2 + (1 + \frac{iv}{2})^2}{p^{1+2\sigma}} + \frac{(2 + iu)(2 + iv) \cos(t \log p)}{2 p^{1+2\sigma}}$$

$$+ O\left(\sum_{j=3}^{\infty} \frac{(4 + |u| + |v|)^j}{j!(p^{1/2+\sigma} - 1)^j}\right)$$

$$= 1 + \frac{(1 + iu/2)^2 + (1 + iv/2)^2}{p^{1+2\sigma}}$$

$$+ \frac{(2 + iu)(2 + iv) \cos(t \log p)}{2 p^{1+2\sigma}} + D_p(u, v),$$

where $D_p(u, v)$ satisfies $|D_p(u, v)| \ll \frac{1+|u|^3+|v|^3}{p^{3/2+3\sigma}} \ll \frac{1+|u|^3+|v|^3}{p^{3/2}}$, and its partial derivatives satisfy $|\frac{\partial D_p(u,v)}{\partial u}| \ll \frac{1+u^2+v^2}{p^{3/2}}$, $|\frac{\partial D_p(u,v)}{\partial v}| \ll \frac{1+u^2+v^2}{p^{3/2}}$ and $|\frac{\partial D_p(u,v)}{\partial u \partial v}| \ll \frac{1+|u|+|v|}{p^{3/2}}$.

The conclusion of Lemma 6 now follows as in the proof of Lemma 1, using the independence of f on distinct primes and the standard Chebychev-type estimate $\sum_{p>x} 1/p^{3/2} \ll 1/(\sqrt{x} \log x)$. $\qquad\square$

We remark that in our applications of Lemma 1, we took the shift t to be rather small so that $\cos(t \log p) \approx 1$. In this setting, the expression $\sum_{x < p \leqslant y} \frac{1+iu \cos(t \log p) - u^2/4}{p^{1+2\sigma}}$ from Lemma 1 is approximately the expression

$\sum_{x<p\leqslant y} \frac{(1+iu/2)^2}{p^{1+2\sigma}}$ in Lemma 6. When we come to apply Lemma 6, we will take the shift t there to be somewhat large compared with x so that the sum $\sum_{x<p\leqslant y} \frac{(2+iu)(2+iv)\cos(t\log p)}{2p^{1+2\sigma}}$ is small. The important point is that this regime corresponds to the product and the shifted product in Lemma 6 behaving fairly independently.

Next, we want to adapt the Girsanov-type calculations from Section 3.2 to a 'tilted' probability measure weighted by two Euler products. Thus, for each $t \in \mathbb{R}$ and for large x and $-1/100 \leqslant \sigma \leqslant 1/100$, let us define a probability measure $\tilde{\mathbb{P}}_t^{\text{St},(2)} = \tilde{\mathbb{P}}_{x,\sigma,t}^{\text{St},(2)}$ by setting

$$\tilde{\mathbb{P}}_t^{\text{St},(2)}(A) := \frac{\mathbb{E}\mathbf{1}_A \prod_{p\leqslant x^{1/e}} \left|1 - \frac{f(p)}{p^{1/2+\sigma}}\right|^{-2} \left|1 - \frac{f(p)}{p^{1/2+\sigma+it}}\right|^{-2}}{\mathbb{E} \prod_{p\leqslant x^{1/e}} \left|1 - \frac{f(p)}{p^{1/2+\sigma}}\right|^{-2} \left|1 - \frac{f(p)}{p^{1/2+\sigma+it}}\right|^{-2}}$$

for each event A, where f is a Steinhaus random multiplicative function. We write $\tilde{\mathbb{E}}_t^{\text{St},(2)}$ to denote expectation with respect to this measure. The superscript (2) in this notation reflects the fact that we are weighting by two products and is designed to be distinctive from the notation used in the one product Rademacher case in Section 3.3.

As in the one-dimensional case, the exact range of p in the definition of the measure $\tilde{\mathbb{P}}_t^{\text{St},(2)}$ does not matter too much because if the event A does not involve a particular prime, then the expectation of that part of the products will factor out and cancel between the numerator and denominator. *We emphasize this point because, in our later calculations, we will extract various parts of the products and we must be clear that our results about $\tilde{\mathbb{P}}_t^{St,(2)}(A)$ will still be applicable, for appropriate A.*

As before, we shall let $(l_j)_{j=1}^n$ denote a *strictly decreasing* sequence of nonnegative integers, with $l_1 \leqslant \log\log x - 2$, and define a corresponding increasing sequence of real numbers $(x_j)_{j=1}^n$ by setting $x_j := x^{e^{-(l_j+1)}}$. And for each $l \in \mathbb{N} \cup \{0\}$, we set $I_l(s) := \prod_{x^{e^{-(l+2)}} < p \leqslant x^{e^{-(l+1)}}} (1 - \frac{f(p)}{p^s})^{-1}$, the lth increment of the Euler product corresponding to Steinhaus f.

Using Lemma 6 (which will serve as a two-dimensional characteristic function calculation under the measure $\tilde{\mathbb{P}}_t^{\text{St},(2)}$) and a two-dimensional version of the Berry–Esseen inequality, we can prove the following result, which the reader may compare with the one-dimensional case in Lemma 3.

LEMMA 7. *Let the situation be as above, with the restriction that $|t| \leqslant 1$. Suppose that $x_1 \geqslant e^{C/|t|^2}$ is sufficiently large and that $|\sigma| \leqslant 1/\log x_n$. Suppose further that $(u_j)_{j=1}^n$ and $(v_j)_{j=1}^n$ are any sequences of real numbers satisfying*

$$|u_j|, |v_j| \leqslant (1/40)(\log x_j)^{1/4} + 2 \quad \forall 1 \leqslant j \leqslant n.$$

Then we have

$$\tilde{\mathbb{P}}_t^{St,(2)}\left(u_j \leqslant \log\left|I_{l_j}\left(\frac{1}{2}+\sigma\right)\right| \leqslant u_j + \frac{1}{j^2},\right.$$

$$v_j \leqslant \log\left|I_{l_j}\left(\frac{1}{2}+\sigma+it\right)\right| \leqslant v_j + \frac{1}{j^2}, \forall j \leqslant n\right)$$

$$= \left(1 + O\left(\frac{1}{x_1^{1/100}}\right)\right)$$

$$\times \mathbb{P}\left(u_j \leqslant N_j^1 \leqslant u_j + \frac{1}{j^2}, \text{ and } v_j \leqslant N_j^2 \leqslant v_j + \frac{1}{j^2}, \forall 1 \leqslant j \leqslant n\right),$$

where $(N_j^1, N_j^2)_{j=1}^n$ is a sequence of independent bivariate Gaussian random vectors, and the components N_j^1, N_j^2 have mean $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1+\cos(t \log p)}{p^{1+2\sigma}}$, variance $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}}$ and covariance $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t \log p)}{2p^{1+2\sigma}}$.

Furthermore, we also have the alternative estimate

$$\tilde{\mathbb{P}}_t^{St,(2)}\left(u_j \leqslant \log\left|I_{l_j}\left(\frac{1}{2}+\sigma\right)\right| \leqslant u_j + \frac{1}{j^2},\right.$$

$$v_j \leqslant \log\left|I_{l_j}\left(\frac{1}{2}+\sigma+it\right)\right| \leqslant v_j + \frac{1}{j^2}, \forall j \leqslant n\right)$$

$$= \left(1 + O\left(\frac{1}{\sqrt{C}}\right)\right)\mathbb{P}\left(u_j \leqslant N_j^1 \leqslant u_j + \frac{1}{j^2} \forall j \leqslant n\right)$$

$$\cdot \mathbb{P}\left(v_j \leqslant N_j^2 \leqslant v_j + \frac{1}{j^2} \forall j \leqslant n\right),$$

in other words, we may replace the covariance of N_j^1, N_j^2 by zero.

We should perhaps comment on some of the assumptions in Lemma 7, as compared with Lemmas 3 and 5. Note that the allowed range of u_j, v_j here is a multiple of $(\log x_j)^{1/4}$, as opposed to our previous $\sqrt{\log x_j}$. The assumption that $x_1 \geqslant e^{C/|t|^2}$ is also stronger than previous assumptions of the shape $x_1 \geqslant e^{C/|t|}$. These conditions are used to deduce the second part of Lemma 7, that we may treat N_j^1 and N_j^2 as independent. Some tradeoff between the conditions is possible, so if one restricted u_j, v_j further one could allow x_1 to be smaller. But in our application, we will set things up so both conditions are anyway satisfied with room to spare.

Proof of Lemma 7. By independence, both the probability on the left and the ones on the right factor as a product over j. So, bearing in mind that we have

$x_j \geqslant x_1^{e^{j-1}}$ for all j, it will suffice to prove the equalities

$$\tilde{\mathbb{P}}_t^{\text{St},(2)} \left(u_j \leqslant \log \left| I_{l_j} \left(\frac{1}{2} + \sigma \right) \right| \leqslant u_j + \frac{1}{j^2}, \right.$$

$$\text{and } v_j \leqslant \log \left| I_{l_j} \left(\frac{1}{2} + \sigma + it \right) \right| \leqslant v_j + \frac{1}{j^2} \right)$$

$$= \left(1 + O \left(\frac{1}{x_j^{1/100}} \right) \right)$$

$$\times \mathbb{P} \left(u_j \leqslant N_j^1 \leqslant u_j + \frac{1}{j^2}, \text{ and } v_j \leqslant N_j^2 \leqslant v_j + \frac{1}{j^2} \right)$$

$$= \left(1 + O \left(\frac{1}{\sqrt{Ce^j}} \right) \right)$$

$$\times \mathbb{P} \left(u_j \leqslant N_j^1 \leqslant u_j + \frac{1}{j^2} \right) \cdot \mathbb{P} \left(v_j \leqslant N_j^2 \leqslant v_j + \frac{1}{j^2} \right)$$

for all $1 \leqslant j \leqslant n$.

By two applications of Lemma 6, for any $|u|, |v| \leqslant x_j^{1/10}$ (say), the characteristic function $\tilde{\mathbb{E}}_t^{\text{St},(2)} e^{iu \log |I_{l_j}(1/2+\sigma)| + iv \log |I_{l_j}(1/2+\sigma+it)|}$ is

$$= \exp \left\{ \sum_{x_j^{1/e} < p \leqslant x_j} \frac{iu - \frac{u^2}{4} + iv - \frac{v^2}{4}}{p^{1+2\sigma}} \right.$$

$$\left. + \sum_{x_j^{1/e} < p \leqslant x_j} \frac{(2iu + 2iv - uv) \cos(t \log p)}{2p^{1+2\sigma}} + T(u,v) - T(0,0) \right\}.$$

Without the error term $T(u,v) - T(0,0)$, a standard calculation shows that this would be the characteristic function of the pair (N_j^1, N_j^2), where N_j^1, N_j^2 are Gaussian random variables each having mean $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1 + \cos(t \log p)}{p^{1+2\sigma}}$ and variance $\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}}$, and with covariance $\mathbb{E} N_j^1 N_j^2 - \mathbb{E} N_j^1 \mathbb{E} N_j^2 = \sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t \log p)}{2p^{1+2\sigma}}$. Before proceeding further, we record a few calculations we will use later. First, under our conditions $|t| \leqslant 1$ and $x_j \geqslant x_1 \geqslant e^{C/|t|^2}$ (in fact, it would suffice to have $x_1 \geqslant e^{C/|t|}$ at this stage) and $|\sigma| \leqslant 1/\log x_n$, we have by standard prime number estimates (as in the proofs of Lemmas 4 and 5) that

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t \log p)}{p^{1+2\sigma}} \ll \frac{1}{|t| \log x_j} \ll \frac{1}{e^j |t| \log x_1} \ll \frac{1}{Ce^j}$$

as well as

$$\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{p^{1+2\sigma}} = 1 + O\left(\frac{1}{\log x_j} + \frac{\log x_j}{\log x_n}\right) = 1 + O(C^{-1}e^{-j} + e^{-(n-j)}).$$

Second, by Lemma 6 the 'error' terms $T(u, v)$ in our characteristic functions satisfy $|T(u, v) - T(0, 0)| \ll \frac{1+|u|^3+|v|^3}{\sqrt{x_j^{1/e}} \log x_j}$ as well as $|T(u, 0) - T(0, 0)| \ll \frac{|u|+|u|^3}{\sqrt{x_j^{1/e}} \log x_j}$ and $|T(0, v) - T(0, 0)| \ll \frac{|v|+|v|^3}{\sqrt{x_j^{1/e}} \log x_j}$, and

$$|T(u, v) - T(u, 0) - T(0, v) + T(0, 0)| = \left|\int_0^u \int_0^v \frac{\partial T(u, v)}{\partial u \partial v} \, du \, dv\right|$$
$$\ll \frac{|u||v|(1 + |u| + |v|)}{\sqrt{x_j^{1/e}} \log x_j}.$$

Now by the two-dimensional Berry–Esseen inequality (see Sadikova's paper [24]), we have

$$\left|\tilde{\mathbb{P}}_t^{\mathrm{St},(2)}\left(u_j \leqslant \log\left|I_{l_j}\left(\frac{1}{2}+\sigma\right)\right| \leqslant u_j + \frac{1}{j^2},\right.\right.$$
$$\left.\text{and } v_j \leqslant \log\left|I_{l_j}\left(\frac{1}{2}+\sigma+it\right)\right| \leqslant v_j + \frac{1}{j^2}\right)$$
$$\left.- \mathbb{P}\left(u_j \leqslant N_j^1 \leqslant u_j + \frac{1}{j^2}, \text{ and } v_j \leqslant N_j^2 \leqslant v_j + \frac{1}{j^2}\right)\right|$$
$$\ll \int_{-x_j^{1/50}}^{x_j^{1/50}} \int_{-x_j^{1/50}}^{x_j^{1/50}} \left|\frac{\Delta(u, v)}{uv}\right| \, du \, dv$$
$$+ \int_{-x_j^{1/50}}^{x_j^{1/50}} \left|\frac{\tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iu \log |I_{l_j}(1/2+\sigma)|} - \mathbb{E}e^{iuN_j^1}}{u}\right| \, du$$
$$+ \int_{-x_j^{1/50}}^{x_j^{1/50}} \left|\frac{\tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iv \log |I_{l_j}(1/2+\sigma+it)|} - \mathbb{E}e^{ivN_j^2}}{v}\right| \, dv + \frac{1}{x_j^{1/50}},$$

where

$$\Delta(u, v) := \tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iu \log |I_{l_j}(1/2+\sigma)|+iv \log |I_{l_j}(1/2+\sigma+it)|} - \mathbb{E}e^{iuN_j^1+ivN_j^2}$$
$$- \tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iu \log |I_{l_j}(1/2+\sigma)|} \tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iv \log |I_{l_j}(1/2+\sigma+it)|} + \mathbb{E}e^{iuN_j^1} \mathbb{E}e^{ivN_j^2}.$$

Our general expression for the two-dimensional characteristic function (now with $v = 0$) implies that $\tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iu \log |I_{l_j}(1/2+\sigma)|}$ is equal to $e^{T(u,0)-T(0,0)} \mathbb{E}e^{iuN_j^1}$,

so the second integral here is

$$\leqslant \int_{-x_j^{1/50}}^{x_j^{1/50}} \left| \frac{e^{T(u,0)-T(0,0)} - 1}{u} \right| du \ll \frac{1}{\sqrt{x_j^{1/e} \log x_j}} \int_{-x_j^{1/50}}^{x_j^{1/50}} (1 + u^2)\, du \ll \frac{1}{x_j^{1/50}},$$

and the third integral may be bounded exactly similarly. To bound the first (double) integral, we split into the ranges $|uv| \leqslant 1$ and $|uv| > 1$. On the latter range, to bound $\Delta(u, v)$, we can simply use that (by Lemma 6, as noted earlier in the proof)

$$\tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iu \log |I_{l_j}(1/2+\sigma)| + iv \log |I_{l_j}(1/2+\sigma+it)|} - \mathbb{E} e^{iuN_j^1 + ivN_j^2}$$
$$= \mathbb{E} e^{iuN_j^1 + ivN_j^2} (e^{T(u,v)-T(0,0)} - 1)$$
$$\ll \frac{1 + |u|^3 + |v|^3}{\sqrt{x_j^{1/e} \log x_j}}$$

and similarly that

$$\tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iu \log |I_{l_j}(1/2+\sigma)|} \tilde{\mathbb{E}}_t^{\mathrm{St},(2)} e^{iv \log |I_{l_j}(1/2+\sigma+it)|} - \mathbb{E} e^{iuN_j^1} \mathbb{E} e^{ivN_j^2}$$
$$= \mathbb{E} e^{iuN_j^1} \mathbb{E} e^{ivN_j^2} (e^{T(u,0)-T(0,0)+T(0,v)-T(0,0)} - 1) \ll \frac{1 + |u|^3 + |v|^3}{\sqrt{x_j^{1/e} \log x_j}}.$$

To handle the more delicate range where $|uv| \leqslant 1$, we note that

$$\Delta(u, v) = \mathbb{E} e^{iuN_j^1} \mathbb{E} e^{ivN_j^2}$$
$$\times \left(e^{-uv \sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t \log p)}{2p^{1+2\sigma}} + T(u,v) - T(0,0)} - e^{-uv \sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t \log p)}{2p^{1+2\sigma}}} \right.$$
$$\left. - e^{T(u,0)-T(0,0)+T(0,v)-T(0,0)} + 1 \right).$$

We can write $e^{T(u,0)-T(0,0)+T(0,v)-T(0,0)} = e^{T(u,v)-T(0,0)} e^{-(T(u,v)-T(u,0)-T(0,v)+T(0,0))}$, which is $= e^{T(u,v)-T(0,0)} + O\left(\frac{|u||v|(1+|u|+|v|)}{\sqrt{x_j^{1/e} \log x_j}} \right)$, using our earlier estimation that

$$|T(u, v) - T(u, 0) - T(0, v) + T(0, 0)| \ll \frac{|u||v|(1 + |u| + |v|)}{\sqrt{x_j^{1/e} \log x_j}}.$$

So when $|uv| \leqslant 1$, we have that $|\Delta(u, v)|$ has order at most

$$\left|e^{T(u,v)-T(0,0)} - 1\right\| e^{-uv\sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t\log p)}{2p^{1+2\sigma}}} - 1\right| + \frac{|uv|(1+|u|+|v|)}{\sqrt{x_j^{1/e}\log x_j}}$$

$$\ll \left|e^{T(u,v)-T(0,0)} - 1\right\| uv| + \frac{|uv|(1+|u|+|v|)}{\sqrt{x_j^{1/e}\log x_j}} \ll \frac{|uv|(1+|u|^3+|v|^3)}{\sqrt{x_j^{1/e}\log x_j}},$$

and we obtain overall that

$$\int_{-x_j^{1/50}}^{x_j^{1/50}} \int_{-x_j^{1/50}}^{x_j^{1/50}} \left|\frac{\Delta(u,v)}{uv}\right| \, du \, dv \ll \frac{x_j^{1/10}}{\sqrt{x_j^{1/e}\log x_j}} \ll \frac{1}{x_j^{1/20}} \ll \frac{1}{x_j^{1/50}}.$$

To finish the proof, let us (for concision) temporarily set

$$\mu_j = \sum_{x_j^{1/e} < p \leqslant x_j} \frac{1+\cos(t\log p)}{p^{1+2\sigma}},$$

$$\sigma_j := \sqrt{\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}}}$$

and

$$r_j = \frac{\sum_{x_j^{1/e} < p \leqslant x_j} \frac{\cos(t\log p)}{2p^{1+2\sigma}}}{\sum_{x_j^{1/e} < p \leqslant x_j} \frac{1}{2p^{1+2\sigma}}},$$

and recall we calculated earlier that $\sigma_j \asymp 1$ and $\mu_j \asymp 1$ and $r_j \ll 1/(|t|\log x_j) \ll 1/C$. Then we may note that

$$\mathbb{P}\left(u_j \leqslant N_j^1 \leqslant u_j + \frac{1}{j^2}, \text{ and } v_j \leqslant N_j^2 \leqslant v_j + \frac{1}{j^2}\right)$$

$$= \mathbb{P}\left(\frac{u_j - \mu_j}{\sigma_j} \leqslant \frac{N_j^1 - \mu_j}{\sigma_j} \leqslant \frac{u_j + 1/j^2 - \mu_j}{\sigma_j},\right.$$

$$\left.\text{and } \frac{v_j - \mu_j}{\sigma_j} \leqslant \frac{N_j^2 - \mu_j}{\sigma_j} \leqslant \frac{v_j + 1/j^2 - \mu_j}{\sigma_j}\right)$$

$$= \int_{(u_j-\mu_j)/\sigma_j}^{(u_j-\mu_j+1/j^2)/\sigma_j} \int_{(v_j-\mu_j)/\sigma_j}^{(v_j-\mu_j+1/j^2)/\sigma_j} \frac{1}{2\pi\sqrt{1-r_j^2}} e^{-(x^2-2r_jxy+y^2)/2(1-r_j^2)} \, dx \, dy.$$

In particular, since we have $|u_j|, |v_j| \leqslant (\log x_j)^{1/4}$, we see this probability is $\gg \frac{1}{j^4} e^{-O(\sqrt{\log x_j})}$. Thus, as in the one-dimensional case in Lemma 3, we may

replace the absolute error term $1/x_j^{1/50}$ from the Berry–Esseen theorem by a multiplier $1 + O(1/x_j^{1/100})$, say. This gives the first estimate in Lemma 7. To obtain the second estimate in Lemma 7, we simply note that

$$
\frac{e^{-(x^2-2r_jxy+y^2)/2(1-r_j^2)}}{2\pi\sqrt{1-r_j^2}} = \frac{e^{-(x^2+y^2)/2}}{2\pi\sqrt{1-r_j^2}} e^{O(r_j\sqrt{\log x_j})}
$$

$$
= \frac{e^{-(x^2+y^2)/2}}{2\pi}\left(1 + O\left(\frac{1}{|t|\sqrt{\log x_j}}\right)\right),
$$

and here $\dfrac{1}{|t|\sqrt{\log x_j}} \ll \dfrac{1}{|t|\sqrt{e^j\log x_1}} \ll \dfrac{1}{\sqrt{Ce^j}}$ because of our assumption that $x_1 \geqslant e^{C/|t|^2}$. $\qquad\square$

Now we can swiftly deduce the following result, which is what we shall need for our lower bound proof and is a two-dimensional analogue of Proposition 5.

PROPOSITION 7. *There is a large natural number B such that the following is true.*

Let $t \in \mathbb{R}$ satisfy $|t| \leqslant 1$ and let $D \geqslant 2\log(1/|t|) + (B+1)$ be any natural number. Let $n \leqslant \log\log x - D$ be large, and define the decreasing sequence $(l_j)_{j=1}^n$ of nonnegative integers by $l_j := \lfloor\log\log x\rfloor - D - j$. Suppose also that $|\sigma| \leqslant \frac{1}{e^{D+n}}$.

Then uniformly for any large a and any function $h(j)$ satisfying $|h(j)| \leqslant 10\log j$, and with $I_l(s)$ denoting the increments of the Euler product corresponding to a Steinhaus random multiplicative function (as before), we have

$$
\tilde{\mathbb{P}}_t^{St,(2)}\left(-a - Bj \leqslant \sum_{m=1}^j \log\left|I_{l_m}\left(\frac{1}{2} + \sigma\right)\right|, \sum_{m=1}^j \log\left|I_{l_m}\left(\frac{1}{2} + \sigma + it\right)\right|\right.
$$

$$
\left. \leqslant a + j + h(j) \;\forall j \leqslant n\right)
$$

$$
\asymp \min\left\{1, \frac{a}{\sqrt{n}}\right\}^2.
$$

Proof of Proposition 7. Under the hypotheses of Proposition 7, we have

$$
\log x_1 = e^{-(l_1+1)}\log x \geqslant \frac{e^{B+1}}{|t|^2} \quad\text{and}\quad |\sigma| \leqslant \frac{1}{e^{D+n}} \leqslant \frac{e^{l_n+1}}{\log x} = \frac{1}{\log x_n}.
$$

This means the hypotheses of Lemma 7 will be satisfied, provided the constant B is fixed sufficiently large.

So using Lemma 7, together with a slicing argument exactly as in the deduction of Lemma 4, we may show the following: provided that $-(1/80)(\log x_j)^{1/4} \leqslant u_j \leqslant v_j \leqslant (1/80)(\log x_j)^{1/4}$ for all $1 \leqslant j \leqslant n$, we have

$$
\tilde{\mathbb{P}}_t^{\mathrm{St},(2)} \left(u_j \leqslant \sum_{m=1}^{j} \log \left| I_{l_m} \left(\frac{1}{2} + \sigma \right) \right|, \sum_{m=1}^{j} \log \left| I_{l_m} \left(\frac{1}{2} + \sigma + it \right) \right| \right.
$$
$$
\left. \leqslant v_j, \ \forall 1 \leqslant j \leqslant n \right)
$$
$$
\asymp \mathbb{P} \left(u_j - j + O(1) \leqslant \sum_{m=1}^{j} G_m \leqslant v_j - j + O(1) \ \forall 1 \leqslant j \leqslant n \right)^2,
$$

where $(G_m)_{m=1}^{j}$ are a sequence of independent Gaussian random variables, each having mean zero and variance $\sum_{x_m^{1/e} < p \leqslant x_m} \frac{1}{2 p^{1+2\sigma}}$. The conclusion of Proposition 7 now follows exactly as in the one-dimensional case in Proposition 5. $\square$

5.2. The lower bound in the Steinhaus case.

Let B be the large fixed natural number from Proposition 5 (which we may assume, without loss of generality, is the same as the number B from Proposition 7). For each $t \in \mathbb{R}$, let $L(t) = L_{x,q,V}(t)$ denote the event that

$$
\left(\frac{\log x}{e^{j+1}} \right)^{-B} e^{-\min\{\sqrt{\log\log x}, 1/(1-q)\}} \leqslant \prod_{l=j}^{\lfloor \log\log x \rfloor - B - 2} \left| I_l \left(1/2 + \frac{4V}{\log x} + it \right) \right|
$$
$$
\leqslant \frac{\log x}{e^{j+1}} e^{\min\{\sqrt{\log\log x}, 1/(1-q)\} - 2\log\log(\frac{\log x}{e^{j+1}})}
$$

for all $\lfloor \log V \rfloor + 3 \leqslant j \leqslant \log\log x - B - 2$. Here the quantity V will ultimately be fixed as another large constant, but initially, we allow any $1 \leqslant V \leqslant (\log x)^{1/100}$, say. Let $\mathcal{L}$ denote the random subset of points $|t| \leqslant 1/2$ at which $L(t)$ occurs.

The following is the key restricted second moment estimate that we shall need.

KEY PROPOSITION 5. *With the foregoing notation and uniformly for all large x and $2/3 \leqslant q \leqslant 1$ and $1 \leqslant V \leqslant (\log x)^{1/100}$, we have*

$$\mathbb{E}\left(\int_{\mathcal{L}}\left|F\left(1/2+\frac{4V}{\log x}+it\right)\right|^2 dt\right)^2$$

$$\ll e^{2\min\{\sqrt{\log\log x},\,1/(1-q)\}}\left(\frac{\log x}{V(1+(1-q)\sqrt{\log\log x})}\right)^2.$$

We remark that the factor $e^{2\min\{\sqrt{\log\log x},\,1/(1-q)\}}$ here, which may look rather alarming, will in fact be harmless because, when raised to the power $1-q$, it becomes $\ll 1$. This should become clear imminently when we deduce the Theorem 1 lower bound.

Proof of the lower bound in Theorem 1, assuming Key Proposition 5. As argued at the beginning of Section 5, Hölder's inequality implies that

$$\mathbb{E}\left(\int_{\mathcal{L}}\left|F\left(1/2+\frac{4V}{\log x}+it\right)\right|^2 dt\right)^q$$

$$\geqslant \frac{\left(\mathbb{E}\int_{\mathcal{L}}|F(1/2+\frac{4V}{\log x}+it)|^2 dt\right)^{2-q}}{\left(\mathbb{E}\left(\int_{\mathcal{L}}|F(1/2+\frac{4V}{\log x}+it)|^2 dt\right)^2\right)^{1-q}}.$$

In the Steinhaus case, translation invariance in law implies that we have a simplified expression for the numerator, namely

$$\left(\int_{-1/2}^{1/2}\mathbb{E}\mathbf{1}_{L(t)}\left|F\left(1/2+\frac{4V}{\log x}+it\right)\right|^2 dt\right)^{2-q}$$

$$=\left(\mathbb{E}\mathbf{1}_{L(0)}\left|F\left(1/2+\frac{4V}{\log x}\right)\right|^2\right)^{2-q}$$

$$=\left(\tilde{\mathbb{P}}(L(0))\cdot\mathbb{E}\left|F\left(1/2+\frac{4V}{\log x}\right)\right|^2\right)^{2-q}.$$

We can apply Proposition 5 from Section 3.2 here, taking $n=\lfloor\log\log x\rfloor-(B+1)-(\lfloor\log V\rfloor+3)$ and $\sigma=\frac{4V}{\log x}$ and $t_j\equiv 0$, and $a=\min\{\sqrt{\log\log x},\,1/(1-q)\}+O(1)$ and $h(j)=-2\log j$. Using also (2) and the standard Mertens prime number estimates, which imply that $\mathbb{E}|F(1/2+\frac{4V}{\log x})|^2$ is $\gg\exp\{\sum_{p\leqslant x}\frac{1}{p^{1+8V/\log x}}\}=\exp\{\sum_{p\leqslant x^{1/V}}\frac{1}{p^{1+8V/\log x}}+O(1)\}\gg(\log x)/V$, we

deduce our numerator is

$$\gg \left(\frac{1}{1 + (1-q)\sqrt{\log\log x}} \mathbb{E} \left| F\left(1/2 + \frac{4V}{\log x}\right) \right|^2 \right)^{2-q}$$

$$\gg \left(\frac{\log x}{V(1 + (1-q)\sqrt{\log\log x})} \right)^{2-q}.$$

Inserting the estimate from Key Proposition 5 to upper bound the denominator, and taking $2q$th roots, we deduce overall that

$$\left\| \int_{-1/2}^{1/2} \left| F\left(1/2 + \frac{4V}{\log x} + it\right) \right|^2 dt \right\|_q^{1/2} \gg \sqrt{\frac{\log x}{V(1 + (1-q)\sqrt{\log\log x})}}.$$

Meanwhile, if we argue as we did in Section 4.1 when proving the upper bound part of Theorem 1 (specifically when handling $F_k(s)$ with $k = \lfloor \log V \rfloor$), we have

$$\left\| \int_{-1/2}^{1/2} \left| F\left(1/2 + \frac{2V}{\log x} + it\right) \right|^2 dt \right\|_q^{1/2} \ll \sqrt{\frac{\log x}{V(1 + (1-q)\sqrt{\log\log x})}}.$$

Substituting these two bounds into Proposition 3 and choosing V to be a sufficiently large fixed constant that the term C/e^V there kills off the effect of the implicit constants, we obtain (with V now fixed)

$$\left\| \sum_{n \leqslant x} f(n) \right\|_{2q} \gg \sqrt{\frac{x}{V(1 + (1-q)\sqrt{\log\log x})}} \gg \sqrt{\frac{x}{1 + (1-q)\sqrt{\log\log x}}}.$$

This is the lower bound claimed in Theorem 1. $\qquad\square$

Proof of Key Proposition 5. Expanding out and recalling the definition of $\mathcal{L}$, we find $\mathbb{E}(\int_{\mathcal{L}} |F(1/2 + \frac{4V}{\log x} + it)|^2 dt)^2$ is

$$= \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \mathbb{E} \mathbf{1}_{L(t)} \left| F\left(1/2 + \frac{4V}{\log x} + it\right) \right|^2$$

$$\times \mathbf{1}_{L(s)} \left| F\left(1/2 + \frac{4V}{\log x} + is\right) \right|^2 ds\,dt.$$

In the Steinhaus case, we can use translation invariance in law to simplify this by shifting t to 0 and replacing s by $s - t$. This yields that

$$\mathbb{E}\left(\int_{\mathcal{L}}\left|F\left(1/2+\frac{4V}{\log x}+it\right)\right|^2 dt\right)^2$$

$$\leqslant \int_{-1}^{1}\mathbb{E}\mathbf{1}_{L(0)}\left|F\left(1/2+\frac{4V}{\log x}\right)\right|^2 \mathbf{1}_{L(t)}\left|F\left(1/2+\frac{4V}{\log x}+it\right)\right|^2 dt.$$

Now the point here is that given a shift of size t in the integral, we expect the parts of the two Euler products on primes roughly $\leqslant e^{1/|t|}$ (or on primes $\leqslant x$, if $|t| < 1/\log x$) to behave in almost the same way and the parts on larger primes to behave almost independently. So let us temporarily set $M := \max\{\lfloor\log V\rfloor + 3, \lfloor\log(|t|\log x)\rfloor\}$. Then our strategy is to use the fact that, by definition of $L(t)$, when the event $L(t)$ occurs, we have

$$\prod_{l=M}^{\lfloor\log\log x\rfloor-B-2}\left|I_l\left(\frac{1}{2}+\frac{4V}{\log x}+it\right)\right|^2$$

$$\ll \left(\min\left\{\frac{\log x}{V(\log\log x)^2}, \frac{1}{|t|\log^2(2/|t|)}\right\}e^{\min\{\sqrt{\log\log x}, \frac{1}{1-q}\}}\right)^2, \quad (7)$$

and then use our probabilistic estimates to bound the remaining part

$$\mathbb{E}\mathbf{1}_{L(0)}\left|F\left(1/2+\frac{4V}{\log x}\right)\right|^2 \mathbf{1}_{L(t)}\frac{|F(1/2+\frac{4V}{\log x}+it)|^2}{\prod_{l=M}^{\lfloor\log\log x\rfloor-B-2}|I_l(1/2+\frac{4V}{\log x}+it)|^2} \quad (8)$$

of the integrand.

It turns out that when $|t| \leqslant 1/(\log x)^{1/3}$, say (so in particular $\log(2/|t|) \gg \log\log x$), we can afford to take a crude approach thanks to the saving $1/(\log\log x)^4$ in the minimum in (7), and now throw away the indicator functions $\mathbf{1}_{L(0)}, \mathbf{1}_{L(t)}$ from (8). Having done this, using the independence of $f(p)$ on distinct primes, we find (8) is

$$\leqslant \mathbb{E}\prod_{l=M}^{\lfloor\log\log x\rfloor-B-2}\left|I_l\left(1/2+\frac{4V}{\log x}\right)\right|^2$$

$$\times \mathbb{E}\frac{|F(1/2+\frac{4V}{\log x})|^2|F(1/2+\frac{4V}{\log x}+it)|^2}{\prod_{l=M}^{\lfloor\log\log x\rfloor-B-2}|I_l(1/2+\frac{4V}{\log x})|^2|I_l(1/2+\frac{4V}{\log x}+it)|^2}.$$

Now using (6), the second expectation here is equal to

$$\exp\left\{\sum_{p\leqslant x^{e^{-(\lfloor\log\log x\rfloor-B)}}}\frac{2+2\cos(t\log p)}{p^{1+8V/\log x}} + \sum_{x^{e^{-(M+1)}}<p\leqslant x}\frac{2+2\cos(t\log p)}{p^{1+8V/\log x}} + O(1)\right\},$$

and remembering that B is a fixed constant and using standard Mertens and Chebychev-type estimates for sums over primes, this is

$$\ll \exp\left\{\sum_{\min\{x^{1/V}, e^{1/|t|}\} < p \leqslant x} \frac{2 + 2\cos(t\log p)}{p^{1+8V/\log x}}\right\}$$

$$\ll \exp\left\{\sum_{\min\{x^{1/V}, e^{1/|t|}\} < p \leqslant x^{1/V}} \frac{2 + 2\cos(t\log p)}{p}\right\}.$$

Again, since we may assume here that $p > e^{1/|t|}$, standard estimates for sums over primes (as in the proof of Lemma 5) show the overall contribution from the $\cos(t\log p)$ sum is $\ll 1$, so we finally have a bound $\ll \exp\{\sum_{\min\{x^{1/V}, e^{1/|t|}\} < p \leqslant x^{1/V}} \frac{2}{p}\} \ll (\max\{1, \frac{|t|\log x}{V}\})^2$. Meanwhile, (2) implies that the first expectation $\mathbb{E}\prod_{l=M}^{\lfloor\log\log x\rfloor - B - 2} |I_l(1/2 + \frac{4V}{\log x})|^2$ is equal to $\exp\{\sum_{p \leqslant \min\{x^{1/V}, e^{1/|t|}\}} \frac{1}{p^{1+8V/\log x}} + O(1)\} \ll \min\{\frac{\log x}{V}, \frac{1}{|t|}\}$. So putting together (7) with our above upper bounds for (8), we obtain

$$\int_{|t| \leqslant 1/(\log x)^{1/3}} \mathbb{E}\mathbf{1}_{L(0)}\left|F\left(1/2 + \frac{4V}{\log x}\right)\right|^2 \mathbf{1}_{L(t)}\left|F\left(1/2 + \frac{4V}{\log x} + it\right)\right|^2 dt$$

$$\ll \frac{e^{2\min\{\sqrt{\log\log x}, 1/(1-q)\}}}{(\log\log x)^4}$$

$$\times \int_{|t| \leqslant 1/(\log x)^{1/3}} \left(\min\left\{\frac{\log x}{V}, \frac{1}{|t|}\right\}\right)^3 \left(\max\left\{1, \frac{|t|\log x}{V}\right\}\right)^2 dt$$

$$\ll \frac{e^{2\min\{\sqrt{\log\log x}, \frac{1}{1-q}\}}\log^2 x}{V^2(\log\log x)^4} \int_{|t| \leqslant 1/(\log x)^{1/3}} \min\left\{\frac{\log x}{V}, \frac{1}{|t|}\right\} dt$$

$$\ll \frac{e^{2\min\{\sqrt{\log\log x}, \frac{1}{1-q}\}}\log^2 x}{V^2(\log\log x)^3},$$

which is more than acceptable for Key Proposition 5.

To handle the contribution from $1/(\log x)^{1/3} < |t| \leqslant 1$, we follow the same approach but need to deal with the indicator functions $\mathbf{1}_{L(0)}, \mathbf{1}_{L(t)}$ in (8) more carefully because our saving $1/\log^4(2/|t|)$ from the minimum in (7) is no longer so great. We temporarily set $D = D(t) = \lceil 2\log(1/|t|)\rceil + (B+1)$ so that $B + 1 \leqslant D \leqslant (2/3)\log\log x + B + O(1)$ and $e^D \asymp \frac{e^B}{|t|^2} \asymp \frac{1}{|t|^2}$. If the event $L(0)$ occurs then, comparing the definition of $L(0)$ with $j = \lfloor\log\log x\rfloor - D$ and with general

j (and recalling that B is an absolute constant), we must in particular have

$$\left(\frac{\log x}{e^{j+1}}\right)^{-B} e^{-2\min\{\sqrt{\log\log x}, \frac{1}{1-q}\}|t|^2} \log^2\left(\frac{2}{|t|}\right)$$

$$\ll \prod_{l=j}^{\lfloor\log\log x\rfloor - D - 1} \left|I_l\left(\frac{1}{2} + \frac{4V}{\log x}\right)\right|$$

$$\ll \frac{\log x}{e^{j+1}} e^{-2\log\log(\frac{\log x}{e^{j+1}})} \frac{e^{2\min\{\sqrt{\log\log x}, 1/(1-q)\}}}{|t|^{2B}}$$

for all $\lfloor \log V \rfloor + 3 \leqslant j \leqslant \log\log x - D - 1$. Similarly, if the event $L(t)$ occurs, then we must have the same bounds with $|I_l(1/2 + \frac{4V}{\log x})|$ replaced by $|I_l(1/2 + \frac{4V}{\log x} + it)|$.

Let us write $R(t)$ for the event that one has these bounds for both $|I_l(1/2 + \frac{4V}{\log x})|$ and $|I_l(1/2 + \frac{4V}{\log x} + it)|$. Then for $1/(\log x)^{1/3} < |t| \leqslant 1$, we can upper bound (8) by

$$\mathbb{E}\mathbf{1}_{R(t)}\left|F\left(\frac{1}{2} + \frac{4V}{\log x}\right)\right|^2 \frac{|F(1/2 + \frac{4V}{\log x} + it)|^2}{\prod_{l=\lfloor\log(|t|\log x)\rfloor}^{\lfloor\log\log x\rfloor - B - 2} |I_l(1/2 + \frac{4V}{\log x} + it)|^2}.$$

Now, here we have $\log\log x - D - 1 < \lfloor\log(|t|\log x)\rfloor$. Since $f(p)$ is independent on distinct primes, and so the event $R(t)$ is independent of $\prod_{l=\lfloor\log(|t|\log x)\rfloor}^{\lfloor\log\log x\rfloor - B - 2} |I_l(1/2 + \frac{4V}{\log x})|^2$, we may pull $\mathbb{E}\prod_{l=\lfloor\log(|t|\log x)\rfloor}^{\lfloor\log\log x\rfloor - B - 2} |I_l(1/2 + \frac{4V}{\log x})|^2$ out from the product $|F(1/2 + \frac{4V}{\log x})|^2$ above and use (2) to show it is $= \exp\{\sum_{p \leqslant e^{1/|t|}} \frac{1}{p^{1+8V/\log x}} + O(1)\} \ll 1/|t|$. So employing our tilted probability notation from Section 5.1, we have that (8) is

$$\ll (1/|t|)\mathbb{E}\mathbf{1}_{R(t)} \frac{|F(1/2 + \frac{4V}{\log x})|^2}{\prod_{l=\lfloor\log(|t|\log x)\rfloor}^{\lfloor\log\log x\rfloor - B - 2} |I_l(1/2 + \frac{4V}{\log x})|^2}$$

$$\times \frac{|F(1/2 + \frac{4V}{\log x} + it)|^2}{\prod_{l=\lfloor\log(|t|\log x)\rfloor}^{\lfloor\log\log x\rfloor - B - 2} |I_l(1/2 + \frac{4V}{\log x} + it)|^2}$$

$$= \frac{1}{|t|}\tilde{\mathbb{P}}_t^{\mathrm{St},(2)}(R(t))\mathbb{E}\frac{|F(1/2 + \frac{4V}{\log x})|^2}{\prod_{l=\lfloor\log(|t|\log x)\rfloor}^{\lfloor\log\log x\rfloor - B - 2} |I_l(1/2 + \frac{4V}{\log x})|^2}$$

$$\times \frac{|F(1/2 + \frac{4V}{\log x} + it)|^2}{\prod_{l=\lfloor\log(|t|\log x)\rfloor}^{\lfloor\log\log x\rfloor - B - 2} |I_l(1/2 + \frac{4V}{\log x} + it)|^2}.$$

We may deploy Proposition 7 to bound $\tilde{\mathbb{P}}_t^{\mathrm{St},(2)}(R(t))$, taking $n = \lfloor\log\log x\rfloor - D - (\lfloor\log V\rfloor + 3) \gg \log\log x$, $\sigma = \frac{4V}{\log x}$, $a = 2\min\{\sqrt{\log\log x}, 1/(1-q)\} + $

$O(\log(2/|t|))$ and $h(j) = -2\log j$, and obtaining that

$$\tilde{\mathbb{P}}_t^{\mathrm{St},(2)}(R(t)) \ll \left(\frac{1}{1 + (1-q)\sqrt{\log\log x}} + \frac{\log(2/|t|)}{\sqrt{\log\log x}} \right)^2$$

$$\ll \frac{\log^2(2/|t|)}{(1 + (1-q)\sqrt{\log\log x})^2}.$$

Furthermore, as we did earlier when $|t| \leqslant 1/\log^{1/3} x$, we may use (6) and standard prime number estimates to obtain that the expectation in the preceding display is $\ll (\frac{|t|\log x}{V})^2$.

Finally, putting together (7) with the above upper bounds for (8) yields that

$$\int_{1/(\log x)^{1/3} < |t| \leqslant 1} \mathbb{E}\mathbf{1}_{L(0)} \left| F\left(1/2 + \frac{4V}{\log x}\right) \right|^2 \mathbf{1}_{L(t)} \left| F\left(1/2 + \frac{4V}{\log x} + it\right) \right|^2 dt$$

$$\ll \frac{e^{2\min\{\sqrt{\log\log x}, 1/(1-q)\}}}{(1 + (1-q)\sqrt{\log\log x})^2}$$

$$\times \int_{1/(\log x)^{1/3} < |t| \leqslant 1} \left(\frac{1}{|t|\log^2(2/|t|)} \right)^2 \frac{\log^2(2/|t|)}{|t|} \left(\frac{|t|\log x}{V} \right)^2 dt$$

$$\ll \frac{e^{2\min\{\sqrt{\log\log x}, 1/(1-q)\}}\log^2 x}{V^2(1 + (1-q)\sqrt{\log\log x})^2} \int_{1/(\log x)^{1/3} < |t| \leqslant 1} \frac{1}{|t|\log^2(2/|t|)} dt$$

$$\ll \frac{e^{2\min\{\sqrt{\log\log x}, \frac{1}{1-q}\}}\log^2 x}{V^2(1 + (1-q)\sqrt{\log\log x})^2}.$$

Because of the factor $\log^2(2/|t|)$ (which we remark descends from the term $-2\log\log(\frac{\log x}{e^{j+1}})$ in the original definition of the set $\mathcal{L}$), the integral over t here was $\ll 1$ and we obtain the upper bound claimed in Key Proposition 5. □

5.3. The lower bound in the Rademacher case. As in our upper bound arguments, there is not too much difference between the Rademacher and Steinhaus cases, except the former is more notationally complicated because we no longer have 'translation invariance in law' of the Euler products. We will only sketch the changes that are needed to deduce the Rademacher lower bound.

In place of Lemma 6, in the Rademacher case, one needs to compute

$$\mathbb{E} \prod_{x < p \leqslant y} \left| 1 + \frac{f(p)}{p^{1/2+\sigma+it_1}} \right|^{2+iu} \left| 1 + \frac{f(p)}{p^{1/2+\sigma+i(t_1+t_2)}} \right|^{2+iv},$$

where t_1, t_2, u, v are real. If one does this (proceeding as in the proofs of Lemmas 2 and 6), one obtains an expression that is, up to error terms, the exponential of

$$\sum_{x < p \leqslant y} \left(\frac{(1 + iu/2)^2 + (1 + iv/2)^2}{p^{1+2\sigma}} + \frac{(2 + iu)(2 + iv) \cos(t_1 \log p) \cos((t_1 + t_2) \log p)}{p^{1+2\sigma}} \right.$$
$$\left. + \frac{((2 + iu)^2/2 - (2 + iu)) \cos(2t_1 \log p) + ((2 + iv)^2/2 - (2 + iv)) \cos(2(t_1 + t_2) \log p)}{2p^{1+2\sigma}} \right),$$

which can be simplified a little by observing it is

$$= \sum_{x < p \leqslant y} \left(\frac{(1 + iu/2)^2 + (1 + iv/2)^2}{p^{1+2\sigma}} + \frac{(2 + iu)(2 + iv) \cos(t_2 \log p)}{2p^{1+2\sigma}} \right.$$
$$+ \frac{(2 + iu)(iu/2) \cos(2t_1 \log p) + (2 + iv)(iv/2) \cos(2(t_1 + t_2) \log p)}{2p^{1+2\sigma}}$$
$$\left. + \frac{(2 + iu)(2 + iv) \cos((2t_1 + t_2) \log p)}{2p^{1+2\sigma}} \right).$$

Note that the first two fractions here are precisely analogous to the Steinhaus case in Lemma 6, whereas the second two terms are new.

For each real t, we can define the event $L(t) = L_{x,q,V}(t)$ exactly as in the Steinhaus case, but now we let $\mathcal{L}$ be the random subset of points $t \in [1/3, 1/2]$ (rather than $[-1/2, 1/2]$) at which $L(t)$ occurs. The point of this change is that we will then always have $t_1, t_1 + t_2 \in [1/3, 1/2]$, and therefore $2/3 \leqslant 2t_1$, $2(t_1 + t_2), 2t_1 + t_2 \leqslant 1$, in our calculations. This means that the two new terms in the Rademacher characteristic functions will only cause negligible changes in mean, variance and covariance as compared with the Steinhaus case (because $\sum_{x < p \leqslant y} \frac{\cos(2t_1 \log p)}{p^{1+2\sigma}}$, $\sum_{x < p \leqslant y} \frac{\cos(2(t_1 + t_2) \log p)}{p^{1+2\sigma}}$, $\sum_{x < p \leqslant y} \frac{\cos((2t_1 + t_2) \log p)}{p^{1+2\sigma}}$ all exhibit significant cancellation, indeed they are all $\ll 1/\log x$, say). Once one is in this situation, the Rademacher lower bound follows simply by imitating the arguments from the Steinhaus case.

6. Proof of Corollary 2, lower bound

In this section, we shall prove our last remaining result that for all $2 \leqslant \lambda \leqslant e^{\sqrt{\log \log x}}$, we have

$$\mathbb{P}\left(\left| \sum_{n \leqslant x} f(n) \right| \geqslant \lambda \frac{\sqrt{x}}{(\log \log x)^{1/4}} \right) \gg \frac{1}{\lambda^2 (\log \log x)^{O(1)}}.$$

The proof is almost the same in the Steinhaus and Rademacher cases, but for definiteness, let us first think of $f(n)$ as a Steinhaus random multiplicative

function. As before, $F(s)$ will denote the Euler product of $f(n)$ over x-smooth numbers. The proof will consist of three main steps: we will prove that it will suffice to establish a suitable lower bound for $\mathbb{P}(\int_{-1/2}^{1/2} |F(1/2 + \frac{\log\log x}{\log x} + it)|^2 \, dt \geqslant \beta \log x)$; then we will show it will actually suffice to prove a corresponding lower bound for a discrete maximum of $|F(1/2 + \frac{\log\log x}{\log x} + it)|^2$ (roughly speaking); finally, we will deduce that from some existing results on maxima of random processes.

Let $\hat{\mathbb{E}}$ denote expectation conditional on the values $(f(p))_{p \leqslant \sqrt{x}}$ and let $\hat{\mathbb{P}}$ denote the corresponding conditional probability. Thus, if we let $P(n)$ denote the largest prime factor of n, then $\hat{\mathbb{E}}|\sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n)|$ is a function of the random variables $(f(p))_{p \leqslant \sqrt{x}}$, and in fact we have

$$
\hat{\mathbb{E}} \left| \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \right| = \hat{\mathbb{E}} \left| \sum_{\sqrt{x} < p \leqslant x} f(p) \sum_{m \leqslant x/p} f(m) \right| \asymp \sqrt{\sum_{\sqrt{x} < p \leqslant x} \left| \sum_{m \leqslant x/p} f(m) \right|^2},
$$

by Khintchine's inequality (as in the proof of Proposition 3).

LEMMA 8. *Let $f(n)$ be a Steinhaus or Rademacher random multiplicative function, and let $\mathcal{A}$ denote the event that*

$$
\left| \sum_{n \leqslant x} f(n) \right| \geqslant (1/2)\hat{\mathbb{E}} \left| \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \right|.
$$

Then for any realization of $(f(p))_{p \leqslant \sqrt{x}}$, we have the uniform lower bound

$$
\hat{\mathbb{P}}(\mathcal{A}) \gg 1.
$$

Once we have proved Lemma 8, to prove the lower bound in Corollary 2, it will suffice to show that

$$
\mathbb{P} \left(\sqrt{\sum_{\sqrt{x} < p \leqslant x} \left| \sum_{m \leqslant x/p} f(m) \right|^2} \geqslant \lambda \frac{\sqrt{x}}{(\log\log x)^{1/4}} \right)
$$

$$
\gg \frac{1}{\lambda^2 (\log\log x)^{O(1)}} \quad \forall \, 2 \leqslant \lambda \leqslant e^{\sqrt{\log\log x}}
$$

or, equivalently, that

$$\mathbb{P}\left(\sum_{\sqrt{x} < p \leqslant x} \left| \sum_{m \leqslant x/p} f(m) \right|^2 \geqslant \beta x \right) \gg \frac{1}{\beta (\log \log x)^{O(1)}} \quad \forall \, 2 \leqslant \beta \leqslant e^{2\sqrt{\log \log x}}.$$

(9)

(Strictly speaking, since β here corresponds to $\frac{\lambda^2}{\sqrt{\log \log x}}$ the range for β in (9) should be $\frac{4}{\sqrt{\log \log x}} \leqslant \beta \leqslant \frac{e^{2\sqrt{\log \log x}}}{\sqrt{\log \log x}}$. However, by adjusting the $O(1)$ exponents suitably, it will certainly suffice, once we have Lemma 8, to prove (9).)

Proof of Lemma 8. To establish that $\hat{\mathbb{P}}(\mathcal{A}) \gg 1$ for any realization of $(f(p))_{p \leqslant \sqrt{x}}$, note that (as in the proof of Proposition 3) we have

$$\sum_{n \leqslant x} f(n) = \sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n) + \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \stackrel{d}{=} \sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n) + \epsilon \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n),$$

where ϵ is an auxiliary Rademacher random variable independent of everything else and $\stackrel{d}{=}$ denotes equality in distribution. By the triangle inequality, we have

$$2 \left| \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \right| \leqslant \left| \sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n) + \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \right|$$

$$+ \left| \sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n) - \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \right|,$$

so for at least one of the values ± 1 of ϵ, we must have $|\sum_{\substack{n \leqslant x, \\ P(n) \leqslant \sqrt{x}}} f(n) + \epsilon \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n)| \geqslant |\sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n)|$. So if we let $\mathcal{B}$ denote the event that

$$\left| \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \right| \geqslant (1/2)\hat{\mathbb{E}} \left| \sum_{\substack{n \leqslant x, \\ P(n) > \sqrt{x}}} f(n) \right|,$$

then we have $\hat{\mathbb{P}}(\mathcal{A}) \geqslant (1/2)\hat{\mathbb{P}}(\mathcal{B})$. Meanwhile, we can lower bound $\hat{\mathbb{P}}(\mathcal{B})$ in a standard (Paley–Zygmund type) way using the Cauchy–Schwarz inequality,

noting that

$$\hat{\mathbb{P}}(\mathcal{B}) \geqslant \frac{(\hat{\mathbb{E}}\mathbf{1}_{\mathcal{B}}|\sum_{\substack{n\leqslant x,\\ P(n)>\sqrt{x}}} f(n)|)^2}{\hat{\mathbb{E}}|\sum_{\substack{n\leqslant x,\\ P(n)>\sqrt{x}}} f(n)|^2}$$

$$= \frac{(\hat{\mathbb{E}}|\sum_{\substack{n\leqslant x,\\ P(n)>\sqrt{x}}} f(n)| - \hat{\mathbb{E}}\mathbf{1}_{\mathcal{B}\text{ fails}}|\sum_{\substack{n\leqslant x,\\ P(n)>\sqrt{x}}} f(n)|)^2}{\hat{\mathbb{E}}|\sum_{\substack{n\leqslant x,\\ P(n)>\sqrt{x}}} f(n)|^2}$$

$$\gg \frac{(\hat{\mathbb{E}}|\sum_{\substack{n\leqslant x,\\ P(n)>\sqrt{x}}} f(n)|)^2}{\hat{\mathbb{E}}|\sum_{\substack{n\leqslant x,\\ P(n)>\sqrt{x}}} f(n)|^2}.$$

Writing $\sum_{\substack{n\leqslant x,\\ P(n)>\sqrt{x}}} f(n) = \sum_{\sqrt{x}<p\leqslant x} f(p)\sum_{m\leqslant x/p} f(m)$ and applying Khintchine's inequality again, we obtain that the conditional expectation (squared) in the numerator and the conditional expectation in the denominator have the same order. So we have shown $\hat{\mathbb{P}}(\mathcal{A}) \geqslant (1/2)\hat{\mathbb{P}}(\mathcal{B}) \gg 1$, as we wanted. □

To establish (9), we again begin similarly as in the proof of Proposition 3, noting (with $X = e^{\sqrt{\log x}}$) that $\sum_{\sqrt{x}<p\leqslant x} \left|\sum_{m\leqslant x/p} f(m)\right|^2$ is

$$\geqslant \sum_{\sqrt{x}<p\leqslant x} \frac{\log p}{\log x} \left|\sum_{m\leqslant x/p} f(m)\right|^2$$

$$= \frac{1}{\log x} \sum_{\sqrt{x}<p\leqslant x} \log p \frac{X}{p} \int_p^{p(1+1/X)} \left|\sum_{m\leqslant x/p} f(m)\right|^2 dt$$

$$\geqslant \frac{1}{\log x} \sum_{\sqrt{x}<p\leqslant x} \log p \frac{X}{p} \int_p^{p(1+1/X)} \frac{1}{2} \left|\sum_{m\leqslant x/t} f(m)\right|^2 dt$$

$$- \frac{1}{\log x} \sum_{\sqrt{x}<p\leqslant x} \log p \frac{X}{p} \int_p^{p(1+1/X)} \left|\sum_{x/t<m\leqslant x/p} f(m)\right|^2 dt.$$

The subtracted term here satisfies the first moment bound

$$\frac{1}{\log x} \sum_{\sqrt{x}<p\leqslant x} \log p \frac{X}{p} \int_p^{p(1+1/X)} \mathbb{E}\left|\sum_{x/t<m\leqslant x/p} f(m)\right|^2 dt$$

$$\ll \frac{1}{\log x} \sum_{\sqrt{x} < p \leqslant x} \log p \frac{X}{p} \int_p^{p(1+1/X)} \left(\frac{x}{pX} + 1 \right) dt$$

$$= \frac{1}{\log x} \sum_{\sqrt{x} < p \leqslant x} \log p \left(\frac{x}{pX} + 1 \right) \ll \frac{x}{\log x},$$

so by Markov's inequality, the probability it is larger than $x/\sqrt{\log x}$ is $\ll \frac{1}{\sqrt{\log x}}$. Since $x/\sqrt{\log x}$ is negligible compared with the term βx in (9) and $\frac{1}{\sqrt{\log x}}$ is negligible compared with our target probability lower bound $\frac{1}{\beta(\log\log x)^{O(1)}}$, we may ignore this subtracted term. Meanwhile, still broadly following the proof of Proposition 3, we have that

$$\frac{1}{\log x} \sum_{\sqrt{x} < p \leqslant x} \log p \frac{X}{p} \int_p^{p(1+1/X)} \left| \sum_{m \leqslant x/t} f(m) \right|^2 dt$$

$$\gg \frac{1}{\log x} \int_{\sqrt{x}}^x \left| \sum_{m \leqslant x/t} f(m) \right|^2 dt$$

$$= \frac{x}{\log x} \int_1^{\sqrt{x}} \left| \sum_{m \leqslant z} f(m) \right|^2 \frac{dz}{z^2}$$

$$\geqslant \frac{x}{\log x} \int_1^{\sqrt{x}} \left| \sum_{\substack{m \leqslant z, \\ m \text{ is } x \text{ smooth}}} f(m) \right|^2 \frac{dz}{z^{2+2\log\log x/\log x}}.$$

We would like to complete the integral to the range $\int_1^\infty$, so we can apply Harmonic Analysis Result 1 and have a lower bound $\gg \frac{x}{\log x} \int_{-1/2}^{1/2} |F(1/2 + \frac{\log\log x}{\log x} + it)|^2 dt$, where F is the Euler product of $f(n)$ over x-smooth numbers. It turns out that the shift $(\log\log x)/\log x$ that we introduced makes it acceptable to perform this completion, as we have the first moment bound

$$\frac{x}{\log x} \int_{\sqrt{x}}^\infty \mathbb{E} \left| \sum_{\substack{m \leqslant z, \\ x \text{ smooth}}} f(m) \right|^2 \frac{dz}{z^{2+2\log\log x/\log x}} \leqslant \frac{x}{\log x} \int_{\sqrt{x}}^\infty \frac{dz}{z^{1+2\log\log x/\log x}}$$

$$\ll \frac{x}{\log x \log\log x},$$

so by Markov's inequality, $\int_{\sqrt{x}}^{\infty}$ is negligible as in our discussion of the subtracted term above.

Reviewing the arguments in the previous paragraph, we see that we have established the following reduction.

LEMMA 9. *To prove (9) (and therefore the lower bound in Corollary 2), it will suffice to show that*

$$\mathbb{P}\left(\int_{-1/2}^{1/2}\left|F\left(1/2 + \frac{\log\log x}{\log x} + it\right)\right|^2 dt \geqslant \beta \log x\right)$$
$$\gg \frac{1}{\beta(\log\log x)^{O(1)}} \quad \forall\, 2 \leqslant \beta \leqslant e^{2\sqrt{\log\log x}}. \tag{10}$$

Now to prove (10), our arguments begin to differ from what we have done previously. We would like to say that the integral is essentially the same as $\sum_{|k| \leqslant \frac{\log x - 1}{2}} \frac{1}{\log x}|F(1/2 + \frac{\log\log x}{\log x} + i\frac{k}{\log x})|^2$, so we could lower bound the probability on the left by the probability that the maximum of the $|F(1/2 + \frac{\log\log x}{\log x} + i\frac{k}{\log x})|^2$ is $\gg \beta \log^2 x$. It seems technically tricky to establish an approximation quite like that, but since we only want a lower bound and since the exponential function is convex, we can exploit Jensen's inequality, obtaining that

$$\int_{-1/2}^{1/2}\left|F\left(\frac{1}{2} + \frac{\log\log x}{\log x} + it\right)\right|^2$$

$$\geqslant \sum_{|k| \leqslant \frac{\log x - 1}{2}} \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}}\left|F\left(\frac{1}{2} + \frac{\log\log x}{\log x} + i\frac{k}{\log x} + it\right)\right|^2 dt$$

$$= \frac{1}{\log x}\sum_{|k| \leqslant \frac{\log x - 1}{2}}\left(\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}}\exp\left\{2\log\left|F\left(1/2 + \frac{\log\log x}{\log x} + i\frac{k}{\log x} + it\right)\right|\right\} dt\right)$$

$$\geqslant \frac{1}{\log x}\sum_{|k| \leqslant \frac{\log x - 1}{2}}\exp\left\{2\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}}\log\left|F\left(1/2 + \frac{\log\log x}{\log x} + i\frac{k}{\log x} + it\right)\right| dt\right\}.$$

The exponential in the sum here will behave, for our purposes, in exactly the same way as $|F(1/2 + \frac{\log\log x}{\log x} + i\frac{k}{\log x})|^2$, but we may calculate precisely that (in the Steinhaus case) it is

$$\exp\left\{-2\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \sum_{p \leqslant x} \Re \log\left(1 - \frac{f(p)}{p^{1/2 + \frac{\log\log x}{\log x} + i\frac{k}{\log x} + it}}\right) dt\right\}$$

$$= \exp\left\{2\sum_{p \leqslant x}\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}}\left(\Re\frac{f(p)p^{-i\frac{k}{\log x} - it}}{p^{1/2 + \frac{\log\log x}{\log x}}} + \Re\frac{(f(p)p^{-i\frac{k}{\log x} - it})^2}{2p^{1 + \frac{2\log\log x}{\log x}}} + O\left(\frac{1}{p^{3/2}}\right)\right) dt\right\}.$$

We can simplify this expression a bit further by noting in the second term that

$$\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} e^{-2it\log p}\, dt = \log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} (1 + O(|t|\log p))\, dt$$

$$= 1 + O((\log p)/\log x)$$

and that $\sum_{p \leqslant x} \frac{\log p}{p^{1 + \frac{2\log\log x}{\log x}}\log x} \leqslant \sum_{p \leqslant x}\frac{\log p}{p\log x} = O(1)$ by Chebychev's estimate, as is $\sum_{p \leqslant x}\frac{1}{p^{3/2}}$. Thus, the exponential is

$$= \exp\left\{2\left(\sum_{p \leqslant x}\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re\frac{f(p)p^{-i\frac{k}{\log x} - it}}{p^{1/2 + \frac{\log\log x}{\log x}}}\, dt + \sum_{p \leqslant x}\Re\frac{(f(p)p^{-i\frac{k}{\log x}})^2}{2p^{1 + \frac{2\log\log x}{\log x}}}\right) + O(1)\right\},$$

and to prove (10) (and therefore also (9)), it will suffice to show that

$$\mathbb{P}\left(\max_{|k| \leqslant \frac{\log x - 1}{2}}\left(\sum_{p \leqslant x}\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re\frac{f(p)p^{-\frac{ik}{\log x} - it}}{p^{1/2 + \frac{\log\log x}{\log x}}} + \sum_{p \leqslant x}\Re\frac{(f(p)p^{-\frac{ik}{\log x}})^2}{2p^{1 + \frac{2\log\log x}{\log x}}}\right)\right.$$

$$\geqslant \frac{\log\beta}{2} + \log\log x\Bigg)$$

$$\gg \frac{1}{\beta(\log\log x)^{O(1)}} \quad \forall\, 2 \leqslant \beta \leqslant e^{2\sqrt{\log\log x}}.$$

As a final simplification, we shall tweak things a little so that we can ignore the nuisance second sum and the contribution from small primes. Let us note that

$$\left|\sum_{p \leqslant \log^{10} x}\frac{(f(p)p^{-i\frac{k}{\log x}})^2}{p^{1 + \frac{2\log\log x}{\log x}}}\right| \leqslant \sum_{p \leqslant \log^{10} x}\frac{1}{p} = \log\log\log x + O(1)$$

by the Mertens estimate, and

$$\mathbb{E}\left|\sum_{\log^{10} x < p \leqslant x}\frac{(f(p)p^{-i\frac{k}{\log x}})^2}{p^{1 + \frac{2\log\log x}{\log x}}}\right|^2 = \sum_{\log^{10} x < p \leqslant x}\frac{1}{p^{2 + \frac{4\log\log x}{\log x}}} \ll \frac{1}{\log^{10} x},$$

so we have

$$\mathbb{P}\left(\max_{|k|\leqslant\frac{\log x-1}{2}}\left|\sum_{\log^{10} x<p\leqslant x}\Re\frac{(f(p)p^{-i\frac{k}{\log x}})^2}{2p^{1+\frac{2\log\log x}{\log x}}}\right|\geqslant 1\right)\ll 1/\log^9 x$$

by the union bound and Chebychev's inequality. Thus, since $1/\log^9 x$ is negligible compared with our target probability lower bound $\frac{1}{\beta(\log\log x)^{O(1)}}$ and since replacing $\frac{\log\beta}{2}+\log\log x$ by $\frac{\log\beta}{2}+\log\log\log x+O(1)+\log\log x$ (essentially replacing β by $\beta e^{2\log\log\log x+O(1)}$) does not change the form of that lower bound, we can omit the second sum $\sum_{p\leqslant x}\Re\frac{(f(p)p^{-i\frac{k}{\log x}})^2}{2p^{1+\frac{2\log\log x}{\log x}}}$ entirely when trying to prove the lower bound above. Furthermore, since the contributions from different primes are independent, and for any given k, the symmetry of the $f(p)$ implies that

$$\mathbb{P}\left(\sum_{p\leqslant\log^{10} x}\log x\int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}}\Re\frac{f(p)p^{-i\frac{k}{\log x}-it}}{p^{1/2+\frac{\log\log x}{\log x}}}\,dt\geqslant 0\right)=1/2,$$

we are free to omit the part of the first sum over primes $p\leqslant\log^{10} x$ (say) as well. Thus, we have established the following.

LEMMA 10. *Let $f(n)$ be a Steinhaus random multiplicative function. To prove (10) (and therefore the lower bound in Corollary 2), it will suffice to show that*

$$\mathbb{P}\left(\max_{|k|\leqslant\frac{\log x-1}{2}}\left(\sum_{\log^{10} x<p\leqslant x}\log x\int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}}\Re\frac{f(p)p^{-i\frac{k}{\log x}-it}}{p^{1/2+\frac{\log\log x}{\log x}}}\,dt\right)\geqslant\frac{\log\beta}{2}+\log\log x\right)$$

$$\gg\frac{1}{\beta(\log\log x)^{O(1)}}\quad\forall\,2\leqslant\beta\leqslant e^{2\sqrt{\log\log x}}.\tag{11}$$

The lower bound (11) is the form in which we shall complete our proof. The idea is that the random sums corresponding to different k should behave, for the purpose of these tail probabilities, roughly like $\log x$ *independent* random variables each having a $N(0,(1/2)\log\log x)$ distribution, and, therefore, the probability that their maximum is larger than $\frac{\log\beta}{2}+\log\log x$ should be roughly

$$\log x\cdot\mathbb{P}\left(N(0,(1/2)\log\log x)\geqslant\frac{\log\beta}{2}+\log\log x\right)$$

$$\gg\log x\cdot\frac{e^{-(\frac{\log\beta}{2}+\log\log x)^2/\log\log x}}{\sqrt{\log\log x}}$$

$$\gg\frac{1}{\beta\sqrt{\log\log x}},$$

on our range of β.

To prove this, using a multivariate central limit theorem (for example, Theorem 2.1 of Reinert and Röllin [**22**], applied with the same kind of exchangeable pair construction as in Appendix B of Harper [**10**]), we may replace the collection of sums

$$\left(\sum_{\log^{10} x < p \leqslant x} \log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re \frac{f(p) p^{-i\frac{k}{\log x} - it}}{p^{1/2 + \frac{\log \log x}{\log x}}} \, dt \right)_{|k| \leqslant \frac{\log x - 1}{2}}$$

in (11) by a collection of Gaussian random variables $(X(k))_{|k| \leqslant \frac{\log x - 1}{2}}$ with the same means, variances and covariances. (More precisely, the multivariate central limit theorem allows us to lower bound our probability in (11) by the corresponding Gaussian probability, with $\frac{\log \beta}{2} + \log \log x$ replaced by $\frac{\log \beta}{2} + 1 + \log \log x$, and up to an error term saving a power of $\log x$ which is therefore negligible compared with $\frac{1}{\beta(\log \log x)^{O(1)}}$.) Furthermore, the mean is clearly 0 for each k, and for any j and k, the covariance is

$$\mathbb{E} \left(\sum_{\log^{10} x < p \leqslant x} \log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re \frac{f(p) p^{-i\frac{k}{\log x} - it}}{p^{1/2 + \frac{\log \log x}{\log x}}} \right)$$

$$\times \left(\sum_{\log^{10} x < p \leqslant x} \log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re \frac{f(p) p^{-i\frac{j}{\log x} - it}}{p^{1/2 + \frac{\log \log x}{\log x}}} \right)$$

$$= \sum_{\log^{10} x < p \leqslant x} \frac{\mathbb{E}(\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re f(p) p^{-i\frac{k}{\log x} - it} \, dt)(\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re f(p) p^{-i\frac{j}{\log x} - it} \, dt)}{p^{1 + \frac{2\log \log x}{\log x}}}$$

$$= \sum_{\log^{10} x < p \leqslant x} \frac{\mathbb{E}(\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re f(p) p^{-it} \, dt)(\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re f(p) p^{-i\frac{(j-k)}{\log x} - it} \, dt)}{p^{1 + \frac{2\log \log x}{\log x}}}$$

since in the Steinhaus case, $f(p) p^{-ik/\log x}$ has the same distribution as $f(p)$. Note that this means that the covariance is a function of $j - k$ (and x) only, in other words, our random variables are *stationary*. Furthermore, using that $\mathbb{E}(\Re f(p) p^{-it_1})(\Re f(p) p^{-i\frac{(j-k)}{\log x} - it_2})$ is

$$(1/2) \cos \left(\left(\frac{j - k}{\log x} + (t_2 - t_1) \right) \log p \right)$$

$$= (1/2) \cos \left(\frac{j - k}{\log x} \log p \right) + O(|t_1 - t_2| \log p),$$

we find the covariance is

$$\sum_{\log^{10} x < p \leqslant x} \frac{\log^2 x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \cos((\frac{j-k}{\log x} + (t_2 - t_1)) \log p)}{2p^{1+\frac{2\log\log x}{\log x}}}$$

$$= \sum_{\log^{10} x < p \leqslant x} \frac{\cos(\frac{(j-k)\log p}{\log x})}{2p^{1+\frac{2\log\log x}{\log x}}} + O(1).$$

Here the final equality also used Chebychev's bound $\sum_{p \leqslant x} \frac{\log p}{p} \ll \log x$.

Later, we will also need to know that the covariance is a decreasing function of $|j - k|$, on a suitable range of $|j - k|$. To establish this, we just note that if we differentiate,

$$\frac{d}{dh} \sum_{\log^{10} x < p \leqslant x} \frac{\log^2 x \int \int \cos((h + (t_2 - t_1)) \log p)}{2p^{1+\frac{2\log\log x}{\log x}}}$$

$$= - \sum_{\log^{10} x < p \leqslant x} \frac{\log p \log^2 x \int \int \sin((h + (t_2 - t_1)) \log p)}{2p^{1+\frac{2\log\log x}{\log x}}},$$

then

$$-\sum_{\log^{10} x < p \leqslant x} \frac{\log p \sin(h \log p)}{p^{1+\frac{2\log\log x}{\log x}}} = \Im \sum_{p} \frac{\log p}{p^{1+\frac{2\log\log x}{\log x}+ih}} + O(\log\log x)$$

$$= -\Im \frac{\zeta'}{\zeta} \left(1 + \frac{2\log\log x}{\log x} + ih\right) + O(\log\log x),$$

where $\zeta(s)$ denotes the Riemann zeta function. In particular, if $|h|$ is small, then it is known that

$$-\frac{\zeta'}{\zeta} \left(1 + \frac{2\log\log x}{\log x} + ih\right) = \frac{1}{(2\log\log x)/\log x + ih} + O(1).$$

See Theorem 6.7 of Montgomery and Vaughan [20], for example. So provided $(\log\log x)/\log x \leqslant h \leqslant 1/(500 \log\log x)$, say, the derivative will be negative. This translates into a range $\log\log x \leqslant |j - k| \leqslant (\log x)/(500 \log\log x)$.

Now tidying up a bit, the contribution to the covariance sum from primes $p > x^{1/\log\log x}$ is

$$\ll \frac{\log\log x}{\log x} \sum_{p \leqslant x} \frac{\log p}{p^{1+\frac{2\log\log x}{\log x}}},$$

and using a Chebychev-type prime number estimate again this is

$$\ll \frac{\log \log x}{\log x} \sum_{n \leqslant x} \frac{1}{n^{1 + \frac{2 \log \log x}{\log x}}} \ll 1.$$

When $p \leqslant x^{1/\log\log x}$, writing $\frac{1}{p^{1 + \frac{2\log\log x}{\log x}}} = \frac{1 + O((\log p \log\log x)/\log x)}{p}$ and applying the Chebychev bound

$$\sum_{p \leqslant x^{1/\log\log x}} \frac{\log p}{p} \ll \frac{\log x}{\log\log x},$$

we see that we can replace $\frac{1}{p^{1 + \frac{2\log\log x}{\log x}}}$ by $\frac{1}{p}$ in our expression for the covariance. *Thus, when $|j - k| \leqslant \log\log x$, and in particular for the variance where $j = k$,* we get that

$$\mathbb{E}X(j)X(k) = \sum_{\log^{10} x < p \leqslant x^{1/\log\log x}} \frac{1 + O((\frac{(j-k)\log p}{\log x})^2)}{2p} + O(1)$$

$$= \sum_{\log^{10} x < p \leqslant x^{1/\log\log x}} \frac{1}{2p} + O(1)$$

$$= (1/2)(\log\log x - 2\log\log\log x) + O(1),$$

using the Mertens estimate for $\sum 1/p$. Similarly, *when $\log\log x < |j - k| \leqslant (\log x)/10\log\log x$, say,* then standard prime number estimates (see the proof of Lemma 5 or Section 6.1 of Harper [10]) imply that

$$\sum_{x^{1/|j-k|} < p \leqslant x^{1/\log\log x}} \frac{\cos(\frac{(j-k)\log p}{\log x})}{2p} \ll 1,$$

so the covariance

$$\mathbb{E}X(j)X(k) = \sum_{\log^{10} x < p \leqslant x^{1/|j-k|}} \frac{1 + O((\frac{(j-k)\log p}{\log x})^2)}{2p} + O(1)$$

$$= (1/2)(\log\log x - \log|j - k| - \log\log\log x) + O(1).$$

Collecting everything together, we see that for any large parameter $E \in \mathbb{N}$, the probability on the left-hand side of (11) is

$$\mathbb{P}\left(\max_{|k| \leqslant \frac{\log x - 1}{2}} X(k) \geqslant \frac{\log \beta}{2} + \log\log x \right)$$

$$\geqslant \mathbb{P}\left(\max_{|j|\leqslant \frac{\log x}{1000E(\log\log x)^2}} \frac{X(jE\lfloor\log\log x\rfloor)}{\sqrt{\mathbb{E}X(jE\lfloor\log\log x\rfloor)^2}} \geqslant u\right),$$

where $u = u(\beta, x)$ is given by

$$u := \frac{(1/2)\log\beta + \log\log x}{\sqrt{\mathbb{E}X(jE\lfloor\log\log x\rfloor)^2}}$$

$$= \frac{(1/2)\log\beta + \log\log x}{\sqrt{(1/2)(\log\log x - 2\log\log\log x) + O(1)}} \geqslant \sqrt{2\log\log x}.$$

Recall here that $\mathbb{E}X(jE\lfloor\log\log x\rfloor)^2$ is the same for all j and E because our random variables are stationary, so $u(\beta, x)$ does not depend on j, E. We also always have $u(\beta, x) \ll \sqrt{\log\log x}$, on our range $2 \leqslant \beta \leqslant e^{2\sqrt{\log\log x}}$.

We will choose the parameter E and explain the reason for introducing it, shortly. We can ease notation by setting $Z(j) := X(jE\lfloor\log\log x\rfloor)/\sqrt{\mathbb{E}X(jE\lfloor\log\log x\rfloor)^2}$. Then the $Z(j)$ are mean zero, variance one, stationary Gaussian random variables whose covariance function $r(m) := \mathbb{E}Z(j)Z(j+m)$ satisfies

$$0 \leqslant r(m) = \frac{\log\log x - \log(mE\lfloor\log\log x\rfloor) - \log\log\log x + O(1)}{\log\log x - 2\log\log\log x + O(1)}$$

$$\leqslant 1 - \frac{(2\log(mE) + O(1))}{u^2},$$

for all $1 \leqslant m \leqslant (\log x)/(500E(\log\log x)^2)$.

At this point, having seen that $r(m)$ is nonnegative and decreasing on our range of m, we can use Theorem 1 of Harper [10], obtaining that $\mathbb{P}(\max_{|j|\leqslant \frac{\log x}{1000E(\log\log x)^2}} Z(j) \geqslant u)$ is

$$\gg \frac{\log x}{E(\log\log x)^2} \frac{e^{-u^2/2}}{u} \sqrt{\frac{1 - r(1)}{u^2 r(1)}}$$

$$\times \prod_{1\leqslant m\leqslant \frac{\log x}{500E(\log\log x)^2}} \Phi\left(u\sqrt{1 - r(m)}\left(1 + O\left(\frac{1}{u^2(1 - r(m))}\right)\right)\right)$$

$$\gg \frac{\sqrt{\log E}}{E} \frac{\log x}{(\log\log x)^{7/2}} e^{-u^2/2}$$

$$\times \prod_{1\leqslant m\leqslant \frac{\log x}{500E(\log\log x)^2}} \Phi\left(\sqrt{u^2(1 - r(m))}\left(1 + O\left(\frac{1}{u^2(1 - r(m))}\right)\right)\right),$$

where Φ denotes the standard normal cumulative distribution function. Here we see, by writing $(1/2)\log\beta + \log\log x = (1/2)\log\beta + 2\log\log\log x + (\log\log x - 2\log\log\log x)$, that

$$e^{-u^2/2} = \exp\left\{-\frac{((1/2)\log\beta + \log\log x)^2}{\log\log x - 2\log\log\log x + O(1)}\right\} \asymp \frac{1}{\beta\log x(\log\log x)^2}.$$

And by our previous calculations, together with the fact that

$$\Phi(t) = 1 - \int_t^\infty \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx = 1 - O\left(\frac{e^{-t^2/2}}{t}\right) \quad \text{for } t \geqslant 1,$$

the product over m is

$$\geqslant \prod_{m \leqslant \frac{\log x}{500E(\log\log x)^2}} \Phi\left(\sqrt{2\log(mE)}\left(1 + O\left(\frac{1}{\log(mE)}\right)\right)\right)$$

$$= \prod_{m \leqslant \frac{\log x}{500E(\log\log x)^2}} \left(1 - O\left(\frac{1}{mE\sqrt{\log(mE)}}\right)\right).$$

Now this product is $= \exp\{-\sum_{m \leqslant \frac{\log x}{500E(\log\log x)^2}} O(\frac{1}{mE\sqrt{\log(mE)}})\}$, so if we choose $E = \sqrt{\log\log x}$, say, then the sum is uniformly bounded and so the product is $\gg 1$. *This is the step where the spacing parameter E is important to obtain a good bound.* Putting everything together again, we have shown that the left-hand side of (11) is

$$\gg \frac{\sqrt{\log E}}{E}\frac{\log x}{(\log\log x)^{7/2}} \cdot \frac{1}{\beta\log x(\log\log x)^2} = \frac{1}{\beta(\log\log x)^{O(1)}},$$

as we wanted. $\qquad\square$

The Rademacher case is exactly the same as the Steinhaus case as far as (10) and the subsequent application of Jensen's inequality. At that point, things change a little because the Rademacher Euler product is slightly different than the Steinhaus one, and specifically in place of the term

$$2\sum_{p \leqslant x}\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re\frac{(f(p)p^{-i\frac{k}{\log x}-it})^2}{2p^{1+\frac{2\log\log x}{\log x}}}\,dt,$$

one has the negative of that term, and because $f(p)^2 \equiv 1$ in the Rademacher case, that becomes the deterministic quantity

$$-\sum_{p \leqslant x}\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re\frac{p^{-2i\frac{k}{\log x}-2it}}{p^{1+\frac{2\log\log x}{\log x}}}\,dt.$$

But this is simply

$$-\log x \int_{-\frac{1}{2\log x}}^{\frac{1}{2\log x}} \Re \log \zeta \left(1 + \frac{2 \log \log x}{\log x} + i \frac{2k}{\log x} + 2it \right) dt + O(1),$$

and if one restricts attention to the range $9(\log x)/20 \leqslant k \leqslant (\log x)/2$, then standard estimates for the zeta function (see Theorem 6.7 of Montgomery and Vaughan [20]) show this is all $\ll 1$, so may be discarded. Moreover, the restriction to the range $9(\log x)/20 \leqslant k \leqslant (\log x)/2$ makes no essential difference to the subsequent arguments since rather than looking at points $jE\lfloor \log \log x \rfloor$ with $|j| \leqslant \frac{\log x}{1000E(\log \log x)^2}$, one can simply look at points of the form $9(\log x)/20 + jE\lfloor \log \log x \rfloor$.

In the calculation of covariances that follows (11), one runs into the usual complication in the Rademacher case that the distribution of $f(n)n^{it}$ is not the same as the distribution of $f(n)$. Here this means that the covariances are not perfectly stationary, that is, they are a function of j and k rather than only a function of $j - k$. However, they are almost (that is, up to error terms) a function of $j - k$ and, in fact, exactly the same function one gets in the Steinhaus case, so one can adapt the analysis accordingly. Rather than giving further details, we refer the interested reader to Sections 6.1–6.2 of Harper [10] for an example of such an argument in an extremely similar context (basically the same as here except that only a special choice of u, slightly smaller than $\sqrt{2 \log \log x}$, is considered).

Acknowledgements

Conflict of Interest: None

<div align="center">

Appendix A.   Proofs of the Probability Results

</div>

We prove the results on Gaussian random walks that we stated in Section 3.

A.1.   Proof of Probability Result 1.   We shall prove that for large a, we have

$$\mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 10 \log j \; \forall 1 \leqslant j \leqslant n\right) \ll \min\left\{1, \frac{a}{\sqrt{n}}\right\},$$

which implies the upper bound part of Probability Result 1. The proof of the lower bound is very similar, replacing $a + 10 \log j$ by $a - 10 \log j$ and proving a lower bound $\gg \min\{1, \frac{c}{\sqrt{n}}\}$ for the probability.

We may assume that $\sqrt{n} \geqslant a$, otherwise the upper bound is trivial (and the lower bound follows immediately from (A1)). We shall temporarily adopt the convention that $\log_0 n = n$, and then for $k \in \mathbb{N}$, let $\log_k n := \log(\log_{k-1} n)$ denote the k-fold iterated logarithm.

We shall need the following results:

(A1) constant barrier case: under the hypotheses of Probability Result 1, for any $c \geqslant 1$, we have $\mathbb{P}(\sum_{m=1}^{j} G_m \leqslant c \; \forall 1 \leqslant j \leqslant n) \asymp \min\{1, \frac{c}{\sqrt{n}}\}$.

(A2) under the hypotheses of Probability Result 1, for any b and any $c \geqslant 1$, we have $\mathbb{P}(\max_{j \leqslant n} \sum_{m=1}^{j} G_m \in [b, b+c]) \ll c/\sqrt{n}$.

The result (A1) is standard, but, annoyingly, it seems difficult to find a reference when the G_m have unequal (though comparable) variances. The continuous time analogue, with $(\sum_{m=1}^{j} G_m)_{1 \leqslant j \leqslant n}$ replaced by a Brownian motion on the time interval $[0, \sum_{m=1}^{n} \mathbb{E}G_m^2]$, is completely standard (see, for example, Section 13.4 of Grimmett and Stirzaker [7]) and implies the lower bound in (A1) because if Brownian motion stays below c on the continuous interval, it certainly does so at the discrete points corresponding to $\sum_{m=1}^{j} G_m$. For the upper bound, it suffices to handle the case $c = 1$ since we can group the G_m into subsums of variance $\approx c^2$ (replacing n by n/c^2) and then multiply through by $1/c$. This case follows as in the proof of Lemma 5.1.8 of Lawler and Limic [19].

To deduce (A2), we can let $J = J(b)$ denote the smallest $1 \leqslant j \leqslant n$ at which $\sum_{m=1}^{j} G_m \geqslant b$, if such j exists, and then $\mathbb{P}(\max_{j \leqslant n} \sum_{m=1}^{j} G_m \in [b, b+c])$ is

$$= \sum_{k \leqslant n/2} \mathbb{P}\left(\max_{j \leqslant n} \sum_{m=1}^{j} G_m \in [b, b+c], \text{ and } J = k\right)$$

$$+ \mathbb{P}\left(\max_{j \leqslant n} \sum_{m=1}^{j} G_m \in [b, b+c], \text{ and } J > \frac{n}{2}\right)$$

$$\leqslant \sum_{k \leqslant n/2} \mathbb{P}(J = k)\mathbb{P}\left(\sum_{m=k+1}^{j} G_m \leqslant c \; \forall k + 1 \leqslant j \leqslant n\right)$$

$$+ \mathbb{P}\left(\max_{n/2 < j \leqslant n} \sum_{m=1}^{j} G_m \in [b, b+c]\right)$$

since the random variables $(G_m)_{m=k+1}^{n}$ are independent of the event $\{J = k\}$. By result (A1), the first sums here are $\ll \sum_{k \leqslant n/2} \mathbb{P}(J = k)c/\sqrt{n} \leqslant c/\sqrt{n}$. And if we let $M := \max_{n/2 < j \leqslant n} \sum_{n/2 < m \leqslant j} G_m$, then we have $\mathbb{P}(\max_{n/2 < j \leqslant n} \sum_{m=1}^{j} G_m \in [b, b+c]) = \mathbb{P}(\sum_{m \leqslant n/2} G_m \in [b - M, b - M + c])$. Since $\sum_{m \leqslant n/2} G_m$ is a Gaussian random variable with variance $\asymp n$, that is independent of M, the probability it lies in the interval $[b - M, b - M + c]$ is also $\ll c/\sqrt{n}$.

Using (A1) and (A2), we can prove the following lemma, from which the upper bound part of Probability Result 1 will quickly follow.

PROBABILITY LEMMA 1. *In the setting of Probability Result 1, and with the above conventions, the following is true. Uniformly for any $k \in \mathbb{N} \cup \{0\}$ such that the k-fold iterated logarithm $\log_k n \geqslant a/1000$, we have*

$$\mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 10 \log j \; \forall 1 \leqslant j \leqslant n\right)$$

$$= \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 10 \log(\min\{j, \log_k^{20} n\}) \; \forall 1 \leqslant j \leqslant n\right)$$

$$+ O\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{k} \frac{1}{\log_i^2 n}\right).$$

Proof of Probability Lemma 1. We shall prove the lemma by induction on k, for each given a. The conclusion of the lemma is trivial for $k = 0$. (And, incidentally, since we assume that $\sqrt{n} \geqslant a$, we certainly do have $\log_0 n = n \geqslant a/1000$.)

For the inductive step, suppose that the condition $\log_{k+1} n \geqslant a/1000$ is satisfied and that we have already established the lemma at the 'level' k. Let

$$S_{k+1} := \sum_{m \leqslant \log_{k+1}^{20} n} G_m, \quad \text{and let } s_{k+1}^2 := \sum_{m \leqslant \log_{k+1}^{20} n} \mathbb{E}G_m^2 \asymp \log_{k+1}^{20} n$$

denote its variance. Then conditioning on the behaviour of S_{k+1}, we can write $\mathbb{P}(\sum_{m=1}^{j} G_m \leqslant a + 10\log(\min\{j, \log_k^{20} n\})\ \forall j \leqslant n)$ as

$$\frac{1}{\sqrt{2\pi}s_{k+1}}$$

$$\times \int_{-\infty}^{\infty} e^{-\frac{u^2}{2s_{k+1}^2}} \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 10\log(\min\{j, \log_k^{20} n\})\ \forall j \leqslant n \mid S_{k+1} = u\right) du$$

$$= \frac{1}{\sqrt{2\pi}s_{k+1}}$$

$$\times \int_{-\infty}^{\infty} e^{-\frac{u^2}{2s_{k+1}^2}} \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 10\log j\ \forall 1 \leqslant j \leqslant \log_{k+1}^{20} n \mid S_{k+1} = u\right) \cdot$$

$$\cdot \mathbb{P}\left(\sum_{\log_{k+1}^{20} n < m \leqslant j} G_m \leqslant a + 10\log(\min\{j, \log_k^{20} n\}) - u\ \forall \log_{k+1}^{20} n < j \leqslant n\right) du,$$

where we used the independence of $\sum_{\log_{k+1}^{20} n < m \leqslant j} G_m$ and $\sum_{m \leqslant \log_{k+1}^{20} n} G_m$ to remove the conditioning on S_{k+1} from the final probability here.

Now we may restrict the integration to the range $(-\infty, a + 200\log_{k+2} n]$ since if u is larger, then the first conditional probability is zero because S_{k+1} would be too large. Meanwhile, in the part of the integral where $-\log_{k+1}^{4} n \leqslant u \leqslant a + 200\log_{k+2} n$, we can trivially upper bound the integrand by

$$e^{-\frac{u^2}{2s_{k+1}^2}} \mathbb{P}\left(\sum_{\log_{k+1}^{20} n < m \leqslant j} G_m \leqslant a + 10\log(\min\{j, \log_k^{20} n\}) + \log_{k+1}^{4} n\right.$$

$$\left. \forall \log_{k+1}^{20} n < j \leqslant n\right),$$

which is

$$\ll e^{-\frac{u^2}{2s_{k+1}^2}} (a + \log_{k+1}^{4} n)/\sqrt{n} \ll e^{-\frac{u^2}{2s_{k+1}^2}} \frac{\log_{k+1}^{4} n}{\sqrt{n}}$$

by result (A1) and the fact that $\log_{k+1} n \geqslant a/1000$. So the total contribution from this part of the integral is

$$\ll \frac{1}{\sqrt{2\pi}s_{k+1}} \int_{-\log_{k+1}^{4} n}^{a+200\log_{k+2} n} e^{-\frac{u^2}{2s_{k+1}^2}} \frac{\log_{k+1}^{4} n}{\sqrt{n}}\, du$$

$$\ll \frac{(a + \log_{k+1}^4 n) \log_{k+1}^4 n}{s_{k+1}} \frac{\log_{k+1}^4 n}{\sqrt{n}} \ll \frac{1}{\sqrt{n} \log_{k+1}^2 n}.$$

This may be absorbed into the 'big Oh' term in the statement of the lemma.

Finally, when $u \leqslant -\log_{k+1}^4 n$ in the integral, we have

$$\mathbb{P}\left(\sum_{\log_{k+1}^{20} n < m \leqslant j} G_m \leqslant a + 10 \log(\min\{j, \log_k^{20} n\}) - u \; \forall \log_{k+1}^{20} n < j \leqslant n\right)$$

$$= \mathbb{P}\left(\sum_{\log_{k+1}^{20} n < m \leqslant j} G_m \leqslant a + 10 \log(\log_{k+1}^{20} n) - u \; \forall \log_{k+1}^{20} n < j \leqslant n\right)$$

$$+ O\left(\frac{\log_{k+1} n}{\sqrt{n}}\right),$$

by result (A2) applied with $b = a + 10 \log(\log_{k+1}^{20} n) - u$ and $c = 10(\log(\log_k^{20} n) - \log(\log_{k+1}^{20} n)) \ll \log_{k+1} n$. Now we may trivially rewrite the preceding line as

$$\mathbb{P}\left(\sum_{\log_{k+1}^{20} n < m \leqslant j} G_m \leqslant a + 10 \log(\min\{j, \log_{k+1}^{20} n\}) - u \; \forall \log_{k+1}^{20} n < j \leqslant n\right)$$

$$+ O\left(\frac{\log_{k+1} n}{\sqrt{n}}\right),$$

and the crucial point is that, by result (A1) and the fact that $u \leqslant -\log_{k+1}^4 n$, the probability here is $\gg (\log_{k+1}^4 n)/\sqrt{n}$, so *we can replace the absolute error term* $O(\frac{\log_{k+1} n}{\sqrt{n}})$ *by a multiplicative factor* $1 + O(1/\log_{k+1}^3 n)$.

Putting everything back together, essentially reversing the above argument, we have shown that

$$\mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 10 \log(\min\{j, \log_k^{20} n\}) \; \forall j \leqslant n\right)$$

$$= \left(1 + O\left(\frac{1}{\log_{k+1}^3 n}\right)\right)$$

$$\times \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 10 \log(\min\{j, \log_{k+1}^{20} n\}) \; \forall j \leqslant n\right)$$

$$+ O\left(\frac{1}{\sqrt{n} \log_{k+1}^2 n}\right).$$

Finally, using result (A1), we have

$$
\frac{1}{\log_{k+1}^3 n} \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 10\log(\min\{j, \log_{k+1}^{20} n\}) \; \forall j \leqslant n\right)
$$

$$
\ll \frac{a + \log_{k+2} n}{\sqrt{n}\log_{k+1}^3 n} \ll \frac{1}{\sqrt{n}\log_{k+1}^2 n},
$$

which gives an acceptable error term and completes our induction. □

Proof of Probability Result 1, upper bound. We apply Probability Lemma 1 with k chosen as large as possible so that $\log_k n \geqslant a/1000$ but $\log_{k+1} n < a/1000$. This yields that $\mathbb{P}(\sum_{m=1}^{j} G_m \leqslant a + 10\log j \; \forall 1 \leqslant j \leqslant n)$ is

$$
\leqslant \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + 200\log_{k+1} n \; \forall 1 \leqslant j \leqslant n\right) + O\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{k}\frac{1}{\log_i^2 n}\right)
$$

$$
\ll \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant 2a \; \forall 1 \leqslant j \leqslant n\right) + \frac{1}{\sqrt{n}} \ll \frac{a}{\sqrt{n}},
$$

by (A1). □

A.2. Proof of Probability Result 2.

The upper bound part of Probability Result 2 follows immediately from Probability Result 1.

To prove the lower bound part of the result, recalling that $g(j) \leqslant -Bj$ and $h(j) \geqslant -10\log j$, we see that $\mathbb{P}(g(j) \leqslant \sum_{m=1}^{j} G_m \leqslant \min\{a, Bj\} + h(j) \; \forall 1 \leqslant j \leqslant n)$ is

$$
\geqslant \mathbb{P}\left(-Bj \leqslant \sum_{m=1}^{j} G_m \leqslant \min\{a, Bj\} + h(j) \; \forall 1 \leqslant j \leqslant n\right)
$$

$$
\geqslant \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + h(j) \; \forall 1 \leqslant j \leqslant n\right)
$$

$$
- \sum_{k=1}^{n} \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + h(j) \; \forall 1 \leqslant j \leqslant n, \text{ and } \sum_{m=1}^{k} G_m \leqslant -Bk\right)
$$

$$
- \sum_{k=1}^{n} \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + h(j) \; \forall 1 \leqslant j \leqslant n, \text{ and } \sum_{m=1}^{k} G_m \geqslant Bk - 10\log k\right).
$$

By Probability Result 1, the first term here is $\gg \min\{1, \frac{a}{\sqrt{n}}\}$, and we will show that the sum of the subtracted terms is smaller than this, provided B is fixed large enough.

Indeed, since $(1/20)k \leqslant \sum_{m=1}^{k} \mathbb{E}G_m^2 \leqslant 20k$, if we condition on $\sum_{m=1}^{k} G_m$, we see $\mathbb{P}(\sum_{m=1}^{j} G_m \leqslant a + h(j) \; \forall 1 \leqslant j \leqslant n, \text{ and } \sum_{m=1}^{k} G_m \leqslant -Bk)$ is

$$\ll \frac{1}{\sqrt{k}} \int_{-\infty}^{-Bk} e^{-u^2/40k} \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + h(j) \; \forall 1 \leqslant j \leqslant n \mid \sum_{m=1}^{k} G_m = u\right) du$$

$$\leqslant \frac{1}{\sqrt{k}} \int_{-\infty}^{-Bk} e^{-u^2/40k} \mathbb{P}\left(\sum_{m=k+1}^{j} G_m \leqslant a + h(j) - u \; \forall k+1 \leqslant j \leqslant n\right) du.$$

By Probability Result 1, provided that $k \leqslant n/2$ (so that $n - (k+1) \gg n$) this is all

$$\ll \frac{1}{\sqrt{k}} \int_{-\infty}^{-Bk} e^{-u^2/40k} \min\left\{1, \frac{a + |u|}{\sqrt{n}}\right\} du$$

$$\ll \min\left\{1, \frac{a}{\sqrt{n}}\right\} \frac{1}{\sqrt{k}} \int_{-\infty}^{-Bk} e^{-u^2/40k} |u| \, du$$

$$\ll \min\left\{1, \frac{a}{\sqrt{n}}\right\} \sqrt{k} e^{-B^2 k/40}.$$

If $k > n/2$, we can upper bound the probability in the integral trivially by 1 and just use the fact that

$$\frac{1}{\sqrt{k}} \int_{-\infty}^{-Bk} e^{-u^2/40k} \, du \ll e^{-B^2 k/40} \ll e^{-B^2 n/80}.$$

Summing over $1 \leqslant k \leqslant n$, the contribution from all these terms will indeed be small compared with $\min\{1, \frac{a}{\sqrt{n}}\}$, provided B is fixed large enough.

Similarly, we have

$$\mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + h(j) \; \forall 1 \leqslant j \leqslant n, \text{ and } \sum_{m=1}^{k} G_m \geqslant Bk - 10\log k\right)$$

$$\ll \frac{1}{\sqrt{k}} \int_{Bk-10\log k}^{a+h(k)} e^{-u^2/40k} \mathbb{P}\left(\sum_{m=1}^{j} G_m \leqslant a + h(j) \; \forall 1 \leqslant j \leqslant n \mid \sum_{m=1}^{k} G_m = u\right) du$$

$$\leqslant \frac{1}{\sqrt{k}} \int_{Bk-10\log k}^{a+h(k)} e^{-u^2/40k} \mathbb{P}\left(\sum_{m=k+1}^{j} G_m \leqslant a + h(j) - Bk + 10\log k \; \forall k+1 \leqslant j \leqslant n\right) du,$$

which is $\ll \min\{1, \frac{a}{\sqrt{n}}\}e^{-(Bk-10\log k)^2/40k}$ when $k \leqslant n/2$ and is $\ll e^{-(Bk-10\log k)^2/40k} \ll e^{-B^2 n/100}$ (say) when $k > n/2$. Again summing over $1 \leqslant k \leqslant n$, the contribution from all these terms will be negligible compared with $\min\{1, \frac{a}{\sqrt{n}}\}$ if B is large enough. $\qquad\square$

References

[1] L.-P. Arguin, D. Belius and A. J. Harper, 'Maxima of a randomized Riemann zeta function, and branching random walks', *Ann. Appl. Probab.* **27**(1) (2017), 178–215.

[2] J. Barral, A. Kupiainen, M. Nikula, E. Saksman and C. Webb, 'Basic properties of critical lognormal multiplicative chaos', *Ann. Probab.* **43**(5) (2015), 2205–2249.

[3] N. Berestycki, 'An elementary approach to Gaussian multiplicative chaos', *Electron. Commun. Probab.* **22**(Paper No. 27) (2017), 12 pp.

[4] A. Bondarenko and K. Seip, 'Helson's problem for sums of a random multiplicative function', *Mathematika* **62**(1) (2016), 101–110.

[5] S. Chatterjee and K. Soundararajan, 'Random multiplicative functions in short intervals', *Int. Math. Res. Not. IMRN* **2012** (2012), 479–492.

[6] B. Duplantier, R. Rhodes, S. Sheffield and V. Vargas, 'Renormalization of critical Gaussian multiplicative chaos and KPZ relation', *Comm. Math. Phys.* **330**(1) (2014), 283–330.

[7] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 3rd edn, (Oxford University Press, New York, 2001).

[8] A. Gut, *Probability: A Graduate Course*, 2nd edn, (Springer Texts in Statistics, New York, 2013).

[9] G. Halász, 'On random multiplicative functions', in *Hubert Delange Colloquium, (Orsay, 1982)*, Publications Mathématiques d'Orsay, **83** (Univ. Paris XI, Orsay, 1983), 74–96.

[10] A. J. Harper, 'Bounds on the suprema of Gaussian processes, and omega results for the sum of a random multiplicative function', *Ann. Appl. Probab.* **23**(2) (2013), 584–616.

[11] A. J. Harper, 'On the limit distributions of some sums of a random multiplicative function', *J. reine angew. Math.* **678** (2013), 95–124.

[12] A. J. Harper, 'Moments of random multiplicative functions, II: High moments', *Algebra Number Theory*, to appear.

[13] A. J. Harper, A. Nikeghbali and M. Radziwiłł, 'A note on Helson's conjecture on moments of random multiplicative functions', in *Analytic Number Theory* (Springer, Cham, 2015), 145–169.

[14] W. Heap and S. Lindqvist, 'Moments of random multiplicative functions and truncated characteristic polynomials', *Q. J. Math.* **67**(4) (2016), 683–714.

[15] H. Helson, 'Hankel forms', *Studia Math.* **198**(1) (2010), 79–84.

[16] B. Hough, 'Summation of a random multiplicative function on numbers having few prime factors', *Math. Proc. Cambridge Philos. Soc.* **150** (2011), 193–214.

[17] Y. Hu and Z. Shi, 'Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees', *Ann. Probab.* **37**(2) (2009), 742–789.

[18] Y.-K. Lau, G. Tenenbaum and J. Wu, 'On mean values of random multiplicative functions', *Proc. Amer. Math. Soc.* **141** (2013), 409–420. Also see www.iecl.univ-lorraine.fr/∼Gerald.Tenenbaum/PUBLIC/Prepublications_et_publications/RMF.pdf for some corrections to the published version.

[19] G. F. Lawler and V. Limic, *Random Walk: a Modern Introduction*, 1st edn, (Cambridge University Press, Cambridge, 2010).

[20] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, 1st edn, (Cambridge University Press, Cambridge, 2007).

[21] N. Ng, 'The distribution of the summatory function of the Möbius function', *Proc. Lond. Math. Soc. (3)* **89**(3) (2004), 361–389.

[22] G. Reinert and A. Röllin, 'Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition', *Ann. Probab.* **37**(6) (2009), 2150–2173.

[23] R. Rhodes and V. Vargas, 'Gaussian multiplicative chaos and applications: A review', *Probab. Surv.* **11** (2014), 315–392.

[24] S. M. Sadikova, 'Two-dimensional analogues of an inequality of Esseen with applications to the Central Limit Theorem', *Theory Probab. Appl.* **11** (1966), 325–335.

[25] E. Saksman and K. Seip, 'Integral means and boundary limits of Dirichlet series', *Bull. Lond. Math. Soc.* **41**(3) (2009), 411–422.

[26] E. Saksman and K. Seip, 'Some open questions in analysis for Dirichlet series', in *Recent Progress on Operator Theory and Approximation in Spaces of Analytic Functions*, Contemp. Mathematics, 679 (American Mathematical Society, Providence, RI, 2016), 179–191.

[27] E. Saksman and C. Webb, 'Multiplicative chaos measures for a random model of the Riemann zeta function', Preprint available online at http://arxiv.org/abs/1604.08378.

[28] E. Saksman and C. Webb, 'The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line', Preprint available online at https://arxiv.org/abs/1609.00027.

[29] K. Soundararajan, 'Moments of the Riemann zeta function', *Ann. of Math.* **170** (2009), 981–993.

[30] M. J. G. Weber, 'L^1-Norm of Steinhaus chaos on the polydisc', *Monatsh. Math.* **181**(2) (2016), 473–483.

[31] A. Wintner, 'Random factorizations and Riemann's hypothesis', *Duke Math. J.* **11** (1944), 267–275.