Almost sure lower bounds for a model problem for multiplicative chaos in number theory

Maxim Gerspach

KTH Royal Institute of Technology, Lindstedtsvägen, Stockholm, Sweden

Correspondence
Maxim Gerspach, Department of Mathematics, KTH Royal Institute of Technology, Lindstedtsvägen 25, 114 28 Stockholm, Sweden.
Email: gerspach@kth.se

Funding information
Swedish Research Council, Grant/Award Number: 2016-05198; Göran Gustafsson Foundation

1 INTRODUCTION

Let \((X(k))_{k \geq 1}\) be a sequence of independent standard complex Gaussian random variables, so that their real and imaginary parts are independent, centered and have variance \(\frac{1}{2}\). Consider the sequence of random variables \((A(n))_{n \geq 0}\) defined through the following formal identity of power series:

\[
\exp \left( \sum_{k \geq 1} \frac{X(k)}{\sqrt{k}} z^k \right) = \sum_{n \geq 0} A(n) z^n.
\]

These random variables have recently been considered as an analogue of a (Steinhaus) random multiplicative function that seems to be rich enough to share most of their properties,
while often allowing for simpler proofs thereof. Recall that a Steinhaus random multiplicative function is a (random) completely multiplicative function \( f : \mathbb{N} \to \mathbb{C} \) such that its values at the primes are independent and uniformly distributed on the complex unit circle (Steinhaus distributed).

In this setting, there has been a lot of attention in recent years regarding the moments and almost sure bounds for mean values of random multiplicative functions. In particular, Harper proved in [4] that for any function \( V(x) \) tending to infinity with \( x \) there are, almost surely, arbitrarily large values of \( x \) such that

\[
\left| \frac{1}{\sqrt{x}} \sum_{n \leq x} f(n) \right| \geq \frac{(\log \log x)^{1/4}}{V(x)}.
\] (1.1)

Regarding upper bounds, the best known result to this date is due to Lau, Tenenbaum and Wu [5] and states that, almost surely, we have

\[
\frac{1}{\sqrt{x}} \sum_{n \leq x} f(n) = O_{\varepsilon}((\log \log x)^{2+\varepsilon}).
\]

Moreover, Mastrostefano [6] recently proved an upper bound essentially matching the one of Harper for the same sum when restricted to integers that possess a large prime factor. More precisely, denoting by \( P(n) \) the largest prime factor of an integer \( n \), he showed that we almost surely have

\[
\frac{1}{\sqrt{x}} \sum_{\substack{n \leq x \\colon P(n) > \sqrt{x}}} f(n) = O_{\varepsilon}((\log \log x)^{1/4+\varepsilon}).
\]

These almost sure bounds may be compared to the first moment. Namely, as a corollary to [3, Theorem 1], we have that

\[
\mathbb{E}\left[ \left| \frac{1}{\sqrt{x}} \sum_{n \leq x} f(n) \right| \right] \simeq \frac{1}{(\log \log x)^{1/4}}.
\]

This (suspected) discrepancy of a factor \( \sqrt{\log \log x} \) between the first moment and almost sure behaviour is akin to the law of the iterated logarithm for independent random variables. We refer the reader to the discussion in [4] for more details on this comparison.

Inspired by this result, Soundararajan and Zaman consider in [7] the moments of \( A(n) \), and showed that they behave rather similar to the random multiplicative setting. They proved in particular that we have

\[
\mathbb{E}[|A(n)|] \simeq \frac{1}{(\log n)^{1/4}}.
\]

They further raise the question whether one can obtain analogues of the almost sure results on random multiplicative functions for \( A(n) \). The main goal of this work is to prove that this is indeed the case, that is, to prove the following analogue of (1.1).
Theorem 1. For any function $V(n)$ tending to infinity with $n$, there almost surely exist arbitrarily large values of $n$ for which

$$|A(n)| \geq \frac{(\log n)^{1/4}}{V(n)}.$$

The link between mean values of random multiplicative functions and the random variables $A(n)$ is perhaps not obvious at first sight. As explained by Soundararajan and Zaman in [7], it is more apparent in a function field setting of $\mathbb{F}_q[t]$, especially in the limit as $q \to \infty$. This analogy roughly corresponds to imposing that all primes $e^k < p \leq e^{k+1}$ shall have ‘size’ $e^k$, which is more sensible in a function field setting, where one tends to have many irreducible polynomials of the same degree. We refer the reader to their introduction for the details of this analogy. In addition, we will give a more detailed heuristic in the next section after introducing further definitions.

As in the work of Harper [4], we will prove this by showing a lower bound for the maximum of $|A(n)|$ on adequate intervals of $n$, to hold with high probability, and then deduce the Theorem from an invocation of the first Borel–Cantelli Lemma.

Theorem 2. Uniformly for all large $N$ and all $1 \leq W \leq \frac{1}{10} \log \log N$, we have

$$\max_{\frac{8N}{7} \leq n \leq \frac{4N}{3}} |A(n)| \geq \frac{(\log N)^{1/4}}{e^{W/5}}$$

with probability $\geq 1 - O(e^{-W/10})$.

In fact, the deduction of the main result from this theorem follows [4] to the letter aside from some choices of constants, but will be included here for completeness.

Proof of Theorem 1 assuming Theorem 2. Assume without loss of generality that $V(n) \leq (\log n)^{1/10}$ for sufficiently large $n$. Set

$$W(N) := \min_{n \in \left[\frac{8N}{7}, \frac{4N}{3}\right]} \frac{5 \log V(n)}{6} - 1,$$

so that $W(N)$ goes to infinity with $N$, but also satisfies $W(N) \leq \frac{1}{10} \log \log N$ when $N$ is sufficiently large. Hence, by Theorem 2, the probability that for a given sufficiently large $N$ there is no $n \in \left[\frac{8N}{7}, \frac{4N}{3}\right]$ such that

$$|A(n)| \geq \frac{(\log N)^{1/4}}{e^{W(N)/5}} \geq \frac{(\log n)^{1/4}}{V(n)}$$

is $O(e^{-W(N)/10})$ (since $n$ can also be assumed sufficiently large). Hence, by choosing an adequate subsequence of values of $N$ we can ensure these probabilities to be summable. Thus, the first Borel–Cantelli Lemma tells us that, almost surely, only finitely many of these complementary events occur on our subsequence. Excluding these finitely many events, we have (almost surely) found our sequence of $n$ going to infinity that satisfies the claimed inequality. \hfill \Box

Overall, our proof will closely follow the works of Harper [3, 4], but there will be some simplifications arising from the fact that we are often dealing with Gaussian random variables in
place of random variables that are approximately Gaussian. This will allow the argument to be essentially self-contained with somewhat less effort than in these works. We will keep many of the constants analogous to the ones that arise there, mostly for easier comparison, since we do have even more flexibility in their choices. We will moreover employ some of the ideas from the work of Soundararajan and Zaman [7].

One of the differences to the random multiplicative function setting is that one needs to restrict to a (geometric) subsequence of values of $x$ in $[X^{8/7}, X^{4/3}]$ in order to have any hope that $\frac{1}{\sqrt{x}} \sum_{n \leq x} f(n)$ at different $x$ have small correlations. In our setting, this would correspond to taking a linear (arithmetic) subsequence, and it turns out that one can just as well use all values of $A(n)$ at this step.

It seems very much a possibility that one can also transfer the aforementioned almost sure upper bounds [5, 6] to the random variables $A(N)$ by going through the respective arguments and making similar adjustments as in this work or as in [7].

Finally, we also want to point the reader towards the work [1], where the authors numerically study the asymptotics of the first moment of $A(n)$ (which are not known to exist) by developing an algorithm that allows a more effective computation.

2 | PRELIMINARIES AND BASIC ESTIMATES

Let $t \in \mathbb{R}$ and $m \in \mathbb{N}$. Let $\sigma \geq 0$, which the reader should think of as small compared to $N$, say $o(1)$. In the following, we will frequently be working with the random variables

$$Z_t(m) = Z_{t, \sigma}(m) := \sum_{e^{m-1} < k \leq e^m} \Re X(k) e^{ikt} \sqrt{k \sigma},$$

where (somewhat arbitrarily) we will throughout include the term $k = 1$ in the sum when $m = 1$, and in corresponding sums. We also set $Z(m) := Z_0(m)$. We note that these random variables have mean 0 and variance

$$\sigma^2_m = \sigma_{m, \sigma}^2 := \mathbb{E}[Z_t(m)^2] = \sum_{e^{m-1} < k \leq e^m} \frac{1}{2k e^{2k \sigma}},$$

independent of $t$ (we hope that little confusion can arise from these two uses of the letter $\sigma$). Moreover, we define

$$\rho_{m, t} \sigma^2_m = \rho_{m, t, \sigma} \sigma^2_m := \mathbb{E}[Z(m)Z_t(m)] = \sum_{e^{m-1} < k \leq e^m} \frac{\cos(kt)}{2k e^{2k \sigma}}$$

and remark that $\mathbb{E}[Z_t(m)Z_u(m)] = \rho_{m, u-t} \sigma^2_m$.

It will be useful to derive an explicit expression for $A(n)$ (which also appears in [7]). Note that in the implicit definition, we can expand the exponential and see that

$$\sum_{n \geq 0} A(n) z^n = \sum_{m \geq 0} \frac{1}{m!} \left( \sum_{k \geq 1} \frac{X(k)}{\sqrt{k}} z^k \right)^m = \sum_{m \geq 0} \frac{1}{m!} \sum_{n \geq 0} z^n \sum_{k_1 + \cdots + k_m = n} \frac{X(k_1) \cdots X(k_m)}{\sqrt{k_1 \cdots k_m}}.$$
Next, we can order the $k_i$ decreasingly and note that the number of ways of rearranging them is

$$m! \prod_{j \geq 1} (\# \{k_i = j\})!.$$ 

In this way we have linked the previous expression to partitions of $n$. Let us write $(\lambda_1, \lambda_2, \ldots) = \lambda \rightarrow n$ if $\lambda$ is a partition of $n$, so that we have $\lambda_1 + \lambda_2 + \cdots = n$ and $\lambda_1 \geq \lambda_2 \geq \cdots$. Simply comparing the coefficients in the respective power series we see that

$$A(n) = \sum_{\lambda \rightarrow n} \prod_{j \geq 1} \left( \frac{X(j)}{\sqrt{j}} \right)^{m_j} \frac{1}{m_j!} =: \sum_{\lambda \rightarrow n} a(\lambda),$$

where we denote by $m_j = m_j(\lambda)$ the number of parts of $\lambda$ of size $j$.

Further, for $K > 1$ real and $z \in \mathbb{C}$, we set

$$F_K(z) := \exp \left( \sum_{k \leq K} X(k) \frac{z^k}{\sqrt{k}} \right).$$

The reader should think of $z$ as being on or close to the complex unit circle. Most of the time we will be able to assume that $\log K$ is an integer, which implies in particular that

$$|F_K(e^{-\sigma+i})| = \exp \left( \sum_{m=1}^{\log K} Z_t(m) \right).$$

Moreover, one verifies by the same computation as for $A(n)$ that we have

$$F_K(z) = \sum_{n \geq 0} \left( \sum_{\lambda \rightarrow n, \lambda_1 \leq K} a(\lambda) \right) z^n.$$

In particular, these coefficients are simply $A(n)$ whenever $n \leq K$.

The function $F_K(z)$ can be connected to $A(N)$ through an application of Cauchy’s formula. This is analogous to how the average of a random multiplicative function can be related to its (partial) random Euler product through Perron’s formula. To explain this in more detail, note that for $N \leq K$, Cauchy’s formula gives

$$A(N) = \frac{1}{2\pi i} \int_{|z|=1} F_K(z) \frac{dz}{z^{N+1}}. \tag{2.1}$$

Let us compare this to

$$\frac{1}{\sqrt{e^N}} \sum_{n \leq e^N} f(n).$$

Define

$$F_K(s) := \prod_{p \leq e^K} \left( 1 - \frac{f(p)}{p^s} \right)^{-1} = \sum_{p(n) \leq e^K} \frac{f(n)}{n^s}.$$
to be the (random and partial) Euler product associated to \( f \). Perron’s formula tells us that

\[
\frac{1}{\sqrt{e^N}} \sum_{n \leq e^N} f(n) \approx \frac{1}{2\pi i e^{N/2}} \int_{c} \bar{F}_K(s) e^{sN} \frac{ds}{s},
\]

where \( (c) \) denotes the line with real part equal to \( c \) (and in fact we have equality, but that is not important for this informal discussion). One then verifies (as we will later on) that this integral can be cut off at \( \frac{N}{2} \), and that one can moreover shift the resulting integral to the line \( \Re s = \frac{1}{2} \), with negligible error in both cases. Thus,

\[
\frac{1}{\sqrt{e^N}} \sum_{n \leq e^N} f(n) \approx \frac{1}{2\pi} \int_{|t| \leq e^{N/4}} \bar{F}_K \left( \frac{1}{2} + it \right) e^{iNt} \frac{dt}{\frac{1}{2} + it}.
\]

Ideally one would like to cut off this integral at a constant (such as \( \pi \)), and one can show that with high probability one can at least restrict to \( |t| < N^2 \) (and even \( (\log N)^2 \) in an appropriate sense). Comparing this to (2.1) after setting \( z = e^{-it} \), these expressions seem to look rather similar if one chooses to believe that \( F_K(e^{-it}) \) and \( \bar{F}_K(1/2 + it) \) behave in a similar fashion. And indeed, note that

\[
\log \bar{F}_K(1/2 + it) = - \sum_{p \leq e^k} \log \left( 1 - \frac{f(p)}{p^{1/2 + it}} \right) \approx \sum_{p \leq e^k} \frac{f(p)}{p^{1/2 + it}} = \sum_{k \leq K} \sum_{e^{k-1} < p \leq e^k} \frac{f(p)}{p^{1/2 + it}}.
\]

For any \( t \) and sufficiently large \( k \), these sums over \( p \) are a sum of many random variables with mean 0 and variance \( \frac{1}{p} \) (whose real and imaginary parts are uncorrelated and have variance \( \frac{1}{2p} \)), thus the central limit theorem implies that for each \( k \),

\[
\sum_{e^{k-1} < p \leq e^k} \frac{f(p)}{p^{1/2 + it}}
\]

is approximately a (complex) Gaussian with mean 0 and variance

\[
\sum_{e^{k-1} < p \leq e^k} \frac{1}{p} \approx \log k - \log(k - 1) \approx \frac{1}{k}
\]

(whose real and imaginary part are independent real Gaussians with mean 0 and variance \( \approx \frac{1}{2k} \)). The same holds true for \( \frac{X(k)}{\sqrt{k}} e^{itk} \) for any \( t \). In fact one can also identify the correlations of the respective processes for different values of \( t \).

Denote for \( t \in \mathbb{R} \) by \( \| t \| \) the distance from \( t \) to a nearest integer multiple of \( 2\pi \). We continue by recording the following elementary estimates.

**Proposition 2.1.** For \( \frac{1}{2} \leq x < y \), \( \sigma \geq 0 \) and \( t \in (0, 2\pi) \), we have

\[
\left| \sum_{x < k \leq y} \frac{e^{itk}}{ke^{2k\sigma}} \right| \leq \frac{3\pi}{x \| t \|}.
\]
Proof. Note that
\[ \left| \sum_{k \leq u} e^{ik} \right| \leq \frac{2}{|1 - e^{it}|} \leq \frac{\pi}{\|t\|} \]
for all (say) \( u \geq \frac{1}{2} \) and \( t \in (0, 2\pi) \). The claim then follows from partial summation. \qed

**Proposition 2.2.** Let \( t \in \mathbb{R}, m \in \mathbb{N} \) and \( \sigma \geq 0 \). Then we have
\[ e^{-2\sigma e^m} \left( \frac{1}{2} + O(e^{-m}) \right) \leq \sigma^2_m \leq e^{-2\sigma e^{m-1}} \left( \frac{1}{2} + O(e^{-m}) \right). \]
If moreover \( 0 \leq \sigma \leq 10e^{-m} \) (say), then we have \( \sigma^2_m \asymp 1 \) and
\[ \rho_{m,t} \ll \frac{1}{\|t\|e^m}. \]

**Proof.** Recall that we are still following the convention that the term corresponding to \( k = 1 \) is included in the sum when \( m = 1 \).

To deduce the first part, one simply bounds \( e^{-2k\sigma} \) trivially and uses that
\[ \sum_{e^{m-1} < k < e^m} \frac{1}{2k} = \frac{1}{2} + O(e^{-m}). \]

We note that this sum is in fact always between \( \frac{1}{4} \) and 1, say, as follows, for example, from [8, Theorem] and a short computation. This certainly gives the next claim after bounding \( e^{-2k\sigma} \) trivially again.

As for the estimate on \( \rho_{m,t} \), note first that
\[ \rho_{m,t} \asymp \rho_{m,t} \sigma^2_m = \Re \sum_{e^{m-1} < k < e^m} \frac{e^{ikt}}{2ke^{2k\sigma}}. \]

One then applies partial summation, using again that
\[ \sum_{k \leq u} e^{ikt} \ll \frac{1}{\|t\|} \]
and the claim follows. \qed

**Proposition 2.3.** For any \( m \in \mathbb{N} \), real \( \alpha_1, \ldots, \alpha_m \) and \( t_1, \ldots, t_m \), for \( \frac{1}{2} \leq x \leq y \) (say) and \( \sigma \geq 0 \), we have
\[ \mathbb{E} \left[ \exp \left( \sum_{j=1}^{m} 2\alpha_j \sum_{x < k \leq y} \Re \mathbb{X}(k)e^{ikt_j} \frac{\sqrt{k}e^{ik\sigma}}{ke^{2ik\sigma}} \right) \right] = \exp \left( \sum_{j=1}^{m} \sum_{\ell=1}^{m} \sum_{x < k \leq y} \alpha_j \alpha_\ell \cos \left( k(t_j - t_\ell) \right) \frac{ke^{2ik\sigma}}{ke^{2ik\sigma}} \right). \]
In particular, uniformly for \( \frac{1}{2} \leq x \leq y, \sigma \geq 0 \) and \( t_1, t_2 \in \mathbb{R} \), we have

\[
\mathbb{E} \left[ \exp \left( \sum_{x < k \leq y} \frac{\Re X(k) e^{ikt_1}}{\sqrt{k e^{i k \sigma}}} + \sum_{x < k \leq y} \frac{\Re X(k) e^{ikt_2}}{\sqrt{k e^{i k \sigma}}} \right) \right] \ll \sqrt{\frac{y}{x}} \min \left\{ \sqrt{\frac{y}{x}}, \sqrt{\frac{1}{\|t_2 - t_1\| x}} \right\}.
\]

Proof. The first part of this statement is completely elementary since we are dealing with true Gaussian random variables, and only included for easier comparison to the work of Harper [4]. Namely, one notes that the sum

\[
\sum_{j=1}^m 2 \alpha_j \sum_{x < k \leq y} \frac{\Re X(k) e^{ikt_j}}{\sqrt{k e^{i k \sigma}}}
\]

inside the exponential is still Gaussian, with mean 0 and variance

\[
\sum_{x < k \leq y} \mathbb{E} \left[ \left( 2 \sum_{j=1}^m \alpha_j \frac{\Re X(k) e^{ikt_j}}{\sqrt{k e^{i k \sigma}}} \right)^2 \right] = 4 \sum_{j=1}^m \sum_{\ell=1}^m \alpha_j \alpha_\ell \frac{\mathbb{E} \left[ (\Re X(k) e^{ikt_j})(\Re X(k) e^{ikt_\ell}) \right]}{k e^{2ik\sigma}}.
\]

One then verifies that

\[
\mathbb{E} \left[ (\Re X(k) e^{ikt_j})(\Re X(k) e^{ikt_\ell}) \right] = \frac{\cos \left( k(t_\ell - t_j) \right)}{2},
\]

for example, by using that \( \Re X(k) e^{ikt} = (\Re X(k)) \cos(kt) - (\Im X(k)) \sin(kt) \) as well as \( \mathbb{E}[\Re X(k) \Im X(k)] = 0 \) and \( \cos(x - y) = \cos x \cos y + \sin x \sin y \). Noting that if \( G \) is a Gaussian with mean 0 and variance \( \sigma^2 \) then \( \mathbb{E}[\exp(G)] = \exp(\sigma^2/2) \) gives the first claim.

As for the second part, note that the first part, with \( \alpha_1 = \alpha_2 = \frac{1}{2} \), implies

\[
\mathbb{E} \left[ \exp \left( \sum_{x < k \leq y} \frac{\Re X(k) e^{ikt_1}}{\sqrt{k e^{i k \sigma}}} + \sum_{x < k \leq y} \frac{\Re X(k) e^{ikt_2}}{\sqrt{k e^{i k \sigma}}} \right) \right] = \exp \left( \sum_{x < k \leq y} \frac{1 + \cos(k(t_2 - t_1))}{2ke^{2k\sigma}} \right).
\]

Moreover, we certainly have

\[
\sum_{x < k \leq y} \frac{1}{2ke^{2k\sigma}} \leq \frac{1}{2}(\log y - \log x) + O(1),
\]

which gives rise to the factor \( \sqrt{\frac{y}{x}} \) in our estimate and also the factor \( \sqrt{\frac{y}{x}} \) inside the minimum, simply by bounding the cosine by 1. We may thus assume from now on that \( \|t_2 - t_1\| \geq \frac{1}{y} \), and note that by Proposition 2.1 we have

\[
\sum_{\frac{1}{\|t_2 - t_1\|} < k \leq y} \frac{\cos(k(t_2 - t_1))}{2ke^{2k\sigma}} = \Re \sum_{\frac{1}{\|t_2 - t_1\|} < k \leq y} \frac{e^{ikt_1(t_2 - t_1)}}{2ke^{2k\sigma}} \ll 1.
\]
If \( \|t_2 - t_1\| \geq \frac{1}{x} \), this already gives the claim. Otherwise, we can simply bound

\[
\sum_{x < k \leq \frac{1}{\|t_2 - t_1\|}} \cos \left( k(t_2 - t_1) \right) \leq \frac{1}{2} \left( \log 1/\|t_2 - t_1\| - \log x \right) + O(1),
\]

which completes the proof. \(\square\)

\section{Probabilistic Tools}

We begin by recording the following results, which are simply Probability Results 1 and 2 from \cite{3} and Probability Result 1 from \cite{2} as well as Normal Comparison Result 1 from \cite{4}.

**Proposition 3.1** (Harper). Let \( a \geq 1 \). For any integer \( n \geq 1 \), let \( G_1, \ldots, G_n \) be independent real Gaussian random variables, each having mean zero and variance between \( \frac{1}{20} \) and 20, say. Let \( h \) be a function such that \( |h(j)| \leq 10 \log j \). Then we have

\[
P \left[ \sum_{m=1}^{j} G_m \leq a + h(j) \forall 1 \leq j \leq n \right] \approx \min \left\{ 1, \frac{a}{\sqrt{n}} \right\}.
\]

**Proposition 3.2** (Harper). There is an absolute constant \( B \) such that the following is true. Let \( a \) and \( n \) be large, \( n \) being an integer, and let \( G_1, \ldots, G_n \) be Gaussian random variables, each with mean between \( \frac{1}{20} \) and 20, say. Then uniformly for any functions \( h(j) \) respectively \( g(j) \) satisfying \( |h(j)| \leq 10 \log j \) resp. \( g(j) \leq -Bj \) for all \( 1 \leq j \leq n \), we have

\[
P \left[ g(j) \leq \sum_{m=1}^{j} G_m \leq \min\{a, Bj\} + h(j) \forall 1 \leq j \leq n \right] \approx \min \left\{ 1, \frac{a}{\sqrt{n}} \right\}.
\]

The reader should think of this barrier condition essentially as a barrier of the shape \( \sum_{m=1}^{j} G_m \leq a \) for all \( 1 \leq j \leq n \). The random walk typically varies on a scale of \( \sqrt{j} \), so that the imposed lower bound is essentially irrelevant. On the other hand, the shift by \( h(j) \) is so small compared to the typical variation of the random walk that we should not expect it to make a big difference either, as is indeed reflected by the proposition. The added flexibility will nonetheless be rather useful for our purposes.

**Proposition 3.3** (Harper). Let \( a, b \in \mathbb{R} \) and \( n \in \mathbb{N} \) all be large, and let \( G_1, \ldots, G_n \) be independent real Gaussian random variables, each having mean zero and variance between \( \frac{1}{20} \) and 20, say. Then we have the uniform upper bound

\[
P \left[ \sum_{m=1}^{j} G_m \leq a \forall 1 \leq j \leq n, \quad a - b \leq \sum_{m=1}^{n} G_m \leq a \right] \ll \min \left\{ 1, \frac{a}{\sqrt{n}} \right\} \min \left\{ 1, \frac{b}{\sqrt{n}} \right\}^2.
\]

While the previous statements all concerned barrier conditions on random walks, the last result will pertain to estimating the maximum of a collection of Gaussian random variables with
small correlations. It generalises the fairly well-known result that the maximum of $n$ independent standard Gaussians is close to $\sqrt{2 \log n}$ with high probability.

**Proposition 3.4** (Harper). There exists an absolute constant $c > 0$ such that the following holds: suppose that $n \geq 2$, and that $\varepsilon > 0$ is sufficiently small. Let $G_1, \ldots, G_n$ be mean 0, variance 1, real, jointly Gaussian random variables, and suppose that $\mathbb{E}[G_i G_j] \leq \varepsilon$ whenever $i \neq j$. Then for any $100 \varepsilon \leq \delta \leq \frac{1}{100}$, we have

$$
P\left[ \max_{1 \leq i \leq n} G_i \leq \sqrt{(2 - \delta) \log n} \right] \ll \exp\left( -\frac{cn^\delta/20}{\sqrt{\log n}} \right) + n^{-\delta^2/50\varepsilon}. \quad (3.1)$$

For our analysis we will require a lower bound for $\int_0^{2\pi} |F_N(e^{-\sigma + i t})|^2 \, dt$ that holds with high probability. For this purpose, we have the following

**Proposition 3.5.** Uniformly for all large $N$, all $0 \leq \sigma \leq \frac{1}{N^{1/100}}$ and all $1 \leq W \leq (\log N)^{1/3}$ (say), we have

$$
P\left[ \int_0^{2\pi} |F_N(e^{-\sigma + i t})|^2 \, dt \geq \frac{1}{e^{21W/10}} \min\{N, \frac{1}{\sigma}\} \sqrt{\log N} \right] \geq 1 - O(e^{-W/10}).$$

In the end, the proof of this proposition will boil down to a somewhat more precise version of [7, Proposition 11.1]. Unlike there, it does not seem like we can require the fairly simple constant barrier condition that is imposed, but instead need a slightly modified upper bound and a very weak lower bound of the same type as in [4, Multiplicative Chaos Result 4] in order to make the argument work.

Define $N_\sigma$ such that $\log N_\sigma$ is the largest integer $\leq \log \min\{N, \frac{1}{\sigma}\}$, and let $C$ be a (sufficiently large) constant such that $W + C$ is an integer. Let $M := \log N_\sigma - W - C$, and for $1 \leq m \leq M$ and $0 \leq t \leq 2\pi$, set

$$Z_t^{(W)}(m) = Z_t^{(W)}(m) := Z_{t,\sigma}(W + m) = \sum_{e^{W+C+m-1} < k \leq e^{W+C+m}} \frac{\mathfrak{R}X(k)e^{itk}}{\sqrt{k}e^{k\sigma}}$$

and $Z_t^{(W)}(m) := Z_0^{(W)}(m)$. We also define $\sigma_{m,W}^2 := \sigma_{m,W+C}^2$ as well as $\rho_{m,t}^{(W)} := \rho_{m+W+C,t}$ and $\rho_{m}^{(W)} := \rho_{m+W+C}$. Later on, we might alter the shift $W$ in the superscript and keep the analogous notation.

Further, let $\mathcal{L} \subseteq [0, 2\pi]$ denote the (random) set of $t$ satisfying the barrier condition

$$-Cj \leq \sum_{m=1}^{j} \left( Z_t^{(W)}(m) - 2\sigma_{m,W}^2 \right) \leq C - 3 \log(j) \quad (3.2)$$

for all $1 \leq j \leq M$. Also, set

$$F^{(\text{med})}(e^{-\sigma + i t}) := \exp\left( \sum_{e^{W+C} < k \leq N_\sigma} \frac{X(k)e^{itk}}{\sqrt{k}e^{k\sigma}} \right),$$
so that
\[ |F^{(med)}(e^{-\sigma+i\tau})| = \exp \left( \sum_{m=1}^{M} Z_t^{(W)}(m) \right). \]

The idea behind these definitions is that we consider separately three different sizes of \( k \) in \( F_N \).
Very small values of \( k \), namely those \( < e^{W+C} \), have the property that \( e^{ik} \) hardly varies when we change \( t \) by at most \( e^{-W} \), which will allow us to treat this range in a more direct matter. Very large \( k \) on the other hand, those with \( N\sigma < k \leq N \), contribute very little because the factor \( e^{-k\sigma} \) will make them rather small in a deterministic sense even after summing over \( k \). This leaves us with the task of treating intermediate values of \( k \) with \( e^{W+C} < k \leq N\sigma \).

Our basic strategy for proving Proposition 3.5 is to apply the following lower bounds
\[
\int_0^{2\pi} |F_N(e^{-\sigma+i\tau})|^2 d\tau \geq \left( \int_{\mathcal{L}} |F^{(med)}(e^{-\sigma+i\tau})|^2 d\tau \right)^2
\]
\[
\int_{\mathcal{L}} |F^{(med)}(e^{-\sigma+i\tau})|^2 \exp \left( -2 \sum_{k \leq e^{W+C}} \text{or} \sum_{N\sigma < k \leq N} \text{or} \sum_{\sqrt{k}e^{k\sigma}} \right) d\tau,
\]
where the second bound follows from Cauchy–Schwarz. Thus, we require an upper bound for the denominator here, as well as a lower bound for the numerator. The former will be a fairly direct matter to deal with, whereas the latter requires more careful analysis.

In order to obtain said lower bound for the numerator, the general strategy will be to apply Chebyshev’s inequality in order to show that with high probability, it does not deviate too much from its mean (say at most half of the mean). Thus, we need an upper bound for the variance
\[
E \left[ \left( \int_{\mathcal{L}} |F^{(med)}(e^{-\sigma+i\tau})|^2 d\tau - E \left( \int_{\mathcal{L}} |F^{(med)}(e^{-\sigma+i\tau})|^2 d\tau \right) \right)^2 \right]
\]
\[
= 2\pi \int_0^{2\pi} E \left[ 1(0 \in \mathcal{L}) |F^{(med)}(e^{-\sigma})|^2 1(t \in \mathcal{L}) |F^{(med)}(e^{-\sigma+i\tau})|^2 \right] d\tau - E \left[ \int_{\mathcal{L}} |F^{(med)}(e^{-\sigma+i\tau})|^2 d\tau \right]^2.
\]

Hence, the next step is to acquire good bounds for the correlation
\[
E \left[ 1(0 \in \mathcal{L}) |F^{(med)}(e^{-\sigma})|^2 1(t \in \mathcal{L}) |F^{(med)}(e^{-\sigma+i\tau})|^2 \right].
\]

We will obtain various bounds depending on the distance \( ||t|| \) of \( t \) to the nearest multiple of \( 2\pi \). We first show that when \( ||t|| \) is not too small, the two quantities are almost uncorrelated.

**Proposition 3.6.** Suppose that \( ||t|| \geq e^{-W} \). Then we have
\[
E \left[ 1(0 \in \mathcal{L}) |F^{(med)}(e^{-\sigma})|^2 1(t \in \mathcal{L}) |F^{(med)}(e^{-\sigma+i\tau})|^2 \right]
\]
\[
= \left( 1 + O \left( \frac{1}{||t||e^{W}} \right) \right) E \left[ 1(0 \in \mathcal{L}) |F^{(med)}(e^{-\sigma})|^2 \right] E \left[ 1(t \in \mathcal{L}) |F^{(med)}(e^{-\sigma+i\tau})|^2 \right].
\]
Proof. First, we note that by Proposition 2.2, on this range of $t$ we have $\varepsilon_{m,W}^2 \approx 1$ as well as

$$|\rho_{m,W}(t)| \leq \frac{1}{10^4 \|t\| e^{W+m}} \leq \frac{1}{10^4} \tag{3.7}$$

for all $1 \leq m \leq M$, by taking $C$ sufficiently large.

By independence of the $Z_t^{(W)}(m)$ for different $m$, plugging in the density of the bivariate Gaussian we have

$$\mathbb{E} \left[ 1(0 \in \mathcal{L}) |F^{(med)}(e^{-\sigma})| \mathbb{1}(t \in \mathcal{L}) |F^{(med)}(e^{-\sigma+it})|^2 \right] = \left( \prod_{m=1}^{M} \frac{1}{2\pi \sigma_{m,W}^2 \sqrt{1-\rho_{m}^{(W)^2}}} \right)$$

$$\times \int \cdots \int_{\mathcal{R}} \exp \left( \sum_{m=1}^{M} (2x_m + 2y_m) - \sum_{m=1}^{M} \frac{x_m^2 - 2\rho_{m,W} x_m y_m + y_m^2}{2(1-\rho_{m}^{(W)^2})\sigma_{m,W}^2} \right) dx \, dy,$$

where

$$\mathcal{R} := \left\{ (x, y) \in \mathbb{R}^M \times \mathbb{R}^M : -C \leq \sum_{m=1}^{j} x_m - 2\sigma_{m,W}^2, \sum_{m=1}^{j} y_m - 2\sigma_{m,W}^2 \leq C - 3 \log(j) \forall 1 \leq j \leq M \right\}.$$

Substituting $x'_m := x_m - 2\sigma_{m,W}^2$ and $y'_m := y_m - 2\sigma_{m,W}^2$, this equates to

$$\exp \left( 4 \sum_{m=1}^{M} \sigma_{m,W}^2 + 4 \sum_{m=1}^{M} \rho_{m,W}^{(W)} \sigma_{m,W}^2 \right) \left( \prod_{m=1}^{M} \frac{1}{2\pi \sigma_{m,W}^2 \sqrt{1-\rho_{m}^{(W)^2}}} \right)$$

$$\times \int \cdots \int_{\mathcal{R}'} \exp \left( - \sum_{m=1}^{M} \frac{x_m' - 2\rho_{m,W} x'_m y'_m + y'_m^2}{2(1-\rho_{m}^{(W)^2})\sigma_{m,W}^2} + \sum_{m=1}^{M} \frac{2\rho_{m,W}(x'_m + y'_m)}{1 + \rho_{m}^{(W)^2}} \right) dx' \, dy' \tag{3.8}$$

with

$$\mathcal{R}' := \left\{ (x', y') \in \mathbb{R}^M \times \mathbb{R}^M : -C \leq \sum_{m=1}^{j} x'_m, \sum_{m=1}^{j} y'_m \leq C - 3 \log(j) \forall 1 \leq j \leq M \right\}.$$

At this point we make crucial use of the weak lower bound that we implemented in the barrier condition. It implies in particular that $-C \leq x'_1 \leq C$ and that for any $2 \leq j \leq M$ we have

$$x'_j = \sum_{m=1}^{j} x'_m - \sum_{m=1}^{j-1} x'_m \leq C + C(j-1) \ll j$$
and similarly

\[ x'_j \geq -Cj - C, \]

hence \(|x'_j|, |y'_j| \ll j\) for all \(1 \leq j \leq M\). Note further that

\[
\prod_{m=1}^{M} \frac{1}{\sqrt{1 - \rho_m^W}} = \prod_{m=1}^{M} \left(1 + O(\rho_m^W)^2\right) = 1 + O\left(\sum_{m=1}^{M} \rho_m^W\right)
\]

as well as

\[
\exp\left(-\sum_{m=1}^{M} \frac{x'^2_m}{2(1 - \rho_m^W)^2}\sigma^2_{m,W}\right) = \exp\left(-\sum_{m=1}^{M} \frac{x'^2_m}{2\sigma^2_{m,W}} - \sum_{m=1}^{M} \frac{\rho_m^W x'^2_m}{2\sigma^2_{m,W}(1 - \rho_m^W)^2}\right)
\]

\[
= \exp\left(-\sum_{m=1}^{M} \frac{x'^2_m}{2\sigma^2_{m,W}}\right) \exp\left(O\left(\sum_{m=1}^{M} \rho_m^W\right)\right).
\]

These estimates combined with (3.7) imply that the expression in (3.8) is

\[
= \exp\left(O\left(\sum_{m=1}^{M} m^2\rho_m^W\right)\right) \exp\left(4\sum_{m=1}^{M} \frac{\sigma^2_{m,W}}{2\sigma^2_{m,W}}\right) \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi\sigma^2_{m,W}}} \int_{\mathcal{R}'_1} \cdots \int_{\mathcal{R}'_I} \exp\left(-\sum_{m=1}^{M} \frac{x'^2_m + y'^2_m}{2\sigma^2_{m,W}}\right) \, dx' \, dy'
\]

\[
= \left(1 + O\left(\frac{1}{\|\ell\|_W}\right)\right) \exp\left(2\sum_{m=1}^{M} \sigma^2_{m,W}\right) \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi\sigma^2_{m,W}}} \int_{\mathcal{R}'_1} \cdots \int_{\mathcal{R}'_I} \exp\left(-\sum_{m=1}^{M} \frac{x'^2_m}{2\sigma^2_{m,W}}\right) \, dx'
\]

\[
\times \exp\left(2\sum_{m=1}^{M} \sigma^2_{m,W}\right) \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi\sigma^2_{m,W}}} \int_{\mathcal{R}'_1} \cdots \int_{\mathcal{R}'_I} \exp\left(-\sum_{m=1}^{M} \frac{y'^2_m}{2\sigma^2_{m,W}}\right) \, dy',
\]

(3.9)

where \(\mathcal{R}'_1\) and \(\mathcal{R}'_2\) describe the individual conditions for \(x'_m\) respectively \(y'_m\) that appear in \(\mathcal{R}'\).

On the other hand, we have

\[
\mathbb{E}\left[1(0 \in \mathcal{L}) |\rho^{(\text{med})}(e^{-\sigma})|^2\right] = \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi\sigma^2_{m,W}}} \int_{\mathcal{R}} \cdots \int_{\mathcal{R}} \exp\left(\sum_{m=1}^{M} \left(2\sigma_{m} - \frac{z^2_{m}}{2\sigma^2_{m,W}}\right)\right) \, dz,
\]

where

\[
\mathcal{R} := \left\{ (z_m)_{m} : -C \leq \sum_{m=1}^{j} \left(z_m - 2\sigma^2_{m,W}\right) \leq C - 3 \log(j) \forall 1 \leq j \leq M \right\}
\]
(and the same at $t$ by translation invariance in law). Again substituting $z'_m = z_m - 2\sigma^2_{m,W}$, this equates to

$$\exp \left( 2 \sum_{m=1}^{M} \sigma^2_{m,W} \right) \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi \sigma^2_{m,W}}} \int_{R'} \cdots \int_{R'} \exp \left( - \sum_{m=1}^{M} \frac{z'^2_m}{2\sigma^2_{m,W}} \right) dz', \quad (3.10)$$

where

$$R' := \left\{ (z'_m)_m : -Cj \leq \sum_{m=1}^{j} z'_m \leq C - 3\log(j) \forall 1 \leq j \leq M \right\}.$$

Comparing (3.9) and (3.10), the claim follows.

We would like to record here for future use that by Proposition 2.2 and (3.10) we have

$$\mathbb{E} \left[ 1(0 \in \mathcal{L}) | F^{(med)}(e^{-\sigma}) |^2 \right] \lesssim \exp \left( \sum_{m=1}^{M} (1 + O(e^{-W-m})) \right) \mathbb{P}[B] \leq \frac{N\sigma}{e^W \mathbb{P}[B]}, \quad (3.11)$$

where $B$ denotes the event that a sequence of Gaussians $G_1, \ldots, G_M$ with means 0 and variances $\sigma^2_{m,W}$ satisfies the barrier condition defined by $R'$, that is, that for all $1 \leq j \leq M$ we have

$$-Cj \leq \sum_{m=1}^{j} G_m \leq C - 3\log j.$$

Thus, Proposition 3.2 implies that

$$\mathbb{E} \left[ 1(0 \in \mathcal{L}) | F^{(med)}(e^{-\sigma}) |^2 \right] \asymp \frac{N\sigma}{e^W \sqrt{\log N\sigma}}. \quad (3.12)$$

Next, we will deal with the range $0 < \|t\| \leq e^{-W}$.

**Proposition 3.7.** For $0 < \|t\| \leq N^{-1/3}_\sigma$, we have

$$\mathbb{E} \left[ 1(0 \in \mathcal{L}) | F^{(med)}(e^{-\sigma}) |^2 1(t \in \mathcal{L}) | F^{(med)}(e^{-\sigma+i\|t\|}) |^2 \right] \ll \frac{N^2 \sigma}{e^{3W} (\log N\sigma)^6} \min \left\{ N\sigma, \frac{1}{\|t\|} \right\}. \quad (3.13)$$

For $N^{-1/3}_\sigma \leq \|t\| \leq e^{-W}$, we have

$$\mathbb{E} \left[ 1(0 \in \mathcal{L}) | F^{(med)}(e^{-\sigma}) |^2 1(t \in \mathcal{L}) | F^{(med)}(e^{-\sigma+i\|t\|}) |^2 \right] \ll \frac{N^2 \sigma}{e^{3W} \|t\| \log N\sigma (\log 1/\|t\| - W + 1)^2}.$$

$$\quad (3.14)$$
Proof. We begin with the proof of (3.13), and make use of the parameter

\[ M_1 := \min \{ M, \lfloor \log(1/\|t\|) \rfloor - [W] + 1 \}, \]

which satisfies \( 1 \leq M_1 \leq M \). The purpose of \( M_1 \) is to distinguish the values of \( m \) where \( Z_m^{(W)}(0) \) and \( Z_m^{(W)}(t) \) are strongly correlated from those where they are almost uncorrelated. Our barrier condition (3.2) implies that for \( t \in \mathcal{E} \) we have

\[
\exp \left( 2 \sum_{m=1}^{M_1} Z_m^{(W)}(t) \right) \ll \exp \left( 4 \sum_{m=1}^{M_1} \sigma_{m,W}^2 \right) \frac{1}{M_1^6} \ll \frac{e^{2M_1}}{M_1^6}. \tag{3.15}
\]

Note that since \( \|t\| \leq N_\sigma^{-1/3} \) we have \( M_1 \asymp \log N_\sigma \). By discarding the indicator functions after applying (3.15), we therefore obtain that

\[
\mathbb{E} \left[ 1(0 \in \mathcal{E}) \left| F^{(med)}(e^{-\sigma}) \right|^2 1(t \in \mathcal{E}) \left| F^{(med)}(e^{-\sigma+it}) \right|^2 \right] \\
\ll \frac{e^{2M_1}}{M_1^6} \mathbb{E} \left[ \exp \left( 2 \sum_{m=1}^{M} Z_m^{(W)}(m) + 2 \sum_{m=M_1+1}^{M} Z_t^{(W)}(m) \right) \right] \\
= \frac{e^{2M_1}}{M_1^6} \mathbb{E} \left[ \exp \left( 2 \sum_{m=1}^{M_1} Z_m^{(W)}(m) \right) \right] \mathbb{E} \left[ \exp \left( 2 \sum_{m=M_1+1}^{M} Z_m^{(W)}(m) + 2 \sum_{m=M_1+1}^{M_1} Z_t^{(W)}(m) \right) \right].
\]

Using Proposition 2.2 and 2.3, this is simply

\[
\ll \frac{e^{2M_1}}{M_1^6} e^{2(M-M_1)} = \frac{e^{M_1+2M}}{M_1^6} \ll \frac{1}{(\log N_\sigma)^6} \frac{1}{e^W} \frac{N_\sigma^2}{e^{2W}} \ll \min \left\{ N_\sigma, \frac{1}{\|t\|} \right\}.
\]

We would like to point out here that we have made crucial use of the small shift \(-3 \log j \) in our barrier condition, which has produced this additional factor \((\log N_\sigma)^6 \) in the denominator. Without this, our bound on the contribution from small \( \|t\| \) would turn out to be insufficient. This will be even more apparent on the next range.

Now we shall deal with the range \( N_\sigma^{-1/3} \leq \|t\| \leq e^{-W} \). This time, we will modify \( M_1 \) slightly by setting

\[ M_2 := M_1 + 2 \lfloor \log M_1 \rfloor = \lfloor \log(1/\|t\|) \rfloor - [W] + 1 + 2 \lfloor \log(\lfloor \log(1/\|t\|) \rfloor - [W] + 1) \rfloor. \]

This now allows us to distinguish between weakly and strongly correlated \( Z_t^{(W)} \) variables with some margin.

Here we have \( 1 \leq M_2 \leq \frac{1}{2} \log N_\sigma \) (say) and hence \( M - M_2 \asymp \log N_\sigma \). We need to exploit more information from our barrier condition (3.2). Note that it implies in particular that for all \( M_2 + 1 \leq j \leq M \) we have

\[
-C(j + 1) + 3 \log M_2 \leq \sum_{m=M_2+1}^{j} (Z_m^{(W)}(m) - \sigma_{m,W}^2) \leq C(M_2 + 1) - 3 \log j \tag{3.16}
\]
and the same event for $Z_t^{(W)}(m)$ in place of $Z^{(W)}(m)$. Denoting this event (at both 0 and $t$) by $\mathcal{L}'$, we obtain

\[
\mathbb{E}\left[1(0 \in \mathcal{L})|F^{(\text{med})}(e^{-\sigma})|^21(t \in \mathcal{L})|F^{(\text{med})}(e^{-\sigma+i\tau})|^2 \right] \\
\ll \frac{e^3M_2}{M_6^2} \mathbb{E}\left[1(\mathcal{L}'_t) \exp \left(2 \sum_{m=M_2+1}^M Z^{(W)}(m) + 2 \sum_{m=M_2+1}^M Z_t^{(W)}(m) \right) \right].
\]

But in order to estimate this, we can employ a similar strategy as in the proof of Proposition 3.6. Note that for any $1 \leq m' \leq M - M_2$, Proposition 2.2 and the assumption that $C$ be sufficiently large imply that we have

\[
|\rho^{(W+M_2)}_{m'}| \leq \frac{1}{10^4||t||e^{m'+W+M_2}} \leq \frac{1}{10^3},
\]

say. Plugging in the bivariate Gaussian density to obtain a $2(M - M_2)$-fold integral and making the same substitution, we arrive at

\[
\exp \left(4 \sum_{m'=1}^{M-M_2} \sigma_{m',W+M_2}^2 + 4 \sum_{m'=1}^{M-M_2} \frac{T_{m'}^{(W+M_2)} \sigma_{m',W+M_2}^2}{1 + \rho_{m'}^{(W+M_2)}} \prod_{m'=1}^{M-M_2} \frac{1}{2\pi \sigma_{m',W+M_2}^2 \sqrt{1 - \rho_{m'}^{(W+M_2)}^2}} \right) \\
\times \int \cdots \int_{\mathcal{R}'} \exp \left(- \sum_{m'=1}^{M-M_2} \frac{x_{m'}^2 - 2\rho_{m'}^{(W+M_2)} x_{m'} y_{m'} + y_{m'}^2}{2(1 - \rho_{m'}^{(W+M_2)}^2) \sigma_{m',W+M_2}^2} + \sum_{m'=1}^{M-M_2} \frac{2\rho_{m'}^{(W+M_2)} (x_{m'}^2 + y_{m'}^2)}{1 + \rho_{m'}^{(W+M_2)}} \right) dx' dy'
\]

with

\[
\mathcal{R}' := \left\{ (x',y') \in \mathbb{R}^{M-M_2} \times \mathbb{R}^{M-M_2} : -C(j + M_2 + 1) + 3\log M_2 \leq \sum_{m'=1}^j x_{m'}', \sum_{m'=1}^j y_{m'}' \leq C(M_2 + 1) - 3\log(j + M_2) \forall 1 \leq j \leq M - M_2 \right\}.
\]

By considering successive differences similar as in the proof of Proposition 3.6, we therefore have $|x_{m'}', |y_{m'}'| \ll m' + M_2$. Using this and then invoking (3.17), we can see that the previous expression is

\[
\ll \exp \left(O\left(\frac{M_2^2}{||t||e^{W+M_2}}\right)\right) e^{2(M-M_2)\mathbb{P}[B]^2} \ll e^{2(M-M_2)\mathbb{P}[B]^2}. \quad (3.18)
\]
Here, $B$ is the event that a sequence of Gaussians $G_1, \ldots, G_{M-M_2}$ with mean 0 and variance $\sigma_{1, W+M_2}^2, \ldots, \sigma_{M-M_2, W+M_2}^2$ satisfies that for all $1 \leq j \leq M - M_2$ we have

$$\sum_{m'=1}^{j} G_{m'} \leq C(M_2 + 1) - 3 \log(j + M_2).$$

Note that in the final step of (3.18) we crucially make use of the definition of $M_2$ (with our old parameter $M_1$, this error would not be acceptable). In order to understand the event $B$, we can employ Proposition 3.1 with $a = C(M_2 + 1) - 3 \log(M_2 + 1)$ and $n = M - M_2$ to deduce that

$$\mathbb{P}[B] \ll \frac{M_2}{\sqrt{M-M_2}}.$$ We therefore conclude that

$$\mathbb{E} \left[ 1(0 \in \mathcal{L}) |P^{(\text{med})}(e^{-\sigma})|^2 1(t \in \mathcal{L})|P^{(\text{med})}(e^{-\sigma+it})|^2 \right] \ll e^{3M_2} M_2^2 e^{2(M-M_2)} M_2^2 M - M_2 = e^{2M+M_2} (M - M_2) M_2^2 \ll e^{3W \|t\| \log N_\sigma (\log 1/\|t\|-W+1)^2}$$

as claimed. □

We are now in a position to complete the proof of Proposition 3.5.

**Conclusion of the proof of Proposition 3.5.** Most of the work has already been done and we essentially just need to integrate our respective bounds over their ranges of $t$.

We begin by obtaining an upper bound for the denominator in (3.3) that holds with high probability. To this end, we note that by Proposition 2.3 and (3.12) its expectation satisfies

$$\mathbb{E} \left[ \int_{\mathcal{L}} |P^{(\text{med})}(e^{-\sigma+it})|^2 \exp \left( -2 \sum_{k \leq e^{W+C} \text{ or } \mathcal{N}_\sigma < k \leq N} \frac{\Re X(k) e^{ik}}{\sqrt{k e^{k \sigma}}} \right) dt \right]$$

$$= \int_0^{2\pi} \mathbb{E} \left[ 1(t \in \mathcal{L}) |P^{(\text{med})}(e^{-\sigma+it})|^2 \right] \exp \left( -2 \sum_{k \leq e^{W+C} \text{ or } \mathcal{N}_\sigma < k \leq N} \frac{\Re X(k) e^{ik}}{\sqrt{k e^{k \sigma}}} \right) dt$$

$$\ll e^{W} \mathbb{E} \left[ 1(0 \in \mathcal{L}) |P^{(\text{med})}(e^{-\sigma})|^2 \right] \ll \frac{N_\sigma}{\sqrt{\log N_\sigma}}.$$

In the second step we have used that

$$\sum_{k > N_\sigma} \frac{1}{k e^{2k \sigma}} \ll 1$$

to bound the contribution from the large $k$, which follows, for example, by dividing into dyadic intervals and bounding trivially on each of them.
Hence, by Markov’s inequality, with probability \( \geq 1 - O(e^{-W/10}) \) we have

\[
\int_{\mathcal{L}} |F^{(\text{med})}(e^{-\sigma+it})|^2 \exp \left( -2 \sum_{k \leq eW + C \text{ or } N_{\sigma} < k \leq N} \frac{R_X(k)e^{ikt}}{\sqrt{k \sigma}} \right) dt \leq \frac{e^{W/10}N_{\sigma}}{\sqrt{\log N_{\sigma}}}. \tag{3.19}
\]

As for the numerator in (3.3), we note first that the contribution from small \( ||t|| \) to the first integral in (3.4) is small: By Propositions 3.6 and 3.7 as well as by (3.12) we have

\[
\int_{||t|| \leq e^{-W/2}} \mathbb{E} \left[ |(0 \in \mathcal{L})F^{(\text{med})}(e^{-\sigma})|^2 (t \in \mathcal{L}) F^{(\text{med})}(e^{-\sigma+it})|^2 \right] dt \ll \frac{N_{\sigma}^2}{e^{3W}(\log N_{\sigma})^6} \int_{0}^{N_{\sigma}^{-1/3}} \min \left\{ N_{\sigma}, \frac{1}{t} \right\} dt + \frac{N_{\sigma}^2}{e^{3W} \log N_{\sigma}} \int_{N_{\sigma}^{-1/3}}^{e^{-W}} \frac{1}{t^2 (\log 1/t - W + 1)^2} dt + \frac{N_{\sigma}^2}{e^{5W/2} \log N_{\sigma}} \ll \frac{N_{\sigma}^2}{e^{5W/2} \log N_{\sigma}}.
\]

As a consequence we deduce, again invoking Proposition 3.6 and then applying (3.12), that the variance of the numerator in (3.3) (compare also (3.4)) satisfies

\[
\mathbb{E} \left[ \left( \int_{\mathcal{L}} |F^{(\text{med})}(e^{-\sigma+it})|^2 dt - \mathbb{E} \int_{\mathcal{L}} |F^{(\text{med})}(e^{-\sigma+it})|^2 dt \right)^2 \right] \ll e^{-W/2} \mathbb{E} \left[ |(0 \in \mathcal{L})|F^{(\text{med})}(e^{-\sigma})|^2 \right] + \frac{N_{\sigma}^2}{e^{5W/2} \log N_{\sigma}} \ll e^{-W/2} \mathbb{E} \left[ |(0 \in \mathcal{L})|F^{(\text{med})}(e^{-\sigma})|^2 \right]^2.
\]

Hence, by Chebyshev’s inequality, we have that with probability \( \geq 1 - O(e^{-W/2}) \) the numerator in (3.3) satisfies

\[
\int_{\mathcal{L}} |F^{(\text{med})}(e^{-\sigma+it})|^2 dt \geq \frac{1}{2} \mathbb{E} \left[ \int_{\mathcal{L}} |F^{(\text{med})}(e^{-\sigma+it})|^2 dt \right] \gg \frac{N_{\sigma}}{e^{W} \sqrt{\log N_{\sigma}}}.
\]

Combining this with (3.3) and (3.19) (and the fact that \( \log N_{\sigma} \asymp \log N \)), we conclude that with probability \( \geq 1 - O(e^{-W/10}) \) we have

\[
\int_{0}^{2\pi} |F_N(e^{-\sigma+it})|^2 dt \geq \frac{N_{\sigma}}{e^{21W/10} \sqrt{\log N}},
\]

which gives the claim. \( \square \)

## 4 | REDUCTION TO A COVARIANCE ESTIMATE

We continue to proceed in a rather similar fashion as in [4] and [7]. The first step in the reduction is to discard the \( N \)-smooth part. For \( N < n < 2N \), we write

\[
A(n) = \sum_{\lambda \vdash n} a(\lambda) = \sum_{N < k \leq n} \frac{X(k)}{\sqrt{k}} \sum_{\sigma \vdash n-k} a(\sigma) + \sum_{\lambda_1 \vdash n} a(\lambda_1).
\]
Here we used the fact that for $n$ in this range there can be at most one partition of length $> N$, as well as that $a((k)) = \frac{X(k)}{\sqrt{k}}$, where $(k)$ denotes the partition of $k$ with only one part. As for the $N$-smooth part, note that

$$\mathbb{E}
\left[\left\{ \frac{8N}{7} \leq n \leq \frac{4N}{3} : \sum_{\lambda \vdash n, \lambda_1 \leq N} a(\lambda) \geq (\log N)^{1/100} \right\} \right]$$

$$\leq \sum_{\frac{8N}{7} \leq n \leq \frac{4N}{3}} \mathbb{E}
\left[ \left| \sum_{\lambda \vdash n, \lambda_1 \leq N} a(\lambda) \right|^2 \right] \leq \frac{N}{(\log N)^{1/50}}$$

using orthogonality and the fact that $\mathbb{E}[|A(n)|^2] = 1$ for all $n$. Thus, with probability $\geq 1 - O((\log N)^{-1/50})$ there exists a random set $\mathcal{N} \subseteq \left[ \frac{8N}{7}, \frac{4N}{3} \right]$ of cardinality $\geq \frac{N}{7}$ (say) such that

$$\left| \sum_{\lambda \vdash n, \lambda_1 \leq N} a(\lambda) \right| \leq (\log N)^{1/100} \quad (4.2)$$

for all $n \in \mathcal{N}$. So far, our argument would have worked just as well on our original sum, but the crucial point is that the set $\mathcal{N}$ only depends on the random variables $(X(k))_{k \leq N}$. Since we will now condition on these random variables, the set $\mathcal{N}$ will be fixed under this conditioning.

For $n \in \mathcal{N}$, let $Y_n$ be the random variables defined by

$$\Re \sum_{N < k \leq n} \frac{X(k)}{\sqrt{k}} \sum_{\sigma \vdash n-k} a(\sigma)$$

conditional on the values of $(X(k))_{k \leq N}$, so that $Y_n$ follows a Gaussian distribution, now being a linear combination of independent Gaussians. We will show that this real part is large enough with high probability, and from that deduce the claim for the absolute value.

We have

$$\mathbb{E}[Y_n] = 0$$

and

$$\mathbb{E}[Y_n^2] = \frac{1}{4} \mathbb{E}
\left[ \left( \sum_{N < k \leq n} \frac{X(k)}{\sqrt{k}} \sum_{\sigma \vdash n-k} a(\sigma) + \sum_{N < k \leq n} \frac{X(k)}{\sqrt{k}} \sum_{\sigma \vdash n-k} a(\sigma) \right)^2 \mid (X(k))_{k \leq N} \right]$$

$$= \sum_{N < k \leq n} \frac{1}{2k} \left| \sum_{\sigma \vdash n-k} a(\sigma) \right|^2.$$

In a similar manner, we obtain

$$\mathbb{E}[Y_n Y_m] = \Re \sum_{N < k \leq \frac{4N}{3}} \frac{1}{2k} \left( \sum_{\sigma \vdash n-k} a(\sigma) \right) \left( \sum_{\sigma \vdash m-k} \overline{a(\sigma)} \right),$$

where we interpret the respective sums as empty whenever $n - k$ respectively $m - k$ is negative.
The next step is to obtain lower bounds for the variance that hold with high probability.

**Proposition 4.1.** Let $\frac{8N}{7} \leq n \leq \frac{4N}{3}$ and $1 \leq W \leq (\log N)^{1/3}$. Then with probability $1 - O(e^{-W/10})$, we have

$$\sum_{N < k \leq n} \frac{1}{k} \left| \sum_{\sigma \vdash n-k} a(\sigma) \right|^2 \geq \frac{1}{e^{11W/5} \sqrt{\log N}}.$$  

**Proof.** We follow ideas from [4, Proposition 1] and [7, Proposition 8.1]. First, note that on the prescribed range of $k$ we have $\frac{1}{k} \geq \frac{3}{4N}$, and after replacing $k$ by $n - k$ we get

$$\sum_{N < k \leq n} \frac{1}{k} \left| \sum_{\sigma \vdash n-k} a(\sigma) \right|^2 \geq \frac{3}{4N} \sum_{k \leq n-N} \left| \sum_{\sigma \vdash k} a(\sigma) \right|^2 \gg \frac{1}{N} \sum_{k \leq n} \left| \sum_{\sigma \vdash k} a(\sigma) \right|^2,$$

where we inserted the void condition $\sigma_1 \leq N$ for use in the following step. Now for any $0 \leq r \leq 1$, we get that this expression is

$$\geq \frac{1}{N} \sum_{k \geq 0} r^{7k} \left| \sum_{\sigma \vdash k} a(\sigma) \right|^2 - \frac{r^N}{N} \sum_{k \geq 0} \left| \sum_{\sigma \vdash k} a(\sigma) \right|^2.$$

Next, noting that

$$F_N(z) = \sum_{k \geq 0} \left( \sum_{\sigma \vdash k} a(\sigma) \right) z^k,$$

Parseval's identity implies that this equates to

$$\frac{1}{2\pi N} \int_0^{2\pi} \left| F_N(e^{-7V/2N+it}) \right|^2 dt - \frac{e^{-V}}{2\pi N} \int_0^{2\pi} \left| F_N(e^{it}) \right|^2 dt$$

with $r = e^{-V/N}$ and $V \geq 0$. Applying [7, Proposition 3.2] with $q = \frac{2}{3}$ (and $r = 1$), we obtain that the subtracted second term satisfies

$$\mathbb{E} \left[ \left( \frac{1}{2\pi N} \int_0^{2\pi} \left| F_N(e^{it}) \right|^2 dt \right)^{2/3} \right] \ll \left( \frac{1}{\sqrt{\log N}} \right)^{2/3}.$$

Hence Markov inequality tells us (for any $A \geq 1$) that with probability $\geq 1 - O(A^{-2/3})$ we have

$$\int_0^{2\pi} \left| F_N(e^{it}) \right|^2 dt \leq \frac{A}{\sqrt{\log N}}.$$ 

As for the first term, Proposition 3.5 tells us that with probability $\geq 1 - O(e^{-W/10})$,

$$\int_0^{2\pi} \left| F_N(e^{-7V/2N+it}) \right|^2 dt \geq \min\{1, \frac{2}{7V} \} \frac{e^{21W/10} \sqrt{\log N}}{}.$$
Choosing (say) $V = 6W$ and $A = e^{3W}$ and assuming without loss of generality that $W$ is large enough, the claim follows upon combining these estimates.

The covariance estimate which we shall rely on for now and prove in the next section is the following

**Proposition 4.2.** Let $N$ be sufficiently large, and for $n \in \left[ \frac{8N}{7}, \frac{4N}{3} \right]$ define

$$B_n := \left\{ m \in \left[ \frac{8N}{7}, \frac{4N}{3} \right] : \left| \sum_{N<k\leq \frac{4N}{3}} \frac{1}{k} \left( \sum_{\sigma|n-k} a(\sigma) \right) \left( \sum_{\sigma|m-k} \overline{a(\sigma)} \right) \right| \geq \frac{1}{(\log N)^{4/5}} \right\}.$$

Then with probability $\geq 1 - O((\log N)^{-1/10})$, we have $\max_n |B_n| \leq N^{7/10}$.

Assuming this proposition, we are now in a position to prove Theorem 2. Again, this proceeds similar to the proof of Theorem 2 in [4], though with minor simplifications.

**Proof of Theorem 2 assuming Proposition 4.2.** Suppose without loss of generality that $W$ is sufficiently large. Note first that by (4.1) and (4.2) we have

$$\mathbb{P} \left[ \max_{8N/7 \leq n \leq 4N/3} |A(n)| \leq \frac{(\log N)^{1/4}}{e^{6W/5}} \right] \leq \mathbb{P} \left[ \max_{8N/7 \leq n \leq 4N/3} \Re A(n) \leq \frac{2(\log N)^{1/4}}{e^{6W/5}} - (\log N)^{1/100} \right]$$

$$\leq \mathbb{P} \left[ \max_{n \in \mathbb{N}} \Re \sum_{N<k\leq n} \frac{X(k)}{\sqrt{k}} \sum_{\sigma|n-k} a(\sigma) \leq \frac{2(\log N)^{1/4}}{e^{6W/5}} \right] + O((\log N)^{-1/50})$$

$$= \mathbb{E} \left[ \mathbb{P} \left[ \max_{n \in \mathbb{N}} Y_n \leq \frac{2(\log N)^{1/4}}{e^{6W/5}} \right] \right] + O((\log N)^{-1/50}). \quad (4.3)$$

Defining $V := \min_{n \in \mathbb{N}'} \mathbb{E}[Y_n^2]$ (which is really a conditional variance), Proposition 4.1 implies that with probability $\geq 1 - O(e^{-W/10})$, the event

$$E := \left\{ V \geq \frac{1}{e^{6W/5} \sqrt{\log N}} \right\}$$

holds. Denoting $\mathbb{P}_E[A] := \mathbb{P}[A \mid E]$ (and analogously $\mathbb{E}_E[A]$), we thus have $\mathbb{P}[A] \leq \mathbb{P}_E[A] + O(e^{-W/10})$ for any event $A$ and it suffices in (4.3) to consider this conditional probability. We have

$$\mathbb{P}_E \left[ \max_{n \in \mathbb{N}'} Y_n \leq \frac{2(\log N)^{1/4}}{e^{6W/5}} \right] \leq \mathbb{P}_E \left[ \max_{n \in \mathbb{N}'} \frac{Y_n}{\sqrt{\mathbb{E}[Y_n^2]}} \leq \frac{2(\log N)^{1/4}}{\sqrt{V} e^{6W/5}} \right] \leq \mathbb{P}_E \left[ \max_{n \in \mathbb{N}'} \frac{Y_n}{\sqrt{\mathbb{E}[Y_n^2]}} \leq \sqrt{\frac{1}{2} \log N} \right],$$

say, since $W$ is sufficiently large. Note that the random variables $Y_n / \sqrt{\mathbb{E}[Y_n^2]}$ are jointly Gaussian and each of them is normalised to mean 0 and variance 1.
Next, Proposition 4.2 implies that, with probability \( \geq 1 - O((\log N)^{-1/10}) \) the following holds: there exists a subset \( \mathcal{N}' \) of \( \mathcal{N} \) of cardinality \( \gg N^{3/10} \), such that for any \( n \neq m \in \mathcal{N}' \) we have the correlation estimate

\[
E_E \left[ \frac{Y_n Y_m}{\sqrt{E[Y_n^2]} \sqrt{E[Y_m^2]}} \right] \leq \frac{(\log N)^{-4/5}}{2V} \ll \frac{e^{1.1W/5}}{N^{3/10}}.
\]

By our assumption on \( W \), we can thus make this correlation small, say smaller than \( \epsilon \) for some sufficiently small constant \( 0 < \epsilon < 10^{-4} \); this puts us into a position to apply Proposition 3.4 with (say) \( \delta = \sqrt{\epsilon} \), which implies that

\[
P_E \left[ \max_{n \in \mathcal{N}'} \frac{Y_n}{\sqrt{E[Y_n^2]}} \leq \sqrt{\frac{1}{2} \log N} \right] \leq P_E \left[ \max_{n \in \mathcal{N}'} \frac{Y_n}{\sqrt{E[Y_n^2]}} \leq \sqrt{(2 - \sqrt{\epsilon}) \log(\#\mathcal{N}')} \right] \ll \exp \left( -N^{\sqrt{\epsilon}/100} \right) + N^{-3/500} \ll N^{-3/500}.
\]

Putting all of this together implies that we have

\[
\max_{\frac{8N}{7} \leq n \leq \frac{4N}{3}} |A(n)| \geq \frac{(\log N)^{1/4}}{e^{1.2W}}
\]

with probability

\[
\geq 1 - O((\log N)^{-1/50}) - O(e^{-W/10}) - O((\log N)^{-1/10}) - O(N^{-3/500}) = 1 - O(e^{-W/10}). \quad \square
\]

5 | PROOF OF THE COVARIANCE ESTIMATE

We begin by noting that as a consequence of (2.1) (which holds even for negative \( N \) if we set \( A(N) \) to be 0), we have that the covariance sum we are interested in satisfies, for any \( K \geq \frac{N}{3} \) and \( m, n \in \left[ \frac{8N}{7}, \frac{4N}{3} \right] \), the relation

\[
\sum_{N < k \leq \frac{4N}{3}} \frac{1}{k} \left( \sum_{\sigma \rightarrow n-k} a(\sigma) \right) \left( \sum_{\sigma \rightarrow m-k} \overline{a(\sigma)} \right) = \sum_{N < k \leq \frac{4N}{3}} \frac{1}{(2\pi i)^2 k} \int_{|z_1|=1} \int_{|z_2|=1} F_K(z_1) F_K(z_2) \frac{z_1^{m-k+1}}{z_2^{n-k+1}} \, dz_1 \, dz_2

= \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} F_K(e^{it_1}) F_K(e^{it_2}) e^{i(t_1(m+1) - t_2(n+1))} \sum_{N < k \leq \frac{4N}{3}} \frac{e^{i(t_2-t_1)k}}{k} \, dt_1 \, dt_2.
\]

We shall take \( \frac{N}{3} \leq K \leq N \) (say) such that \( \log K \) is an integer. Our first task is to restrict, with high probability, to an integration range where \( \|t_2 - t_1\| \) is fairly small.
Proposition 5.1. Let $\frac{N}{3} \leq K \leq N$ be such that $\log K$ is an integer, and let $\frac{8N}{7} \leq n, m \leq \frac{4N}{3}$. With probability $\geq 1 - O((\log N)^{-25})$ we have

$$
\sum_{N < k \leq \frac{4N}{3}} \frac{1}{k} \left( \sum_{\sigma + n = k} a(\sigma) \right) \left( \sum_{\sigma + m = k} \overline{a(\sigma)} \right) = \frac{1}{4\pi^2} \int_{[0,2\pi] \times [0,2\pi]} \left| F_K(e^{it_1})F_K(e^{it_2})e^{i(t_1(m+1)-t_2(n+1))} \right| \sum_{N < k \leq \frac{4N}{3}} \frac{e^{i(t_2-t_1)k}}{k} \, dt_1 \, dt_2 + O((\log N)^{-25}).
$$

(5.1)

Proof. We proceed by bounding the expectation on the range where $\|t_2 - t_1\| > \frac{\log N^{100}}{N}$ and then applying Markov’s inequality. Note that by Proposition 2.1 we have

$$
\sum_{N < k \leq \frac{4N}{3}} e^{i(t_2-t_1)k} \ll \frac{1}{\|t_2 - t_1\|}.N.
$$

Thus we can bound the expectation of the integral on this complementary range by

$$
\ll \frac{1}{N} \int_{[0,2\pi] \times [0,2\pi]} \frac{\mathbb{E} \left| F_K(e^{it_1})F_K(e^{it_2}) \right|^2}{\|t_2 - t_1\|} \, dt_1 \, dt_2. \quad (5.2)
$$

Then by Propositions 2.3 and 2.1 the expectation in the numerator satisfies

$$
\mathbb{E} \left| F_K(e^{it_1})F_K(e^{it_2}) \right|^2 = \exp \left( \sum_{k \in K} \frac{1 + \cos(k(t_2 - t_1))}{2k} \right) \ll \sqrt{N} \exp \left( \Re \sum_{k \in K} \frac{e^{i(t_2-t_1)k}}{2k} \right) \ll \frac{\sqrt{N}}{\sqrt{\|t_2 - t_1\|}}.
$$

We can therefore bound (5.2) by

$$
\ll \frac{1}{\sqrt{N}} \int_0^{2\pi} \int_{\frac{\log N^{100}}{N}}^{\pi} \frac{1}{t^{3/2}} \, dt \, dv \ll \frac{1}{(\log N)^{30}}.
$$

Hence, by Markov’s inequality, with probability $\geq 1 - O((\log N)^{-25})$ we have

$$
\frac{1}{4\pi^2} \int_{[0,2\pi] \times [0,2\pi]} \left| F_K(e^{it_1})F_K(e^{it_2})e^{i(t_1(m+1)-t_2(n+1))} \right| \sum_{N < k \leq \frac{4N}{3}} \frac{e^{i(t_2-t_1)k}}{k} \, dt_1 \, dt_2 = O((\log N)^{-25})
$$

as claimed. □
The next step is to restrict to the high-probability event that a certain barrier condition holds. Namely, denote

\[ D_t := \left\{ \sum_{m=1}^{j} (Z_t(m) - 2\sigma_m^2) \leq 12 \log \log N \ \forall \ 1 \leq j \leq \log K \right\}. \]

Then it follows directly from [7, Proposition 5.2], taking \( A = 2 \log \log N \leq \sqrt{\log K} \) and noting \( 10 \log j \leq 10 \log \log N \) for all \( j \), that

\[ \mathbb{P}[D_t \text{ fails for any } t] \ll \frac{1}{(\log N)^2}. \]

We can therefore deduce that with probability \( \geq 1 - O((\log N)^{-2}) \) we in fact have

\[ \sum_{N < k \leq \frac{4N}{3}} \frac{1}{k} \left( \sum_{\sigma \rightarrow n-k} a(\sigma) \right) \left( \sum_{\sigma \rightarrow m-k} \overline{a}(\sigma) \right) \]

\[ = \frac{1}{4\pi^2} \int_{[0,2\pi] \times [0,2\pi]} 1(D_{t_1}) F_K(e^{it_1}) 1(D_{t_2}) F_K(e^{it_2}) e^{i(t_1(m+1) - t_2(n+1))} \sum_{N < k \leq \frac{4N}{3}} \frac{e^{i(t_2-t_1)k}}{k} \ dt_1 \ dt_2 \]

\[ + O((\log N)^{-25}) \]

(since under the event that \( D_t \) holds for all \( t \), the integrals with and without indicators coincide).

Next, we want to restrict further to a strengthened barrier condition that may not hold with high probability, but the contribution to the above integral under the complementary event is nonetheless very small. Namely, set

\[ A_t := \left\{ \sum_{m=1}^{j} (Z_t(m) - 2\sigma_m^2) \leq -2000 \log \log N \ \forall \ 0.01 \log K \leq j \leq \log K \right\} \cap D_t. \]

In order to restrict to this event, we will prove the following analogue of [4, Multiplicative Chaos Result 3].

By a slight abuse of notation, from now on we will write \( \sigma_{m,J}^2 := \sigma_m^2 + J \), that is, we are not shifting by an additional sufficiently large constant anymore.

**Proposition 5.2.** For any \( 0 \leq t \leq 2\pi \) we have

\[ \mathbb{E}[1(D_t) 1(A_t \text{ fails})|F_K(e^{it})|^2] \ll \frac{N(\log \log N)^4}{\log N}. \]

**Proof.** The proof proceeds along similar, though somewhat simpler lines since the discretisation step is hidden in the proof of [7, Proposition 5.2] and, as usual, we do not need any Gaussian approximation statements or prime number estimates.

Note that if \( D_t \) holds but \( A_t \) fails then there must exist some \( 0.01 \log K \leq J \leq \log K \) such that

\[ \sum_{m=1}^{J} (Z_t(m) - 2\sigma_m^2) > -2000 \log \log N. \]
We can therefore bound

$$\mathbb{E}[1(D_t)1(A_t \text{ fails})|F_K(e^{it})|^2] \leq \sum_{0.01 \log K \ll J \ll \log K} \mathbb{E}[1(A_t(J))1(B_t(J))|F_K(e^{it})|^2],$$

where

$$A_t(J) := \left\{ \sum_{m=1}^{j} (Z_t(m) - 2\sigma^2_m) \leq 12 \log \log N \ \forall 1 \leq j \leq J, \ \sum_{m=1}^{J} (Z_t(m) - 2\sigma^2_m) > -2000 \log \log N \right\}$$

and

$$B_t(J) := \left\{ \sum_{m=J+1}^{j} (Z_t(m) - 2\sigma^2_m) \leq 2012 \log \log N \ \forall J + 1 \leq j \leq \log K \right\}.$$

Since $A_t(J)$ only concerns the random variables $Z_t(m)$ with $m \leq J$ and $B_t(J)$ only concerns those with $m > J$, we can factor

$$\mathbb{E}[1(A_t(J))1(B_t(J))|F_K(e^{it})|^2] = \mathbb{E}[1(A_t(J))|F_{e J}^t(e^{it})|^2] \mathbb{E}[1(B_t(J))|F_K/F_{e J}^t(e^{it})|^2].$$

For the second expectation, we can invoke the same idea as in the proof of Proposition 3.6 of writing the expectation as a $(\log K - J)$-fold integral over variables $x_1, ..., x_{\log K - J}$ and substituting $x'_m := x_m - 2\sigma^2_{m,J}$ and we obtain that

$$\left[1(B_t(J)) \left| F_K/F_{e J}^t(e^{it}) \right|^2 \right] = \exp \left( 2 \sum_{m=1}^{\log K - J} \sigma^2_{m,J} \right) \mathbb{P}[B(J)],$$

where $B(J)$ is the event that a sequence of Gaussians $(G_m)_{m=1}^{\log K - J}$ with mean 0 and variances $\sigma^2_{m,J}$ (all between $\frac{1}{20}$ and 20) satisfy

$$\sum_{m=1}^{j} G_m \leq 2012 \log \log N$$

for all $1 \leq j \leq \log K - J$. By Proposition 3.1, we know that $\mathbb{P}[B(J)] \ll \frac{\log \log N}{\sqrt{\log K - J + 1}}$ (this bound is trivial for $J = \log K$). We therefore see, using Proposition 2.3, that

$$\mathbb{E} \left[ 1(B_t(J)) \left| F_K/F_{e J}^t(e^{it}) \right|^2 \right] \ll \frac{K \log \log N}{e^i(\sqrt{\log K - J + 1})}.$$

As for $A_t(J)$, the same idea of writing the expectation as a $J$-fold integral and making the same substitution implies that

$$\mathbb{E}[1(A_t(J))|F_{e J}^t(e^{it})|^2] = \exp \left( 2 \sum_{m=1}^{J} \sigma^2_m \right) \mathbb{P}[A(J)],$$
where $A(J)$ is the event that a sequence of Gaussians $(G_m)_{m=1}^J$ with mean 0 and variances $\sigma^2_{m,J}$ (all between $\frac{1}{20}$ and 20) satisfy
\[
\sum_{m=1}^{J} G_m \leq 12 \log \log N \quad \forall \ 1 \leq j \leq J, \quad \sum_{m=1}^{J} G_m > -2000 \log \log N.
\]

So Proposition 3.3 tells us that $P[B(J)] \ll \frac{(\log \log N)^3}{J^{3/2}}$ and hence
\[
\mathbb{E}[1(A_t(J))|F_{e^t(e^{it})}|^2] \ll \frac{e^t(\log \log N)^3}{(\log K)^{3/2}}.
\]

Putting things together, we have shown that
\[
\mathbb{E}[1(D_t)1(A_t \text{ fails})|F_K(e^{it})|^2] \ll \sum_{0.01 \log K \leq f \leq \log K} \frac{K \log \log N}{e^f(\sqrt{\log K - f + 1})} \frac{e^f(\log \log N)^3}{(\log K)^{3/2}}
\]
\[
\ll \frac{N(\log \log N)^4}{\log N}
\]
as claimed. \[\square\]

Our next goal is to boost the estimate (5.3) to an estimate under the stronger events $A_{t_1}$ and $A_{t_2}$ in place of $D_{t_1}$ and $D_{t_2}$. To this end, note that by Proposition 2.1 the error we pick up in doing so is
\[
\ll \int_{[0,2\pi] \times [0,2\pi]} \left| \frac{1}{N} \left\{ 1, \frac{1}{N \| t_2 - t_1 \|} \right\} \right| dt_1 dt_2.
\]
Using $|xy| \leq x^2 + y^2$ (say) on $x = 1(D_{t_1})1(A_{t_1} \text{ fails})|F_K(e^{it_1})|$ and $y = 1(D_{t_2})1(A_{t_2} \text{ fails})|F_K(e^{it_2})|$, noting the symmetry of the two arising expressions, this term has expectation
\[
\ll \int_0^{\frac{(\log \log N)^{100}}{N}} \mathbb{E}[1(D_t)1(A_t \text{ fails})|F_K(e^{it})|^2] \min \left\{ 1, \frac{1}{Nt} \right\} dt \ll \frac{(\log \log N)^5}{\log N}.
\]

We therefore deduce from Markov’s inequality that with probability $\geq 1 - O((\log N)^{-1/10})$, the error is $\ll \frac{(\log \log N)^5}{(\log N)^{9/10}}$. Hence with probability $\geq 1 - O((\log N)^{-1/10})$, we have (say)
\[
\sum_{N < k \leq 4N^3} \frac{1}{k} \left( \sum_{\sigma - n - k} a(\sigma) \right) \left( \sum_{\sigma - m - k} \overline{a(\sigma)} \right) = \frac{1}{4\pi^2} \int_{[0,2\pi] \times [0,2\pi]} \sum_{N < k \leq 4N^3} \frac{e^{i(t_2 - t_1)k}}{k} dt_1 dt_2 + O\left(\frac{(\log N)^{-17/20}}{N}\right).
\] (5.5)
Proposition 5.3. For any large \( N \) and any positive integer \( 2 \leq p \leq N^{1/4} \) (say), with probability \( \geq 1 - O((\log N)^{-2p}) \) have

\[
\max_{\frac{8N}{7} < m \leq \frac{4N}{3}} \sum_{\frac{8N}{7} < n < \frac{4N}{3}} \int_{[0, 2\pi] \times [0, 2\pi]} \int_{[0, \frac{\log N}{N}]} \left| A_{t_1} F_K(e^{it_1}) 1(A_{t_2} F_K(e^{it_2}) e^{it_1(m+1) - it_2(n+1)} \sum_{N<k \leq \frac{4N}{3}} \frac{e^{i(t_2-t_1)k}}{k} dt_1 dt_2 \right|^{2p} \lesssim N^{2/3} \left( \frac{C \log \log N}{N} \int_0^{2\pi} |F_K(e^{it})|^2 dt \right)^{2p} + N^{2/3} \left( \frac{p^{25} (\log N)^{27}}{(\log N)^{985}} \right)^{2p} + N \left( \frac{p^{25}}{(\log N)^{985}} \right)^{2p}.
\]

We shall postpone the proof of this proposition and first complete the proof of Proposition 4.2 (and thus the main theorems) assuming this proposition, as it happens to be somewhat technical.

Conclusion of the proof of Proposition 4.2. Note first that, for example, by [7, Lemma 2.2], we have

\[
\mathbb{E} \left[ \int_0^{2\pi} |F_K(e^{it})|^2 dt \right] \ll N,
\]

and so by Markov’s inequality we have

\[
\int_0^{2\pi} |F_K(e^{it})|^2 dt \leq N \log N
\]

with probability \( \geq 1 - O((\log N)^{-1}) \). As a consequence, we also have that the quantity on the left-hand side of Proposition 5.3 satisfies

\[
\max_{\frac{8N}{7} < m \leq \frac{4N}{3}} \sum_{\frac{8N}{7} < n < \frac{4N}{3}} \int_{[0, 2\pi] \times [0, 2\pi]} \int_{[0, \frac{\log N}{N}]} \left| A_{t_1} F_K(e^{it_1}) 1(A_{t_2} F_K(e^{it_2}) e^{it_1(m+1) - it_2(n+1)} \sum_{N<k \leq \frac{4N}{3}} \frac{e^{i(t_2-t_1)k}}{k} dt_1 dt_2 \right|^{2p} \lesssim N^{2/3} \left( \frac{p^{25} (\log N)^{27}}{(\log N)^{985}} \right)^{2p} + N \left( \frac{p^{25}}{(\log N)^{985}} \right)^{2p}
\]

with probability \( \geq 1 - O((\log N)^{-1}) \). Under this event, for any \( \frac{8N}{7} < m \leq \frac{4N}{3} \), we have

\[
\int_{[0, 2\pi] \times [0, 2\pi]} \int_{[0, \frac{\log N}{N}]} \left| A_{t_1} F_K(e^{it_1}) 1(A_{t_2} F_K(e^{it_2}) e^{it_1(m+1) - it_2(n+1)} \sum_{N<k \leq \frac{4N}{3}} \frac{e^{i(t_2-t_1)k}}{k} dt_1 dt_2 \right| \leq \frac{1}{\log N}
\]

for all but

\[
\ll N^{2/3} \left( \frac{p^{25} (\log N)^{28}}{(\log N)^{984}} \right)^{2p} + N \left( \frac{p^{25}}{(\log N)^{984}} \right)^{2p}
\]
values of \( \frac{8N}{7} < n \leq \frac{4N}{3} \) and any integer \( 2 \leq p \leq N^{1/4} \). Taking \( p = \left\lfloor \frac{\log N}{B \log \log N} \right\rfloor \) for some absolute constant \( B \), this is
\[
\ll N^{2/3} + 106/B + N^{1-19/18},
\]
which is certainly \( \ll N^{7/10} \) by taking say \( B = 6000 \). Recalling (5.5), this gives the claim. 

It now remains to show Proposition 5.3. Note that its restriction to \( p \geq 2 \) is not really necessary, but since we do not need the case \( p = 1 \) and since some of the indexing later in the proof would require a case distinction, we leave the proof of said case to the interested reader.

**Proof of Proposition 5.3.** Aside from details, the proof is quite similar to that of [4, Proposition 3].

The first step is to expand the 2th power to obtain a 2-fold double integral, say over variables \( u_1, \ldots, u_{2p} \) (corresponding to \( t_1 \)) and \( v_1, \ldots, v_{2p} \) (corresponding to \( t_2 \)). While the arising expression is somewhat complicated, it should be noted that the only part in the above that depends on \( n \) is the factor \( e^{-it_2(n+1)} \), which after expanding the 2th power gives rise to a factor \( e^{-i(v_1 + \cdots + v_p - v_{p+1} - \cdots - v_{2p})(n+1)} \). We can then perform the sum over \( n \), noting that
\[
\sum_{\frac{8N}{7} < n \leq \frac{4N}{3}} e^{-i(v_1 + \cdots + v_p - v_{p+1} - \cdots - v_{2p})(n+1)} \ll \min\left\{ N, \left\| v_1 + \cdots + v_p - v_{p+1} - \cdots - v_{2p} \right\| \right\}.
\]
One then takes absolute values inside the integral. Since the only dependence on \( m \) in the original integral is in the factor \( e^{it_1(m+1)} \), the arising expression in the expanded integral will have absolute value independent of \( m \), so that we can dispose of the maximum. We arrive at the bound
\[
\max_{\frac{8N}{7} < m \leq \frac{4N}{3}} \sum_{\frac{8N}{7} < n \leq \frac{4N}{3}} \left| \int_{[0,2\pi][0,2\pi]} 1(A_t_1)F_K(e^{it_1})1(A_t_2)F_K(e^{it_2})e^{i(n+1)t_1}e^{i(n+1)t_2} \sum_{N < k \leq \frac{4N}{3}} \frac{e^{it_2-kt_1}}{k} dt_1 dt_2 \right|^{2p}
\]
\[
\ll \int \cdots \int \int \cdots \int \prod_{i=1}^{2p} \left( 1(A_{u_i})|F_K(e^{i u_i})|1(A_{v_i})|F_K(e^{i v_i})| \right) \sum_{N < k \leq \frac{4N}{3}} \frac{e^{i(u_1-u_k)}}{k} \right| \right) \times \min\left\{ N, \left\| v_1 + \cdots + v_p - v_{p+1} - \cdots - v_{2p} \right\| \right\} dv_1 \cdots dv_{2p} du_1 \cdots du_{2p}.
\]
In order to deal with this integral, we begin by analysing the part where \( \left\| v_1 + \cdots + v_p - v_{p+1} - \cdots - v_{2p} \right\| \geq N^{-2/3} \), say. On this range, we shall bound the minimum simply by \( N^{2/3} \), and we also discard the indicators. Moreover, we bound
\[
\prod_{i=1}^{2p} |F_K(e^{i u_i})||F_K(e^{i v_i})| \leq \prod_{i=1}^{2p} |F_K(e^{i u_i})|^2 + \prod_{i=1}^{2p} |F_K(e^{i v_i})|^2.
\]
These bounds imply that this portion of the integral is

\[
\ll N^{2/3} \left( \iint_{[0,2\pi] \times [0,2\pi]} |F_K(e^{i t_1})|^2 \left| \sum_{N < k \leq 4N/3} \frac{e^{i(t_2-t_1)k}}{k} \right| dt_1 \ dt_2 \right)^{2p}.
\]

Further, applying Proposition 2.1 to the inner sum (on the range \( \| t_2 - t_1 \| \geq \frac{1}{N} \), otherwise the trivial bound) implies that this is in turn

\[
\ll N^{2/3} \left( C \log \log N \ N \int_0^{2\pi} |F_K(e^{it})|^2 \ dt \right)^{2p},
\]

which is the first term in the required bound.

We will now deal with the portion of the integral where \( \| v_1 + \cdots + v_p - v_{p+1} - \cdots - v_{2p} \| < N^{-2/3} \). On this range, we will simply bound the minimum by \( N \). Note that since we also have \( \| u_i - u_i \| \leq (\log N)^{100} N \), we can see that

\[
\| u_1 + \cdots + u_p - u_{p+1} - \cdots - u_{2p} \| < N^{-2/3} + \frac{2p(\log N)^{100}}{N} \leq \frac{2}{N^{2/3}}.
\]

Again bounding

\[
\prod_{i=1}^{2p} 1(A_{u_i}) |F_K(e^{i u_i})| \prod_{i=1}^{2p} 1(A_{v_i}) |F_K(e^{i v_i})| \leq \prod_{i=1}^{2p} 1(A_{u_i}) |F_K(e^{i u_i})|^2 + \prod_{i=1}^{2p} 1(A_{v_i}) |F_K(e^{i v_i})|^2
\]

and using Proposition 2.1, the contribution from this range is

\[
\ll N \left( \frac{C \log \log N}{N} \right)^{2p} \int \cdots \int 1(A_{t_i}) |F_K(e^{i t_i})|^2 \ dt_1 \ldots dt_{2p}, \quad (5.6)
\]

where we are integrating over the set of \( 0 \leq t_1, \ldots, t_{2p} \leq 2\pi \) satisfying \( \| t_1 + \cdots + t_p - t_{p+1} - \cdots - t_{2p} \| \leq 2N^{-2/3} \). Unfortunately, it seems that some additional technical care is necessary in order to account for some wrap-around issues of variables around the circle.

We shall proceed to subdivide our range of integration. For a permutation \( \sigma \) of length \( 2p \), let

\[
I_\sigma := \{ t_1, \ldots, t_{2p} : \| t_1 + \cdots + t_p - t_{p+1} - \cdots - t_{2p} \| \leq 2N^{-2/3}, 0 < t_{\sigma(1)} < \cdots < t_{\sigma(2p)} < 2\pi \},
\]

so that these sets jointly cover our range of integration. We will sometimes identify \( \sigma(2p + 1) \) with \( \sigma(1) \) and \( \sigma(2p) \) with \( \sigma(0) \). Moreover, for \( 1 \leq h_1, \ldots, h_{2p} \leq \log K \), let \( I_{\sigma, h} \) be the subset of \( I_\sigma \) where

\[
e^{-h_i} \leq \frac{\| t_{\sigma(i+1)} - t_{\sigma(i)} \|}{\pi} \leq e^{-h_{i+1}}
\]
for all $i$, with the convention that the lower bound shall be omitted whenever $h_i = \log K$. Set
\[
\alpha(h) := \begin{cases} 
\frac{(\log N)^{12}}{h}, & \text{if } 0 \leq h < 0.01 \log K, \\
\frac{(\log N)^{-2000}}{h}, & \text{if } 0.01 \log K \leq h \leq \log K,
\end{cases}
\]
and let $1 \leq i_0 \leq 2p$ be an index such that $h_{i_0}$ is minimal.

Then by definition of $\mathbb{A}_i$ we can bound
\[
\prod_{i=1}^{2p} 1(A_{\mathbb{A}_i})|F_K(e^{t_i})|^2 = \prod_{i=1}^{2p} 1(A_{\mathbb{A}_{i_0}})|F_K(e^{t_{i_0}})|^2
\]
\[
\ll \left( \prod_{i=1}^{2p} 1(A_{\mathbb{A}_{i_0}}) \right) \exp \left( 2 \sum_{i=1}^{2p} h_i \sum_{m=1}^{\log K} Z_{\mathbb{A}_{i_0}}(m) + 2 \sum_{m=1}^{\log K} Z_{\mathbb{A}_{i_0}}(m) \right)
\]
\[
\ll \left( \prod_{i=1}^{2p} \alpha(h_i)^2 e^{2h_i} \right) \exp \left( 2 \sum_{i=1}^{2p} h_i \sum_{m=h_i+1}^{\log K} \sum_{m=1}^{\log K} Z_{\mathbb{A}_{i_0}}(m) + 2 \sum_{m=1}^{\log K} Z_{\mathbb{A}_{i_0}}(m) \right)
\]
\[
\ll N \left( \frac{C \log \log N}{N} \right)^{2p} \sum_{\sigma \in S_{2p}} \sum_{0 \leq h_1, ..., h_{2p} \leq \log K} \left( \prod_{i=1 \neq i_0}^{2p} \alpha(h_i)^2 e^{2h_i} \right)
\]
\[
\times \int \cdots \int _{I_{\mathbb{A}, h}} \exp \left( 2 \sum_{i=1 \neq i_0}^{2p} \sum_{m=h_i+1}^{\log K} \sum_{m=1}^{\log K} Z_{\mathbb{A}_{i_0}}(m) + 2 \sum_{m=1}^{\log K} Z_{\mathbb{A}_{i_0}}(m) \right) dt_1 \ldots dt_{2p}.
\]
We will estimate the expectation of this quantity with the intent of later applying Markov’s inequality. Note that by Proposition 2.3 the expectation of the expression inside the integral is
\[
\ll \exp \left( 2 \sum_{i=1 \neq i_0}^{2p} \sum_{m=h_i+1}^{\log K} \sum_{m=1}^{\log K} \frac{\cos(k(t_{\mathbb{A}_i(j)} - t_{\mathbb{A}_j(i)}))}{k} \right)
\]
(with the convention that the term \( k = 1 \) is included in the last inner sum whenever \( h_i = h_j = 0 \)).

Note that if \( i \) (respectively \( j \)) is equal to \( i_0 \) then the last inner sum would start at \( e^{h_j} \) (respectively \( e^{h_i} \)), but since we chose \( h_{i_0} \) to be minimal, we can include it in the maximum. Since the sum is empty whenever \( h_i \) or \( h_j \) are equal to \( \log K \), Proposition 2.1 implies that this is in turn

\[
\ll e^{O(p)} K \prod_{i=1 \atop i \neq i_0}^{2p} \left( \frac{K}{e^{h_i}} \right) \exp \left( 6\pi \sum_{1 \leq i < j \leq 2p \atop h_i, h_j \neq \log K} \frac{1}{\| t_{\sigma(j)} - t_{\sigma(i)} \| e^{\max(h_i, h_j)}} \right). \tag{5.7}
\]

In order to understand the sum inside the exponential, we will split it into a sum over \( i \) and \( j \) such that \( 0 < t_{\sigma(j)} - t_{\sigma(i)} \leq \pi \) and its complementary sum, where \( \pi < t_{\sigma(j)} - t_{\sigma(i)} < 2\pi \). Note that on the former range we have

\[
\| t_{\sigma(j)} - t_{\sigma(i)} \| = \| t_{\sigma(j)} - t_{\sigma(j-1)} \| + \cdots + \| t_{\sigma(i+1)} - t_{\sigma(i)} \| \\
\geq \pi (e^{-h_{j-1}} 1(h_{j-1} \neq \log K) + \cdots + e^{-h_{i+1}} 1(h_{i+1} \neq \log K) + e^{-h_i}).
\]

On the latter range on the other hand, we can wrap around the other way to obtain that

\[
\| t_{\sigma(j)} - t_{\sigma(i)} \| \geq \pi (e^{-h_{i-1}} 1(h_{i-1} \neq \log K) + \cdots + e^{-h_{j+1}} 1(h_{j+1} \neq \log K) + e^{-h_j}),
\]

where the sum wraps from index 1 around to index \( 2p \). On the first range, we will first fix \( i \) and sum over \( j \), whereas on the second range we proceed conversely; since the arguments are symmetrical, we will restrict ourselves to the first case. We have

\[
\sum_{1 \leq i < j \leq 2p \atop h_i, h_j \neq \log K} \frac{1}{\| t_{\sigma(j)} - t_{\sigma(i)} \| e^{\max(h_i, h_j)}} \\
\leq \frac{1}{\pi} \sum_{i} \sum_{j} \left( e^{-h_i} + e^{-h_{i+1}} 1(h_{i+1} \neq \log K) + \cdots + e^{-h_{j-1}} 1(h_{j-1} \neq \log K) \right) e^{\max(h_i, h_j)},
\]

where the ranges of summation coincide. We will now subdivide the sum in the denominator into blocks of consecutive indices \( l \) with \( h_l \neq \log K \). We choose the lengths of these blocks inductively in the following way. We formally set \( l_0 = i \). Supposing that the \( m \)th block ends at an index \( l_m \), we take \( l_{m+1} \) (at most the largest value of \( j \) in the above summation) minimal such that

\[
e^{-h_{m+1}} 1(h_{m+1} \neq \log K) + \cdots + e^{-h_{m+1}} 1(h_{m+1} \neq \log K) \geq e^{-h_i}
\]

if such index exists, and otherwise end the process. Note that within a given block, only the last index \( l_m \) can have the property that \( h_{l_m} \geq h_i \). We can therefore bound

\[
\sum_{j \text{ in block } m} \frac{1}{e^{\max(h_i, h_j)}} \ll 2e^{-h_i}.
\]
Moreover, we have
\[ e^{-h_i} + e^{-h_{i+1}} 1(h_{i+1} \neq \log K) + \cdots + e^{-h_{j-1}} 1(h_{j-1} \neq \log K) \geq m e^{-h_i} \]
uniformly over \( j \) in block \( m \). Since there are at most \( 2p \) blocks, we therefore obtain, for any \( i \), that
\[
\sum_j 1 \left( e^{-h_i} + e^{-h_{i+1}} 1(h_{i+1} \neq \log K) + \cdots + e^{-h_{j-1}} 1(h_{j-1} \neq \log K) \right) e^{\max(h_i, h_j)} 
\leq \sum_{m=1}^{2p} \frac{2}{m} \leq 2 \log(2p) + O(1).
\]
Summing this in turn over \( i \) gives a bound \( \leq 4p \log(2p) + O(p) \). Including the range where \( \pi < t_{\sigma(j)} - t_{\sigma(i)} < 2\pi \) and the factor \( \frac{1}{\pi} \), this shows that the expression in (5.7) is
\[
\ll e^{O(p)} p^{48p} K \prod_{i=1}^{2p} \left( \frac{K}{e^{h_i}} \right).
\]
So far, we have therefore shown that the expectation of the expression in (5.6) is
\[
\ll N \left( Cp^{24} \log \log N \right)^{2p} \sum_{\sigma \in S_{2p}} \sum_{0 \leq h_1, \ldots, h_{2p} \leq \log K} \left( \prod_{i=1}^{2p} \alpha(h_i) e^{h_i} \right) vol(I_{\sigma, h}).
\]
We can proceed in two ways in order to bound the volume of \( I_{\sigma, h} \). One such way would be to simply note that after placing \( t_{\sigma(i_0)} \) arbitrarily, our conditions on the distances of the \( t \) variables gives
\[
vol(I_{\sigma, h}) \leq e^{O(p)} \prod_{i=1}^{2p} e^{-h_i}.
\]
Another way would be to start in the same way, but for an arbitrary index \( l \) with \( l \neq i_0, i_0 + 1 \) we instead use the condition \( \|t_1 + \cdots + t_p - t_{p+1} - \cdots - t_{2p}\| \leq 2N^{-2/3} \) on the variable \( t_{\sigma(l)} \). This gives the bound
\[
vol(I_{\sigma, h}) \ll e^{O(p)} N^{-2/3} \min_{1 \leq l \leq 2p} \prod_{l \neq i_0, i_0+1} e^{-h_i}.
\]
We hence obtain that the expectation of (5.6) is
\[
\ll N \left( Cp^{24} \log \log N \right)^{2p} \sum_{\sigma \in S_{2p}} \sum_{0 \leq h_1, \ldots, h_{2p} \leq \log K} \left( \prod_{i=1}^{2p} \alpha(h_i) \right) \min \left\{ 1, N^{-2/3} \min_{1 \leq l \leq 2p} e^{h_l + h_{l+1}} \right\}.
\]
If there is \( l \neq i_0, i_0 + 1 \) such that \( e^{h_i}, e^{h_{i+1}} \leq N^{1/6} \), then we can bound the minimum in the above expression by \( N^{-1/3} \) and bound the product over \( \alpha \) trivially by \( (\log N)^{48p} \). If on the other hand
there is no such \( l \), then at least half of the remaining \( 2p - 2 \) indices must satisfy \( e^{h_i} > N^{1/6} \) and thus in particular \( h_i > 0.01 \log K \). In that case we bound the minimum trivially by 1 and
\[
\prod_{i=1}^{2p} \alpha(h_i)^2 \leq (\log N)^{24p}(\log N)^{-4000(p-1)} \leq (\log N)^{-988} \] \] ^{2p}.

We thus see, after noting that the number of permutations \( \sigma \) is \( \leq e^{O(p)p^{2p}} \), that the expectation of (5.6) is
\[
\ll N^{2/3}(p^{25}(\log N)^{26})^{2p} + N\left(\frac{p^{25}}{(\log N)^{986}}\right)^{2p}.
\]

Therefore, Markov’s inequality implies that with probability \( \geq 1 - O((\log N)^{-2p}) \), the expression (5.6) is
\[
\ll N^{2/3}(p^{25}(\log N)^{27})^{2p} + N\left(\frac{p^{25}}{(\log N)^{985}}\right)^{2p},
\]
thus implying the claim. \( \square \)

ACKNOWLEDGEMENTS
This work received funding from the Göran Gustafsson Foundation and from the Swedish Research Council Grant No. 2016-05198.

JOURNAL INFORMATION
Mathematika is owned by University College London and published by the London Mathematical Society. All surplus income from the publication of Mathematika is returned to mathematicians and mathematics research via the Society’s research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES
1. D. Aggarwal, U. Subedi, W. Verreault, A. Zaman, and C. Zheng, A conjectural asymptotic formula for multiplicative chaos in number theory, Res. Number Theory 8 (2022), 35.
2. A. J. Harper, On the partition function of the Riemann zeta function, and the Fyodorov–Hiary–Keating conjecture, arXiv:1906.05783 (2019).
3. A. J. Harper, Moments of random multiplicative functions, I: Low moments, better than squareroot cancellation, and critical multiplicative chaos, Forum Math. Pi 8 (2020), e1.
4. A. J. Harper, Almost sure large fluctuations of random multiplicative functions, Int. Math. Res. Not. 11 (2021), rnab299.
5. Y.-K. Lau, G. Tenenbaum, and J. Wu, On mean values of random multiplicative functions, Proc. Amer. Math. Soc. 141 (2013), no. 2, 409–420.
6. D. Mastrostefano, An almost sure upper bound for random multiplicative functions on integers with a large prime factor, Electron. J. Probab. 27 (2022), 1–21.
7. K. Soundararajan and A. Zaman, A model problem for multiplicative chaos in number theory, Enseign. Math. 68 (2022), no. 3/4, 307–340.
8. R. M. Young, Euler’s constant, Math. Gaz. 75 (1991), no. 472, 187–190.