ACTA ARITHMETICA
225.1 (2026)

# Correlations of error terms for weighted prime counting functions

by

Shubhrajit Bhattacharya, Greg Martin
and Reginald M. Simpson

**Abstract.** Standard prime-number counting functions, such as $\psi(x)$, $\theta(x)$, and $\pi(x)$, have error terms with limiting logarithmic distributions once suitably normalized. The same is true of weighted versions of those sums, like $\pi_r(x) = \sum_{p\le x} \frac{1}{p}$ and $\pi_\ell(x) = \sum_{p\le x} \log\bigl(1-\frac{1}{p}\bigr)^{-1}$, that were studied by Mertens. These limiting distributions are all identical, but passing to the limit loses information about how these error terms are correlated with one another.

In this paper, we examine these correlations, showing, for example, that persistent inequalities between certain pairs of normalized error terms are equivalent to the Riemann hypothesis (RH). Assuming both RH and LI, the linear independence of the positive imaginary parts of the zeros of $\zeta(s)$, we calculate the logarithmic densities of the set of real numbers on which two different error terms have prescribed signs. For example, we conditionally show that $\psi(x) - x$ and $\sum_{n\le x} \frac{\Lambda(n)}{n} - (\log x - C_0)$ have the same sign on a set of logarithmic density $\approx 0.9865$.

## Contents

1. Introduction . . . . . 2
2. Explicit formulas for normalized error terms . . . . . 11
3. Persistent inequalities between normalized error terms . . . . . 15
4. Fourier transforms of distributions and densities . . . . . 29
5. A formula for densities related to $\eta_2$ . . . . . 37
6. Calculation and error analysis . . . . . 44
Appendix A. Graphs illustrating the main theorems . . . . . 65
References . . . . . 71

2020 *Mathematics Subject Classification*: Primary 11N64; Secondary 11M06, 11M26, 11Y60.
*Key words and phrases*: prime counting functions, error terms, logarithmic distributions, multivariate distributions.
Received 16 July 2025; revised 12 February 2026.
Published online 10 September 2026 in Open Access (under CC-BY license).

DOI: 10.4064/aa250716-6-5  © Instytut Matematyczny PAN, 2026

**1. Introduction.** A central topic in comparative prime number theory is the relationship between the prime counting function $\pi(x)$ and the logarithmic integral $\mathrm{li}(x) = \int_0^x \frac{dx}{\log x}$, which are asymptotic to each other by the prime number theorem. While empirical data led many to believe that $\pi(x) < \mathrm{li}(x)$ for all $x \geq 2$, Littlewood [10] proved that $\pi(x) - \mathrm{li}(x)$ has infinitely many sign changes. Bays and Hudson [3] gave convincing evidence that the first value of $x$ for which $\pi(x) > \mathrm{li}(x)$ (the true "Skewes number" for this prime number race) is approximately $1.4 \times 10^{316}$. Assuming the Riemann hypothesis (RH), Rubinstein and Sarnak [15], extending work of Wintner [18], showed that a normalized version of the difference $\pi(x) - \mathrm{li}(x)$ has a limiting logarithmic distribution (see Definition 1.15 below); under the additional hypothesis (LI) that the positive numbers $\gamma$ for which $\zeta\big(\frac{1}{2} + i\gamma\big) = 0$ are linearly independent over the rationals, they calculated that the logarithmic density (see Definition 1.7 below) of the set of $x$ for which $\pi(x) > \mathrm{li}(x)$ is a constant $\eta_1$ which is approximately $2.6 \times 10^{-6}$.

These methods extend to various weighted versions of the prime counting function, such as $\theta(x) = \sum_{p \leq x} \log p$, which is asymptotic to $x$ again by the prime number theorem. Assuming RH and LI, a suitably normalized version of $\theta(x) - x$ has the same limiting logarithmic distribution as the one associated to $\pi(x) - \mathrm{li}(x)$, and Rubinstein and Sarnak's work also shows that the logarithmic density of the set of $x$ for which $\theta(x) > x$ is the same constant $\eta_1$. Moreover, Platt and Trudgian [13] gave evidence that the first value of $x$ for which $\theta(x) > x$ (the "Skewes number" for this weighted prime number race) is also approximately $1.4 \times 10^{316}$, though presumably not exactly the same value as for the inequality $\pi(x) > \mathrm{li}(x)$. (See Remark 4.6 below for an insight as to why these limiting distributions are the same.)

The same phenomenon can be observed when examining the Mertens sum $\sum_{p \leq x} \frac{1}{p} \sim \log\log x - C_2$ (for a particular constant $C_2$) and the Mertens product $\prod_{p \leq x} \big(1 - \frac{1}{p}\big)^{-1} \sim e^{C_0} \log x$ (for the Euler–Mascheroni constant $C_0$). In both cases, all calculated values of the prime-related quantities were greater than the corresponding values of their asymptotic main terms; but Robin [14] showed that both differences between the prime-related quantities and their main terms have infinitely many sign changes. Work of Büthe [4] suggests that the first negative value of $\sum_{p \leq x} \frac{1}{p} - (\log\log x - C_2)$ (the "Skewes number" for this Mertens sum race) is approximately $1.9 \times 10^{215}$. In addition, assuming RH and LI, Lay [9] and Lamzouri [8] have shown that appropriate normalizations of these differences have limiting logarithmic distributions that are identical (up to an additive constant) to the distribution associated to $\pi(x) - \mathrm{li}(x)$ and $\theta(x) - x$, and that the logarithmic density of the set of rare sign reversals is the same constant $\eta_1$.

All of these observations raise the question of how these various races are correlated. Each of the four races has an exceptional set (of logarith-

mic density $\eta_1$); for $\pi(x)$ and $\theta(x)$ these sets correspond to regions with significantly more primes than average, while for the Mertens sum and product they correspond to regions with significantly fewer primes than average. Thus, we might expect the first two exceptional sets to be highly correlated (and the same for the second two), but also for the two pairs of exceptional sets to be highly anticorrelated (this idea appears, for example, at the beginning of [4]). Moreover, each of these races has an "unbiased" version involving counting prime powers rather than only primes—the simplest such versions (which were examined by Kaczorowski [7], for example) involve $\psi(x) = \sum_{n\le x} \Lambda(n) \sim x$ and the corresponding Mertens sum $\sum_{n\le x} \frac{\Lambda(n)}{n} \sim \log x - C_0$, where $\Lambda(n)$ is the von Mangoldt function. Both of these functions lie above and below their asymptotic main terms equally often (on the logarithmic scale, assuming RH), and we can ask how correlated the sets are where the two functions exceed their respective main terms, for example.

The purpose of this paper is to answer questions like these concerning the correlations of these error terms in prime number-related races. For example, assuming RH and LI, we show:

- the sets $\{x \ge 2\colon \pi(x) > \operatorname{li}(x)\}$ and $\{x \ge 2\colon \theta(x) > x\}$ are the same up to a set of logarithmic density 0;
- the sets $\{x \ge 2\colon \pi(x) > \operatorname{li}(x)\}$ and $\left\{x \ge 2\colon \sum_{p\le x} \frac{1}{p} > \log\log x - C_2\right\}$ have bounded (and presumably empty) intersection;
- but the $\psi(x)$ and $\sum_{n\le x} \frac{\Lambda(n)}{n}$ races are highly, yet not perfectly correlated: the set of $x \ge 2$ such that $\psi(x) - x$ and $\sum_{n\le x} \frac{\Lambda(n)}{n} - (\log x - C_0)$ have the same sign has a logarithmic density which is approximately 0.9865.

We now proceed to set up the notation and terminology required to make these statements rigorous.

**1.1. Preliminaries.** This section introduces the basic definitions and notation that will be used and referred to frequently throughout the rest of this article.

Definition 1.1. We use the following well-known prime-counting functions:

$$\psi(x) = \sum_{n\le x} \Lambda(n), \qquad \theta(x) = \sum_{p\le x} \log p,$$
$$\Pi(x) = \sum_{n\le x} \frac{\Lambda(n)}{\log n} = \sum_{p^k\le x} \frac{1}{k}, \qquad \pi(x) = \sum_{p\le x} 1.$$

We denote by $\mathcal{B} = \{\psi, \theta, \Pi, \pi\}$ the set of these four *standard functions*.

Furthermore, we define the following *weighted versions* of the above standard functions:

$$\psi_r(x) = \sum_{n\le x} \frac{\Lambda(n)}{n}, \qquad \theta_r(x) = \sum_{p\le x} \frac{\log p}{p},$$
$$\Pi_r(x) = \sum_{n\le x} \frac{\Lambda(n)}{n\log n} = \sum_{p^k\le x} \frac{1}{kp^k}, \qquad \pi_r(x) = \sum_{p\le x} \frac{1}{p}.$$

We also define the *logarithm of the Mertens product*

$$\pi_\ell(x) = \log \prod_{p\le x} \left(1-\frac{1}{p}\right)^{-1} = -\sum_{p\le x} \log\left(1-\frac{1}{p}\right) = \sum_{p\le x}\sum_{k=1}^{\infty} \frac{1}{kp^k},$$

which behaves in many ways like the four weighted versions. We denote by

$$\mathcal{B}_r = \{\psi_r, \theta_r, \Pi_r, \pi_r, \pi_\ell\}$$

the set of these five *reciprocal functions*.

Each of the five reciprocal functions in $\mathcal{B}_r$ has a well-known asymptotic formula, proved by Mertens with elementary arguments, while each of the four standard functions in $\mathcal{B}$ also has a well-known asymptotic formula that is equivalent to the prime number theorem. When the main terms of these asymptotic formulas are subtracted, the resulting error terms have explicit formulas involving a sum over the zeros of the Riemann zeta function (see Section 2); when further divided by the order of magnitude of any one of these summands, the resulting normalized error terms are known (conditionally) to have limiting logarithmic distribution functions. We record these normalized error terms explicitly.

Definition 1.2. The normalized error terms of the four standard functions are

$$E^{\psi}(x) = \frac{\psi(x)-x}{\sqrt{x}}, \qquad E^{\Pi}(x) = \frac{\Pi(x)-\mathrm{li}(x)}{\sqrt{x}/\log x},$$
$$E^{\theta}(x) = \frac{\theta(x)-x}{\sqrt{x}}, \qquad E^{\pi}(x) = \frac{\pi(x)-\mathrm{li}(x)}{\sqrt{x}/\log x}.$$

The normalized error terms of the five reciprocal functions are

$$E^{\psi_r}(x) = \big(\psi_r(x) - (\log x - C_0)\big)\sqrt{x},$$
$$E^{\Pi_r}(x) = \big(\Pi_r(x) - (\log\log x + C_0)\big)\sqrt{x}\log x,$$
$$E^{\theta_r}(x) = \big(\theta_r(x) - (\log x - C_1)\big)\sqrt{x},$$
$$E^{\pi_r}(x) = \big(\pi_r(x) - (\log\log x - C_2)\big)\sqrt{x}\log x,$$
$$E^{\pi_\ell}(x) = \big(\pi_\ell(x) - (\log\log x + C_0)\big)\sqrt{x}\log x.$$

Here $C_0$ is Euler's constant, while $C_1 = C_0 + \sum_p \frac{\log p}{p(p-1)}$ and $C_2 = -C_0 + \sum_{k=2}^{\infty} \sum_p \frac{1}{kp^k}$ (although the exact values will usually not be important to us).

Some of these normalized error terms are biased towards positive or negative values. (Loosely speaking, nature wants to count all prime powers, and it is our human insistence on throwing away squares and higher powers of primes that creates biases in functions that count only primes themselves.) These biases manifest as constant terms in the explicit formulas given in Section 2, which we now name.

DEFINITION 1.3. For the nine functions in $\mathcal{B} \cup \mathcal{B}_r$, we define the *bias terms*

$$\beta_\theta = \beta_\pi = -1, \quad \beta_\psi = \beta_\Pi = \beta_{\psi_r} = \beta_{\Pi_r} = 0, \quad \beta_{\theta_r} = \beta_{\pi_r} = \beta_{\pi_\ell} = 1.$$

These constants are the average values of the corresponding normalized error terms (in ways that can be made precise).

We can now state our first result, which concerns the ranges of values taken by differences of these normalized error terms. The following theorem is proved at the end of Section 2.

THEOREM 1.4. *Assume RH.*

(a) *If $f$ and $g$ are both standard functions, then*

$$E^f(x) - E^g(x) = \beta_f - \beta_g + o(1).$$

*The same is true if $f$ and $g$ are both reciprocal functions.*

(b) *If $f$ is a reciprocal function and $g$ is a standard function, then*

$$\liminf_{x\to\infty}(E^f(x) - E^g(x)) \geq \beta_f - \beta_g - w,$$
$$\limsup_{x\to\infty}(E^f(x) - E^g(x)) \leq \beta_f - \beta_g + w,$$

*where*

$$w = 2 + C_0 - \log 4\pi \approx 0.0461914. \tag{1.1}$$

REMARK 1.5. If we assume LI in addition to RH, then both inequalities in Theorem 1.4(b) are actually equalities, as we show in Theorem 1.17(b) below.

Next, we examine the question of when one normalized error term always exceeds a second when $x$ is sufficiently large; in some cases, such a persistent inequality can be proved unconditionally, but in other cases it is equivalent to RH. The two parts of the following theorem are proved at the ends of Sections 3.1 and 3.3.

THEOREM 1.6.

(a) *Unconditionally, there exists an absolute constant $x_0$ such that each of the inequalities*

$$E^{\theta}(x) < E^{\psi}(x), \quad E^{\pi}(x) < E^{\Pi}(x), \quad E^{\psi_r}(x) < E^{\theta_r}(x),$$
$$E^{\Pi_r}(x) < E^{\pi_\ell}(x) < E^{\pi_r}(x)$$

*holds for all $x > x_0$.*

(b) *Suppose the two functions $f, g \in \mathcal{B} \cup \mathcal{B}_r$ satisfy $\beta_f < \beta_g$. If*

$$(f,g) \notin \big\{(\theta,\psi),\ (\pi,\Pi),\ (\psi_r,\theta_r),\ (\Pi_r,\pi_r),\ (\Pi_r,\pi_\ell)\big\} \tag{1.2}$$

*(that is, if $f$ and $g$ do not form a pair of functions covered by the inequalities in part* (a) *of this theorem), then RH is equivalent to the existence of an absolute constant $x_0$ such that $E^f(x) < E^g(x)$ for all $x > x_0$. More generally, RH is equivalent to the difference $E^g(x) - E^f(x)$ being bounded either above or below.*

**1.2. Main results.** The appropriate method of measuring sets of real numbers involving these error terms turns out to be the (limiting) logarithmic density.

DEFINITION 1.7. For any set $P \subset [2,\infty)$, we define

$$\delta(P) = \lim_{X\to\infty} \frac{1}{\log X} \int\limits_{t\in P\cap[2,X]} \frac{dt}{t}$$

(when the limit exists) to be the *logarithmic density* of $P$. Given functions $f$ and $g$, we use the shorthand notation

$$\delta(f > g) = \delta\big(\{x \ge 2 \colon f(x) > g(x)\}\big),$$

and similarly for chained inequalities involving three or more functions.

NOTATION 1.8. Throughout this paper, $\gamma_1 \le \gamma_2 \le \cdots$ denotes the sequence of positive imaginary parts of $\zeta(s)$, listed with multiplicity if necessary, so that $\gamma_1 \approx 14.135$ for example.

DEFINITION 1.9. Define the real-valued random variable

$$V_1 = \sum_{m=1}^{\infty} 2\Re\frac{Z_m}{\frac12 + i\gamma_m} \quad \text{or equivalently} \quad V_1 = \sum_{m=1}^{\infty} \frac{2\Re Z_m}{\sqrt{\frac14 + \gamma_m^2}},$$

where the $Z_m$ are independent random variables, each uniformly distributed on the unit circle in $\mathbb{C}$. Then, define $\mu_1$ to be the probability distribution associated to $V_1$ (which is symmetric about the origin), and define the constant $\eta_1$ by

$$\eta_1 = \Pr(V_1 > 1) = \int\limits_1^{\infty} d\mu_1.$$

Assuming RH, Rubinstein and Sarnak [15] proved that $\mu_1$ has support equal to all of $\mathbb{R}$ and is absolutely continuous with respect to Lebesgue measure, and calculated that $\eta_1 \approx 2.6 \times 10^{-6}$.

REMARK 1.10. Assuming RH and LI, Rubinstein and Sarnak further proved that the limiting logarithmic distributions of $E^\pi$ and $E^\theta$ (in the sense of Definition 1.15 below) are the same as the distribution of $V_1 - 1$, so that in particular $\delta(\pi(x) > \mathrm{li}(x)) = \delta(\theta(x) > x) = \eta_1$.

Indeed, for each of the nine functions $f \in \mathcal{B} \cup \mathcal{B}_r$, it is known (assuming RH and LI) that the limiting logarithmic distribution of $E^f$ is the same as the distribution of $V_1 + \beta_f$ (which justifies the use of the name "bias term"), so that all nine individual distributions are identical up to the additive constants $\beta_f$. We collect these results in Corollary 4.5.

DEFINITION 1.11. Define the $\mathbb{R}^2$-valued random variable

$$V_2 = \sum_{m=1}^{\infty} \left( 2\Re \frac{Z_m}{\frac{1}{2} + i\gamma_m}, 2\Re \frac{Z_m}{-\frac{1}{2} + i\gamma_m} \right),$$

with the $Z_m$ as in Definition 1.9. Note that $V_2$ is invariant under reflection in either diagonal line $y = \pm x$ (since the distribution of $Z_m$ is invariant under conjugation and negation). Note similarly that each coordinate has the same distribution as $V_1$, although the two coordinates are not independent.

Then, define $\mu_2$ to be the probability distribution associated to $V_2$, and define the constant $\eta_2$ by

$$\eta_2 = \Pr\big(\text{the two coordinates of } V_2 \text{ have different signs}\big)$$

(so that $\eta_2$ is the mass that $\mu_2$ assigns to the union of the second and fourth open quadrants).

Assuming RH, we can calculate (see the end of Section 6) that $\eta_2 \approx 0.01355$. Assuming RH and LI, this distribution is connected to the limiting logarithmic distribution of vector-valued functions such as $\big(E^\psi(x), E^{\psi_r}(x)\big)$, as evidenced in Theorems 1.14 and 1.17(b) below.

Our first three main considerations are the probabilities (in the sense of logarithmic density) of two normalized error terms being jointly positive, jointly negative, or having mixed signs. These probabilities depend upon the bias terms of the functions in question, as well as whether the two functions are standard or reciprocal functions. These three results are proved in Section 4.

The first result of this type, which uses the quantity $\eta_1$ from Definition 1.9, covers the cases where the bias terms are unequal.

THEOREM 1.12. *Assume RH and LI. Let* $f, g \in \mathcal{B} \cup \mathcal{B}_r$.

(a) *If* $\beta_f = -1$ *and* $\beta_g = 0$ (*for example, when* $f = \pi$ *and* $g = \psi$), *then*
$\delta(0 < E^f < E^g) = \eta_1,\ \delta(E^f < 0 < E^g) = \frac{1}{2} - \eta_1,\ \delta(E^f < E^g < 0) = \frac{1}{2}$.

(b) *If* $\beta_f = 0$ *and* $\beta_g = 1$ (*for example, when* $f = \psi$ *and* $g = \pi_r$), *then*
$\delta(0 < E^f < E^g) = \frac{1}{2},\ \delta(E^f < 0 < E^g) = \frac{1}{2} - \eta_1,\ \delta(E^f < E^g < 0) = \eta_1$.

(c) *If* $\beta_f = -1$ *and* $\beta_g = 1$ (*for example, when* $f = \pi$ *and* $g = \pi_r$), *then*
$\delta(0 < E^f < E^g) = \eta_1,\ \delta(E^f < 0 < E^g) = 1 - 2\eta_1,\ \delta(E^f < E^g < 0) = \eta_1$.

Note that in all cases covered by Theorem 1.12, we have $\delta(E^f \geq E^g) = 0$ by Theorem 1.6(b), and so the given sets of inequalities are exhaustive.

Our second such result concerns two functions with equal bias terms, where either both functions are standard or both functions are reciprocal.

THEOREM 1.13. *Assume RH and LI.*

(a) *We have* $\delta(0 < E^\pi$ *and* $0 < E^\theta) = \eta_1$ *and* $\delta(E^\pi < 0$ *and* $E^\theta < 0) = 1 - \eta_1$.

(b) *We have* $\delta(0 < E^\psi$ *and* $0 < E^{\Pi}) = \delta(E^\psi < 0$ *and* $E^{\Pi} < 0) = \frac{1}{2}$, *and the same if* $\psi$ *and* $\Pi$ *are both replaced by* $\psi_r$ *and* $\Pi_r$.

(c) *If* $f, g \in \{\theta_r, \pi_r, \pi_\ell\}$ *then*
$$\delta(0 < E^f \text{ and } 0 < E^g) = 1 - \eta_1 \quad \text{and} \quad \delta(E^f < 0 \text{ and } E^g < 0) = \eta_1.$$

Finally, our third and most novel main result concerns two functions with equal bias terms, where one function is standard and the other reciprocal. The statement uses the constant $\eta_2$ from Definition 1.11.

THEOREM 1.14. *Assume RH and LI. Suppose that* $f$ *is a standard function and* $g$ *is a reciprocal function with* $\beta_f = \beta_g = 0$ (*for example,* $f = \psi$ *and* $g = \psi_r$). *Then*

$$\delta(E^f < 0 < E^g) = \delta(E^g < 0 < E^f) = \frac{\eta_2}{2}$$

*and*

$$\begin{aligned}\delta(E^f < E^g < 0) &= \delta(E^g < E^f < 0) = \delta(0 < E^f < E^g) \\ &= \delta(0 < E^g < E^f) = \frac{1 - \eta_2}{4}.\end{aligned}$$

The regions corresponding to these six sets of inequalities can be visualized in the bottom graph in Figure 1 below.

Our primary tool for evaluating these densities is the limiting logarithmic distribution of a real- or vector-valued function, which we now define formally.

DEFINITION 1.15. We say that a vector-valued function $\vec{\phi}\colon [2, \infty) \to \mathbb{R}^d$ has a *limiting logarithmic distribution* $\mu$ on $\mathbb{R}^d$ if $\mu$ is a probability measure

on $\mathbb{R}^d$ and

$$\lim_{X\to\infty} \int_2^X f\big(\vec{\phi}(x)\big)\,\frac{dx}{x} = \int_{\mathbb{R}^\ell} f(\vec{x})\,d\mu(\vec{x})$$

for all bounded continuous functions $f\colon \mathbb{R}^d \to \mathbb{C}$.

REMARK 1.16. Logarithmic density measures a set of inputs to a function, while logarithmic distributions correspond to the frequency of the function's outputs in various regions. One way they are related is as follows: for any vector-valued function $\vec{\phi}(t)$ possessing a limiting logarithmic distribution $\mu$, and any subset $T \subset \mathbb{R}^d$, set $S = \{x \in [2,\infty)\colon \vec{\phi}(x) \in T\}$. Then $\delta(S) = \mu(T)$, provided that the boundary of $T$ satisfies $\mu(\partial T) = 0$. (This proviso is necessary because the characteristic function of $T$ is not continuous.) In practice, $\mu(\partial T) = 0$ is automatically satisfied if the limiting distribution $\mu$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$, as long as $\partial T$ has Lebesgue measure 0 (which, for example, is guaranteed if $\partial T$ is contained in the union of finitely many hyperplanes).

We are able to determine the precise joint limiting logarithmic distribution of any pair of normalized error terms under consideration, as our final main result describes.

THEOREM 1.17. *Assume RH and LI. Let $f, g \in \mathcal{B} \cup \mathcal{B}_r$.*

(a) *If $f$ and $g$ are both standard functions, then the limiting logarithmic distribution of $\big(E^f(x), E^g(x)\big)$ is singular in $\mathbb{R}^2$ with its support equaling the line $y - \beta_g = x - \beta_f$. The same is true if $f$ and $g$ are both reciprocal functions.*

(b) *If $f \in \mathcal{B}$ and $g \in \mathcal{B}_r$ (or vice versa) then the joint limiting logarithmic distribution of $\big(E^f(x), E^g(x)\big)$ is absolutely continuous on $\mathbb{R}^2$, with support equal to the narrow diagonal strip*

$$\big\{(x,y) \in \mathbb{R}^2\colon |(x-\beta_f) - (y-\beta_g)| \le w\big\}.$$

*In this case, the joint distribution of $(E^f(x), E^g(x)) - (\beta_f, \beta_g)$ equals $\mu_2$ as in Definition* 1.11.

Two examples of Theorem 1.17(b) are shown in Figure 1 (other examples of such graphs appear in Appendix A). This theorem is an immediate consequence of Proposition 4.10, which in fact determines the joint logarithmic distribution of our nine normalized error terms simultaneously.

In Section 2 we exhibit the explicit formulas for our nine normalized error terms, which allows us to prove Theorem 1.4. Using elementary methods combined with Landau's theorem for Mellin transforms, we establish Theorem 1.6 in Section 3. The Fourier transforms of these normalized error terms are examined in Section 4, which leads to the determination of their joint

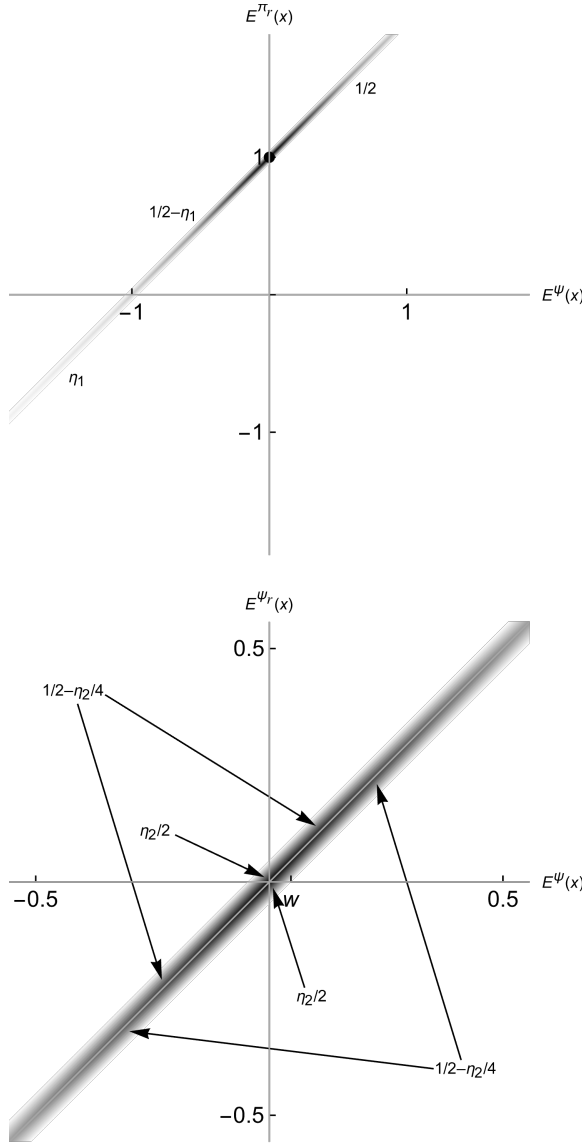



Fig. 1. Two manifestations of $\mu_2$. The top graph shows the joint distribution of $E^{\psi}(x)$ and $E^{\pi_r}(x)$, whose centre is $(\beta_{\psi}, \beta_{\pi_r}) = (0, 1)$; certain regions in the support of $\mu_2 + (0, 1)$ are labeled with their densities from Theorem 1.12(b). The bottom graph shows the joint distribution of $E^{\psi}(x)$ and $E^{\psi_r}(x)$, whose centre is $(\beta_{\psi}, \beta_{\psi_r}) = (0, 0)$; certain regions in the support of $\mu_2$ are labeled with their densities from Theorem 1.14.

limiting logarithmic distribution from which all the theorems in Section 1.2 are then derived. Finally, Sections 5 and 6 contain the numerical calculation of the new constant $\eta_2$, first writing it in terms of the principal value of

a two-variable integral, and then numerically calculating an approximation to that integral with a rigorous error term. Finally, Appendix A contains many graphs like the ones in Figure 1, illustrating all of the possible cases of Theorem 1.17.

**2. Explicit formulas for normalized error terms.** Our first goal is to list explicit formulas for the error terms of the four standard functions and the five reciprocal functions. In this section, $\gamma$ denotes an ordinate of a nontrivial zero of $\zeta(s)$.

PROPOSITION 2.1. *Assume RH. In the notation of Definition* 1.2*, for* $x, X \ge 2$*,*

$$\text{(a)}\quad E^{\psi}(x) = -\Re \sum_{0<\gamma\le X} \frac{2x^{i\gamma}}{\frac12 + i\gamma} + O\bigg(\frac{\sqrt{x}\log^2(xX)}{X} + \frac{\log x}{\sqrt{x}}\bigg),$$

$$\text{(b)}\quad E^{\theta}(x) = -1 - \Re \sum_{0<\gamma\le X} \frac{2x^{i\gamma}}{\frac12 + i\gamma} + O\bigg(\frac{\sqrt{x}\log^2(xX)}{X} + \frac{1}{x^{1/6}}\bigg),$$

$$\text{(c)}\quad E^{\Pi}(x) = -\Re \sum_{0<\gamma\le X} \frac{2x^{i\gamma}}{\frac12 + i\gamma} + O\bigg(\frac{\sqrt{x}\log^2(xX)}{X} + \frac{1}{\log x}\bigg),$$

$$\text{(d)}\quad E^{\pi}(x) = -1 - \Re \sum_{0<\gamma\le X} \frac{2x^{i\gamma}}{\frac12 + i\gamma} + O\bigg(\frac{\sqrt{x}\log^2(xX)}{X} + \frac{1}{\log x}\bigg).$$

*Proof.* (a) This formula is equivalent to [2, (1.6)].

(b) The standard comparison

$$\psi(x) = \sum_{k=1}^{\infty} \theta(x^{1/k}) = \theta(x) + \theta(\sqrt{x}) + O(x^{1/3}) \tag{2.1}$$
$$\implies \theta(x) = \psi(x) - \psi(\sqrt{x}) + O(x^{1/3})$$

shows that this formula follows from part (a).

(c) Partial summation gives us

$$\Pi(x) - \operatorname{li}(x) = \int_{2-}^{x} \frac{1}{\log t}\, d(\psi(t) - t) = \frac{\psi(t)-t}{\log t}\bigg|_{2-}^{x} + \int_{2}^{x} \frac{\psi(t)-t}{t\log^2 t}\, dt$$
$$= \frac{\psi(x)-x}{\log x} + O\bigg(1 + \bigg|\int_{2}^{x} \frac{\psi(t)-t}{t\log^2 t}\, dt\bigg|\bigg),$$

which, upon dividing both sides by $\sqrt{x}/\log x$ becomes

$$E^{\Pi}(x) = E^{\psi}(x) + O\bigg(\frac{\log x}{\sqrt{x}}\bigg|\int_{2}^{x} \frac{E^{\psi}(t)}{\sqrt{t}\log^2 t}\, dt\bigg|\bigg).$$

Using the formula from part (a) in the integrand, we see that

$$E^{\Pi}(x) - E^{\psi}(x) \ll \frac{\log x}{\sqrt{x}} \left| 2\Re \int_2^x \frac{1}{\sqrt{t}\log^2 t} \sum_{0<\gamma\le X} \frac{t^{i\gamma}}{\frac{1}{2}+i\gamma}\, dt \right| + \frac{\log x}{\sqrt{x}} \int_2^x \frac{1}{\sqrt{t}\log^2 t} \left( \frac{\sqrt{t}\log^2(tX)}{X} + \frac{\log t}{\sqrt{t}} \right) dt \tag{2.2}$$

$$\ll \frac{\log x}{\sqrt{x}} \left| \int_2^x \frac{g_X(t)}{\log^2 t}\, dt \right| + \frac{\log x}{\sqrt{x}} \int_2^x \frac{1}{\sqrt{t}\log^2 t} \left( \frac{\sqrt{t}\log^2(tX)}{X} + \frac{\log t}{\sqrt{t}} \right) dt$$

where we have defined

$$g_X(t) = \sum_{0<\gamma\le X} \frac{t^{-1/2+i\gamma}}{\frac{1}{2}+i\gamma}.$$

If we set $G_X(x) = \int_2^x g_X(t)\, dt$, then

$$G_X(x) = \sum_{0<\gamma\le X} \frac{t^{1/2+i\gamma}}{(\frac{1}{2}+i\gamma)^2}\Big|_2^x \ll \sqrt{x},$$

from which it follows by another partial summation that

$$\int_2^x \frac{g_X(t)}{\log^2 t}\, dt = \frac{G_X(t)}{\log^2 t}\Big|_2^x + 2\int_2^x \frac{G_X(t)}{t\log^3 t}\, dt \ll \frac{\sqrt{x}}{\log^2 x} + \int_2^x \frac{1}{\sqrt{t}\log^3 t}\, dt \ll \frac{\sqrt{x}}{\log^2 x} + \frac{\sqrt{x}}{\log^3 x}.$$

Therefore, the first term on the right-hand side of (2.2) is $\ll 1/\log x$; it remains only to estimate the second term on the right-hand side by

$$\frac{\log x}{\sqrt{x}} \int_2^x \frac{1}{\sqrt{t}\log^2 t} \left( \frac{\sqrt{t}\log^2(tX)}{X} + \frac{\log t}{\sqrt{t}} \right) dt \ll \frac{\log x}{X\sqrt{x}} \int_2^x \left( 1 + \frac{\log^2 X}{\log^2 t} \right) dt + \frac{\log x}{\sqrt{x}} \int_2^x \frac{1}{t\log t}\, dt \ll \frac{\sqrt{x}\log x}{X} + \frac{\sqrt{x}\cdot\log^2 X}{X\log x} + \frac{\log x \log\log x}{\sqrt{x}}$$

which is an acceptable error term.

(d) Again, the standard comparison

$$\Pi(x) = \sum_{k=1}^{\infty} \frac{\pi(x^{1/k})}{k} = \pi(x) + \frac{\pi(\sqrt{x})}{2} + O\left(\frac{x^{1/3}}{\log x}\right) \tag{2.3}$$
$$\implies \pi(x) = \Pi(x) - \frac{\Pi(\sqrt{x})}{2} + O\left(\frac{x^{1/3}}{\log x}\right)$$

shows that this formula follows from part (c). ∎

Our goal for the rest of this section is to exhibit analogous explicit formulas for the five reciprocal functions.

LEMMA 2.2. *As a sum over all nontrivial zeros of* $\zeta(s)$*,*

$$\sum_{\rho} \frac{1}{\rho(1-\rho)} = \sum_{\rho} \left(\frac{1}{\rho} + \frac{1}{1-\rho}\right) = 2 + C_0 - \log 4\pi = w$$

*in the notation of equation* (1.1)*.*

*Proof.* This formula is a restatement of [12, (10.30)]. ∎

PROPOSITION 2.3. *For all* $x, T \geq 2$*,*

$$\psi_r(x) = -\sum_{|\gamma| \leq T} \frac{x^{\rho-1}}{\rho - 1} + \log x - C_0 + O\left(\frac{\log x}{x} + \frac{(\log x)^2}{T} + \frac{(\log T)^2}{T \log x}\right).$$

*Proof.* In [8, Lemma 2.4], the author proves that for any $\alpha > 1$,

$$\sum_{n \leq x} \frac{\Lambda(n)}{n^{\alpha}} = -\frac{\zeta'}{\zeta}(\alpha) + \frac{x^{1-\alpha}}{1-\alpha} - \sum_{|\Im\rho| \leq T} \frac{x^{\rho-\alpha}}{\rho - \alpha} + O\big(F(x; \alpha, T)\big) \tag{2.4}$$

where

$$F(x; \alpha, T) = \frac{\log x}{x^{\alpha}} + \frac{x^{1-\alpha}}{T}\left(4^{\alpha} + (\log x)^2 + \frac{(\log T)^2}{\log x}\right) + \frac{1}{T}\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\alpha + 1/\log x}}.$$

We immediately see that

$$\begin{aligned}\lim_{\alpha \to 1^+} F(x; \alpha, T) &= F(x; 1, T) \\ &= \frac{\log x}{x} + \frac{1}{T}\left(4 + (\log x)^2 + \frac{(\log T)^2}{\log x}\right) - \frac{1}{T}\,\frac{\zeta'}{\zeta}\left(1 + \frac{1}{\log x}\right) \\ &\ll \frac{\log x}{x} + \frac{(\log x)^2}{T} + \frac{(\log T)^2}{T \log x} + \frac{\log x}{T}.\end{aligned}$$

We have the formula [12, Corollary 10.14]

$$-\frac{\zeta'}{\zeta}(\alpha) = \frac{C_0}{2} + 1 - \log 2\pi - \frac{1}{1-\alpha} + \frac{1}{2}\,\frac{\Gamma'}{\Gamma}\left(\frac{\alpha}{2} + 1\right) - \sum_{\rho} \left(\frac{1}{\alpha - \rho} + \frac{1}{\rho}\right),$$

as well as the Laurent expansion around $\alpha = 1$ that gives

$$\frac{x^{1-\alpha}}{1-\alpha} = \frac{1}{1-\alpha} + \log x + O_x\big(|1-\alpha|\big).$$

Adding these two equations and taking the limit yields

$$\begin{aligned}\lim_{\alpha\to 1^+}\left(-\frac{\zeta'}{\zeta}(\alpha) + \frac{x^{1-\alpha}}{1-\alpha}\right) &= \frac{C_0}{2} + 1 - \log 2\pi + \log x \\ &\quad + \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{3}{2}\right) - \sum_{\rho}\left(\frac{1}{1-\rho} + \frac{1}{\rho}\right) \\ &= \log x - C_0\end{aligned}$$

by Lemma 2.2 and the special value $\frac{\Gamma'}{\Gamma}\big(\frac{3}{2}\big) = -C_0 - 2\log 2 + 2$. With these evaluations, we can now take the limit of both sides of equation (2.4) to find that

$$\begin{aligned}\psi_r(x) &= \lim_{\alpha\to 1^+}\sum_{n\le x}\frac{\Lambda(n)}{n^\alpha} \\ &= \log x - C_0 - \sum_{|\gamma|\le T}\frac{x^{\rho-1}}{\rho-1} + O\left(\frac{\log x}{x} + \frac{(\log x)^2}{T} + \frac{(\log T)^2}{T\log x}\right)\end{aligned}$$

as claimed. ∎

Proposition 2.4. *Assume RH. In the notation of Definition* 1.2*, for* $x, X \ge 2$*,*

(a) $E^{\psi_r}(x) = -\Re\sum_{0<\gamma\le X}\dfrac{2x^{i\gamma}}{-\frac{1}{2}+i\gamma} + O\Big(\dfrac{\log x}{\sqrt{x}} + \dfrac{\sqrt{x}(\log x)^2}{X} + \dfrac{\sqrt{x}(\log X)^2}{X\log x}\Big)$,

(b) $E^{\pi_\ell}(x) = 1 - \Re\sum_{0<\gamma\le X}\dfrac{2x^{i\gamma}}{-\frac{1}{2}+i\gamma} + O\Big(\dfrac{\sqrt{x}\log^2(xX)}{X} + \dfrac{1}{\log x}\Big)$,

(c) $E^{\Pi_r}(x) = -\Re\sum_{0<\gamma\le X}\dfrac{2x^{i\gamma}}{-\frac{1}{2}+i\gamma} + O\Big(\dfrac{\sqrt{x}\log^2(xX)}{X} + \dfrac{1}{\log x}\Big)$,

(d) $E^{\pi_r}(x) = 1 - \Re\sum_{0<\gamma\le X}\dfrac{2x^{i\gamma}}{-\frac{1}{2}+i\gamma} + O\Big(\dfrac{\sqrt{x}\log^2(xX)}{X} + \dfrac{1}{\log x}\Big)$,

(e) $E^{\theta_r}(x) = 1 - \Re\sum_{0<\gamma\le X}\dfrac{2x^{i\gamma}}{-\frac{1}{2}+i\gamma} + O\Big(\dfrac{\sqrt{x}\log^2(xX)}{X} + \dfrac{1}{\log x}\Big)$.

*Proof.* (a) Assuming RH, we can write

$$\sum_{1<|\Im(\rho)|\le T}\frac{x^{\rho-1}}{\rho-1} = \sum_{1<|\gamma|\le T}\frac{x^{-\frac{1}{2}+i\gamma}}{-\frac{1}{2}+i\gamma}.$$

Then the formula for $E^{\psi_r}(x)$ is a straightforward consequence of Proposition 2.3.

(b) This formula is [8, Corollary 2.2] or [9, Proposition 1], although both those statements contain a sign mistake (confirmed by the author of [8]) that is corrected here.

(c) This formula follows from part (b) and [8, Lemma 2.3], which in our notation is

$$\pi_\ell(x) = \Pi_r(x) + \frac{1}{\sqrt{x}\log x} + O\left(\frac{1}{\sqrt{x}(\log x)^2}\right).$$

(d) This formula follows from part (b) and [9, Lemma 5], which (when both sides are multiplied by $\sqrt{x}\log x$) states in our notation that

$$E^{\pi_r}(x) = E^{\pi_\ell}(x) + O\bigg(\frac{\log x}{\sqrt{x}}\bigg).$$

(e) This formula is [9, Lemma 6]. ∎

These explicit formulas enable us to establish the ranges of values taken by differences of the normalized error terms from Definition 1.2.

*Proof of Theorem 1.4.* Part (a) follows immediately from Propositions 2.1 and 2.4, in light of the bias terms from Definition 1.3. Also from those propositions (setting $X = x$),

(2.5)
$$\begin{aligned}\big|E^{\psi}(x) - E^{\psi_r}(x)\big| &= \bigg|\Re \sum_{0<\gamma\le x} \frac{2x^{i\gamma}}{\frac12 + i\gamma} - \Re \sum_{0<\gamma\le x} \frac{2x^{i\gamma}}{-\frac12 + i\gamma}\bigg| + O\bigg(\frac{\log^2 x}{x}\bigg)\\ &= \bigg|2\Re \sum_{0<\gamma\le x} \frac{x^{i\gamma}}{\frac14 + \gamma^2}\bigg| + o(1) < 2\sum_{\gamma>0} \frac{1}{\frac14+\gamma^2} + o(1)\\ &= \sum_\rho \frac{1}{\rho(1-\rho)} + o(1) = w + o(1)\end{aligned}$$

by Lemma 2.2. This inequality implies the special case of part (b) where $f = \psi_r$ and $g = \psi$, and the general statement follows from this special case and part (a). ∎

**3. Persistent inequalities between normalized error terms.** We begin this section by establishing some unconditional relationships between pairs of normalized error terms, after which we write down the Mellin transforms of those functions for use in Landau's theorem-type arguments. We prove Theorem 1.6(a) at the end of Section 3.1 and Theorem 1.6(b) at the end of Section 3.3.

**3.1. Unconditional relationships.** We first establish some relationships among the error terms from Definition 1.2 that are independent of any assumptions on the zeros of $\zeta(s)$; Theorem 1.6(a) will be an immediate consequence of these formulas.

LEMMA 3.1. *Unconditionally, there exists a positive constant $c$ such that*

$$E^{\psi}(x) - E^{\theta}(x) = 1 + O\big(\exp(-c\sqrt{\log x})\big),$$
$$E^{\Pi}(x) - E^{\pi}(x) = 1 + O\big(\exp(-c\sqrt{\log x})\big).$$

*Proof.* As $\psi(x) = \sum_{k=1}^{\infty} \theta(x^{1/k})$, the standard argument gives

$$\begin{aligned} \psi(x) - \theta(x) &= \sum_{k=2}^{\infty} \theta(x^{1/k}) \\ &= \theta(x^{1/2}) + \theta(x^{1/3}) + \sum_{4 \le k \ll \log x} \theta(x^{1/k}) \\ &= \theta(x^{1/2}) + O(x^{1/3}). \end{aligned} \tag{3.1}$$

By the prime number theorem, there exists $c_0 > 0$ such that

$$\begin{aligned} \psi(x) - \theta(x) &= x^{1/2} + O\big(x^{1/2}\exp(-c_0\sqrt{\log x^{1/2}})\big) + O(x^{1/3}) \\ &= x^{1/2} + O\big(x^{1/2}\exp(-c\sqrt{\log x})\big) \end{aligned}$$

with $c = \frac{c_0}{2}$. The first claim follows upon dividing both sides by $x^{1/2}$.

Similarly, the identity $\Pi(x) = \sum_{k=1}^{\infty} \frac{1}{k}\pi(x^{1/k})$ yields

$$\begin{aligned} \Pi(x) - \pi(x) &= \sum_{k=2}^{\infty} \frac{\pi(x^{1/k})}{k} \\ &= \frac{\pi(x^{1/2})}{2} + \frac{\pi(x^{1/3})}{3} + \sum_{4 \le k \ll \log x} \frac{\pi(x^{1/k})}{k} \\ &= \frac{\pi(x^{1/2})}{2} + O\bigg(\frac{x^{1/3}}{\log x}\bigg). \end{aligned} \tag{3.2}$$

The second claim follows by invoking the prime number theorem for $\pi(x^{1/2})$ and dividing both sides by $x^{1/2}/\log x$. ∎

LEMMA 3.2. *Unconditionally, $E^{\psi_r}(x) - E^{\theta_r}(x) = -1 + O(1/\log x)$.*

*Proof.* By Definition 1.1, it suffices to prove $\sum_p \frac{\log p}{p(p-1)} + \theta_r(x) - \psi_r(x) = x^{-1/2}(1 + O(1/\log x))$, or equivalently

$$\sum_p \sum_{k=2}^{\infty} \frac{\log p}{p^k} + \sum_{p \le x} \frac{\log p}{p} - \sum_{k=1}^{\infty} \sum_{p \le x^{1/k}} \frac{\log p}{p^k} = x^{-1/2}(1 + O(1/\log x)). \tag{3.3}$$

The $k = 1$ term in the third sum precisely cancels the second sum. If we define $\Delta^{\theta_r}(x) = \theta_r(x) - (\log x - C_1)$, so that $\Delta^{\theta_r}(x) \ll 1/\log x$ by Mertens's formula [5, line 3], then the contribution to the left-hand side of $k = 2$ equals

$$\begin{aligned}\sum_{p>\sqrt{x}} \frac{\log p}{p^2} &= \int_{\sqrt{x}}^{\infty} \frac{1}{u}\, d\theta_r(u) = \int_{\sqrt{x}}^{\infty} \frac{1}{u}\, d\big(\log u - C_1 + \Delta^{\theta_r}(u)\big)\\ &= \int_{\sqrt{x}}^{\infty} \frac{1}{u^2}\, du + \frac{\Delta^{\theta_r}(u)}{u}\bigg|_{\sqrt{x}}^{\infty} + \int_{\sqrt{x}}^{\infty} \frac{\Delta^{\theta_r}(u)}{u^2}\, du\\ &= -\frac{1}{u}\bigg|_{\sqrt{x}}^{\infty} + O\bigg(\frac{1/\log u}{u}\bigg|_{\sqrt{x}}^{\infty} + \int_{\sqrt{x}}^{\infty} \frac{1/\log u}{u^2}\, du\bigg)\\ &= \frac{1}{\sqrt{x}} + O\bigg(\frac{1}{\sqrt{x}\log\sqrt{x}}\bigg).\end{aligned}$$

Finally, the contribution to the left-hand side of (3.3) from $k \ge 3$ is, by Rankin's trick,

$$\begin{aligned}\sum_{k=3}^{\infty} \sum_{p>x^{1/k}} \frac{\log p}{p^k} &\le \sum_p \sum_{k\ge 3} \frac{\log p}{p^k} \bigg(\frac{p^k}{x}\bigg)^{3/5} = x^{-3/5} \sum_p \sum_{k\ge 3} \frac{\log p}{p^{2k/5}}\\ &= x^{-3/5} \sum_p \frac{\log p}{p^{6/5} - p^{4/5}} \ll x^{-3/5} \ll \frac{1}{\sqrt{x}\log x},\end{aligned}$$

which concludes the proof of the lemma. ∎

Lemma 3.3. *Unconditionally,*

$$\begin{aligned}E^{\pi_\ell}(x) - E^{\Pi_r}(x) &= 1 + O(1/\log x),\\ E^{\pi_\ell}(x) - E^{\pi_r}(x) &\ll x^{-1/2}\log x,\\ E^{\Pi_r}(x) - E^{\pi_r}(x) &= -1 + O(1/\log x).\end{aligned}$$

*Furthermore, $E^{\pi_\ell}(x) < E^{\pi_r}(x)$ for $x > 1$.*

*Proof.* The first assertion is precisely [8, Lemma 2.3] translated into our notation. Similarly, the second assertion is precisely [9, Lemma 5]; then the third assertion follows from the other two. The final assertion follows from

$$\begin{aligned}\frac{E^{\pi_\ell}(x) - E^{\pi_r}(x)}{\sqrt{x}\log x} &= \big(\pi_\ell(x) - (\log\log x + C_0)\big) - \big(\pi_r(x) - (\log\log x - C_2)\big)\\ &= \pi_\ell(x) - \pi_r(x) - (C_0 + C_2)\\ &= \sum_{p\le x}\sum_{k=2}^{\infty} \frac{1}{kp^k} - \sum_{k=2}^{\infty}\sum_p \frac{1}{kp^k} = -\sum_{k=2}^{\infty}\sum_{p>x} \frac{1}{kp^k} < 0.\end{aligned}$$

∎

*Proof of Theorem 1.6(a).* For the pairs of functions $(f,g) \in \{(\theta,\psi),(\pi,\Pi),(\psi_r,\theta_r),(\Pi_r,\pi_r),(\Pi_r,\pi_\ell)\}$, Lemmas 3.1–3.3 show that $E^g(x) - E^f(x) = 1+o(1)$, which immediately implies that $E^f(x) < E^g(x)$ when $x$ is sufficiently large. ∎

**3.2. Mellin transforms: preliminaries.** A familiar tool in analytic number theory is the Mellin transform $\int_2^\infty f(x)x^{-s-1}\,dx$ of a function $f(x)$, where the integral typically converges in some right half-plane. A famous theorem of Landau (see for example [12, Lemma 15.1]) allows us to connect oscillations of $f(x)$ to the singularities of its Mellin transform. In preparation for carrying out this strategy in Section 3.3 to facilitate the proof of Theorem 1.6(b), we establish some preliminary integral formulas in this section. We consistently choose 2 as the lower endpoint of our integrals to avoid irrelevant technicalities when factors of $\frac{1}{\log x}$ or $\log\log x$ are present.

LEMMA 3.4. *For any complex number $s$ with $\Re s > 1$,*

$$\int_2^\infty x^{-s}\log x\,dx = \frac{2^{1-s}}{(1-s)^2} - \frac{2^{1-s}\log 2}{1-s}.$$

*Proof.* The identity follows from the substitution $v = \log x$ and integrating by parts:

$$\int_2^\infty x^{-s}\log x\,dx = \int_{\log 2}^\infty ve^{(1-s)v}\,dv = v\frac{e^{(1-s)v}}{1-s}\Big|_{\log 2}^\infty - \frac{1}{1-s}\int_{\log 2}^\infty e^{(1-s)v}\,dv$$
$$= -(\log 2)\frac{2^{1-s}}{1-s} + \frac{2^{1-s}}{(1-s)^2}.$$ ∎

LEMMA 3.5. *For any complex number $s$ with $\Re s > 1$,*

$$\int_2^\infty \frac{x^{-s}}{\log x}\,dx = -\log(s-1) + r_0(s)$$

*where $r_0(s)$ is an entire function with $r_0(1) = -C_0 - \log\log 2$.*

REMARK 3.6. The proof is valid when $\Re s > 1$; what we mean by "$r_0(s)$ is an entire function" is that the function $r_0(s)$ implicitly defined by the formula has an analytic continuation to all of $\mathbb{C}$. The analogous interpretation holds for similar statements in the rest of this section and the next.

*Proof of Lemma 3.5.* Using the substitution $u = (s-1)\log x$, we easily verify that

$$\int_2^\infty \frac{x^{-s}}{\log x}\,dx = \int_{(s-1)\log 2}^\infty \frac{e^{-u}}{u}\,du$$
$$= \int_1^\infty \frac{e^{-u}}{u}\,du + \int_0^1 \frac{e^{-u}-1}{u}\,du$$

$$-\int_0^{(s-1)\log 2}\frac{e^{-u}-1}{u}\,du+\int_{(s-1)\log 2}^{1}\frac{du}{u}$$
$$=-C_0-\int_0^{(s-1)\log 2}\frac{e^{-u}-1}{u}\,du-\log((s-1)\log 2),$$

where the evaluation of the first two integrals to $-C_0$ is [12, equation (7.31)]. This calculation establishes the lemma with

$$r_0(s)=-C_0-\int_0^{(s-1)\log 2}\frac{e^{-u}-1}{u}\,du-\log\log 2. \ \blacksquare$$

LEMMA 3.7. *For any complex number $s$ with $\Re s>1$,*

$$\int_2^\infty \mathrm{li}(x)x^{-s-1}\,dx=-\frac{1}{s}\log(s-1)+r_1(s),$$
$$\int_2^\infty \mathrm{li}(x)x^{-s-1}\log x\,dx=\frac{1}{s(s-1)}-\frac{1}{s^2}\log(s-1)-r_{\Pi}(s),$$

*where $r_1(s)$ and $r_{\Pi}(s)$ are analytic for $\Re s>0$.*

*Proof.* Integration by parts gives

$$\int_2^\infty \mathrm{li}(x)x^{-s-1}\,dx=\frac{\mathrm{li}(2)}{s2^s}+\frac{1}{s}\int_2^\infty\frac{x^{-s}}{\log x}\,dx=\frac{\mathrm{li}(2)}{s2^s}+\frac{-\log(s-1)+r_0(s)}{s}$$

by Lemma 3.5, which establishes the first identity with

$$r_1(s)=\frac{\mathrm{li}(2)}{s2^s}+\frac{r_0(s)}{s}.$$

The second formula follows by differentiating both sides of the first equation with respect to $s$, negating both sides, and setting $r_{\Pi}(s)=r_1'(s)$. $\blacksquare$

REMARK 3.8. Our proof of Lemma 3.7 is essentially the same as the one in [12, Theorem 15.2]. However, those authors use the convention that $\mathrm{li}(x)=\int_2^\infty\frac{du}{\log u}$ rather than our convention that $\mathrm{li}(x)=\int_0^\infty\frac{du}{\log u}$; thus we have chosen to include a complete proof (that yields an additional boundary term) to avoid confusion.

LEMMA 3.9. *For any complex number $s$ with $\Re s>1$,*

$$\int_2^\infty x^{-s}\log\log x\,dx=\frac{1}{1-s}(\log(s-1)+C_0)+r_2(s),$$
$$\int_2^\infty x^{-s}\log x\log\log x\,dx=-\frac{\log(s-1)+C_0-1}{(s-1)^2}+r_3(s),$$

*where $r_2(s)$ and $r_3(s)$ are entire.*

*Proof.* Integration by parts gives

$$\begin{aligned}\int_2^\infty x^{-s}\log\log x\,dx &= -\frac{2^{1-s}}{1-s}\log\log 2 - \frac{1}{1-s}\int_2^\infty \frac{x^{1-s}}{x\log x}\,dx\\ &= \frac{2^{1-s}}{s-1}\log\log 2 + \frac{-\log(s-1)+r_0(s)}{s-1}\\ &= \frac{2^{1-s}-1}{s-1}\log\log 2 + \frac{-\log(s-1)-C_0}{s-1} + \frac{r_0(s)-r_0(1)}{s-1}\end{aligned}$$

by Lemma 3.5, which establishes the first formula with

$$r_2(s) = \frac{2^{1-s}-1}{s-1}\log\log 2 + \frac{r_0(s)-r_0(1)}{s-1}.$$

The second formula follows by differentiating both sides of the first equation with respect to $s$, negating both sides, and setting $r_3(s) = -r_2'(s)$. ∎

**3.3. Mellin transforms of number-theoretic functions.** In this section, we determine the Mellin transforms of all of the functions from Definition 1.1 and, more importantly, of all of the normalized error terms from Definition 1.2. These Mellin transforms are needed for the proof of Theorem 1.6(b) at the end of this section. One detail of note is that all of the upcoming Mellin transforms of normalized error terms have removable singularities at $s = 1$; this is equivalent in essence to the fact that the main terms in Definition 1.2 have been chosen appropriately.

The usual Mellin transform identity (see [12, Theorem 1.3]), applied to any sequence $(a_n)$ of complex numbers with $a_1 = 0$, is

$$\int_2^\infty \Big(\sum_{n\le x} a_n\Big) x^{-s-1}\,dx = \frac{1}{s}\sum_{n=1}^\infty \frac{a_n}{n^s} \tag{3.4}$$

whenever $\Re s$ is positive and larger than the abscissa of convergence of the Dirichlet series on the right-hand side. From this formula, it is easy to derive that

$$\int_2^\infty \Big(\sum_{n\le x} \frac{a_n}{n}\Big) x^{-s}\,dx = \frac{1}{s-1}\sum_{n=1}^\infty \frac{a_n}{n^s}. \tag{3.5}$$

LEMMA 3.10. *For any complex number $s$ with $\Re s > 1$,*

$$\int_2^\infty \psi(x)x^{-s-1}\,dx = -\frac{1}{s}\frac{\zeta'}{\zeta}(s),$$

$$\int_2^\infty E^\psi(x)x^{-s-1/2}\,dx = -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{2^{1-s}}{1-s}.$$

*Proof.* The first formula follows immediately from (3.4) with $a_n = \Lambda(n)$. We then calculate

$$\int_2^\infty E^\psi(x)x^{-s-1/2}\,dx = \int_2^\infty \frac{\psi(x)-x}{x^{1/2}}x^{-s-1/2}\,dx \tag{3.6}$$
$$= \int_2^\infty \psi(x)x^{-s-1}\,dx - \int_2^\infty x^{-s}\,dx = -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{2^{1-s}}{1-s}. \ \blacksquare$$

LEMMA 3.11. *For any complex number $s$ with $\Re s > 1$,*

$$\int_2^\infty \theta(x)x^{-s-1}\,dx = -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s}\frac{\zeta'}{\zeta}(2s) + r_\theta(s),$$
$$\int_2^\infty E^\theta(x)x^{-s-1/2}\,dx = -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s}\frac{\zeta'}{\zeta}(2s) + \frac{2^{1-s}}{1-s} + r_\theta(s)$$

*where $r_\theta(s)$ is analytic for $\Re s > \frac{1}{3}$.*

*Proof.* We write

$$\int_2^\infty \theta(x)x^{-s-1}\,dx = \int_2^\infty \psi(x)x^{-s-1}\,dx - \int_2^\infty \psi(\sqrt{x})x^{-s-1}\,dx + r_\theta(s) \tag{3.7}$$

where, by (2.1),

$$r_\theta(s) = \int_2^\infty \big(\theta(x) - \psi(x) + \psi(\sqrt{x})\big)x^{-s-1}\,dx \ll \int_2^\infty x^{1/3}x^{-\sigma-1}\,dx$$

converges to an analytic function for all $\Re s > \frac{1}{3}$. The first integral on the right-hand side of (3.7) is evaluated in Lemma 3.10, and the second integral also follows from Lemma 3.10 since

$$\int_2^\infty \psi(\sqrt{x})x^{-s-1}\,dx = \int_{\sqrt{2}}^\infty \psi(t)(t^2)^{-s-1}\cdot 2t\,dt = 2\int_2^\infty \psi(t)t^{-(2s)-1}\,dt \tag{3.8}$$

(again we may change the lower endpoint from $\sqrt{2}$ to 2 since $\psi(t) = 0$ for intermediate values). The second formula follows from the first by the same calculation as in (3.6). $\blacksquare$

LEMMA 3.12. *For any complex number $s$ with $\Re s > 1$,*

$$\int_2^\infty \Pi(x)x^{-s-1}\,dx = \frac{1}{s}\log\zeta(s),$$
$$\int_2^\infty \Pi(x)x^{-s-1}\log x\,dx = -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s^2}\log\zeta(s),$$

$$\int_2^\infty E^{\Pi}(x)x^{-s-1/2}\,dx = -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s^2}\log\big(\zeta(s)(s-1)\big) - \frac{1}{s(s-1)} + r_{\Pi}(s),$$

*where $r_{\Pi}(s)$ is analytic for $\Re s > 0$.*

*Proof.* Since $\Pi(x) = \sum_{n\le x}\frac{\Lambda(n)}{\log n}$ and $\log\zeta(s) = \sum_{n=1}^{\infty}\frac{\Lambda(n)}{\log n}n^{-s}$ for $\Re s > 1$, the first formula follows from (3.4). The second formula follows from the first by differentiating both sides with respect to $s$ and negating both sides. Then by Lemma 3.7 combined with the second formula,

$$\begin{aligned}(3.9)\quad \int_2^\infty E^{\Pi}(x)x^{-s-1/2}\,dx &= \int_2^\infty \frac{\Pi(x)-\mathrm{li}(x)}{x^{1/2}/\log x}x^{-s-1/2}\,dx\\ &= \int_2^\infty \big(\Pi(x)x^{-s-1}\log x - \mathrm{li}(x)x^{-s-1}\log x\big)\,dx\\ &= -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s^2}\log\zeta(s) - \left(\frac{1}{s(s-1)} - \frac{1}{s^2}\log(s-1) - r_{\Pi}(s)\right)\\ &= -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s^2}\log\big(\zeta(s)(s-1)\big) - \frac{1}{s(s-1)} + r_{\Pi}(s). \ \blacksquare\end{aligned}$$

LEMMA 3.13. *For any complex number $s$ with $\Re s > 1$,*

$$\begin{aligned}\int_2^\infty \pi(x)x^{-s-1}\,dx &= \frac{1}{s}\log\zeta(s) - \frac{1}{2s}\log\zeta(2s) + r_4(s),\\ \int_2^\infty \pi(x)x^{-s-1}\log x\,dx &= -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s^2}\log\zeta(s)\\ &\quad + \frac{1}{s}\frac{\zeta'}{\zeta}(2s) - \frac{1}{2s^2}\log\zeta(2s) + r_4'(s),\\ \int_2^\infty E^{\pi}(x)x^{-s-1/2}\,dx &= -\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s^2}\log\big(\zeta(s)(s-1)\big)\\ &\quad + \frac{1}{s}\frac{\zeta'}{\zeta}(2s) - \frac{1}{2s^2}\log\zeta(2s) - \frac{1}{s(s-1)} + r_{\pi}(s),\end{aligned}$$

*where $r_4(s)$ and $r_{\pi}(s)$ are analytic for $\Re s > \frac{1}{3}$.*

*Proof.* We write

$$(3.10)\quad \int_2^\infty \pi(x)x^{-s-1}\,dx = \int_2^\infty \Pi(x)x^{-s-1}\,dx - \int_2^\infty \tfrac{1}{2}\Pi(\sqrt{x})x^{-s-1}\,dx + r_4(s)$$

where, by (2.3),

$$r_4(s) = \int_2^\infty \big(\pi(x) - \Pi(x) + \tfrac{1}{2}\Pi(\sqrt{x})\big)x^{-s-1}\,dx \ll \int_2^\infty \frac{x^{1/3}}{\log x}x^{-\sigma-1}\,dx$$

converges to an analytic function for all $\Re s > \frac{1}{3}$. The first integral on the right-hand side of (3.10) is evaluated in Lemma 3.12, and the second integral also follows from Lemma 3.10 by a calculation similar to (3.8). The second formula follows from the first by differentiating both sides with respect to $s$ and negating both sides. Finally, the third formula follows from the second by the same calculation as in (3.9), upon setting $r_\pi(s) = r_4'(s) - r_\Pi(s)$. ∎

LEMMA 3.14. *For any complex number $s$ with $\Re s > 1$,*

$$\int_2^\infty \psi_r(x)x^{-s}\,dx = \frac{1}{1-s}\frac{\zeta'}{\zeta}(s),$$

$$\int_2^\infty E^{\psi_r}(x)x^{-s-1/2}\,dx = \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{2^{1-s}}{(1-s)^2} + \frac{2^{1-s}(\log 2 - C_0)}{1-s}.$$

*Proof.* The first formula follows immediately from (3.5) with $a_n = \Lambda(n)$. Then

$$\begin{aligned}\int_2^\infty E^{\psi_r}(x)x^{-s-1/2}\,dx &= \int_2^\infty (\psi_r(x) - \log x + C_0)\sqrt{x}\cdot x^{-s-1/2}\,dx\\ &= \int_2^\infty \psi_r(x)x^{-s}\,dx - \int_2^\infty x^{-s}\log x\,dx + \int_2^\infty C_0 x^{-s}\,dx\\ &= \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \left(\frac{2^{1-s}}{(1-s)^2} - \frac{2^{1-s}\log 2}{1-s}\right) - \frac{C_0 2^{1-s}}{1-s}\end{aligned}$$

by Lemma 3.4. ∎

LEMMA 3.15. *For any complex number $s$ with $\Re s > 1$,*

$$\int_2^\infty \theta_r(x)x^{-s}\,dx = \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{1}{1-s}\frac{\zeta'}{\zeta}(2s) - \frac{r_5(s)}{1-s}, \tag{3.11}$$

$$\begin{aligned}\int_2^\infty E^{\theta_r}(x)x^{-s-1/2}\,dx &= \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{1}{1-s}\left(\frac{\zeta'}{\zeta}(2s) - \frac{\zeta'}{\zeta}(2)\right)\\ &\quad - \frac{2^{1-s}}{(1-s)^2} + \frac{2^{1-s}(\log 2 - C_0)}{1-s} + r_{\theta_r}(s),\end{aligned}$$

*where $r_5(s)$ and $r_{\theta_r}(s)$ are analytic for $\Re s > \frac{1}{3}$.*

*Proof.* We begin by defining

$$\begin{aligned}r_5(s) &= \sum_p \frac{\log p}{p^s} - \left(-\frac{\zeta'}{\zeta}(s) + \frac{\zeta'}{\zeta}(2s)\right)\\ &= \sum_p \frac{\log p}{p^s} - \sum_{p^k} \frac{\log p}{p^{ks}} + \sum_{p^k} \frac{\log p}{p^{2ks}} = -\sum_{\substack{k\ge 3\\ k\text{ odd}}} \sum_p \frac{\log p}{p^{ks}},\end{aligned} \tag{3.12}$$

which converges to an analytic function for $\Re s > \frac{1}{3}$. Equation (3.5) then gives

$$\int_2^\infty \theta_r(x)x^{-s}\,dx = \frac{1}{s-1}\sum_p \frac{\log p}{p^s} = \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{1}{1-s}\frac{\zeta'}{\zeta}(2s) - \frac{r_5(s)}{1-s}. \tag{3.13}$$

By Lemma 3.4 it follows that

$$\begin{aligned}\int_2^\infty E^{\theta_r}(x)x^{-s-1/2}\,dx &= \int_2^\infty \big(\theta_r(x) - (\log x - C_1)\big)\sqrt{x}\cdot x^{-s-1/2}\,dx\\ &= \int_2^\infty \theta_r(x)x^{-s}\,dx - \int_2^\infty x^{-s}\log x\,dx + \int_2^\infty C_1 x^{-s}\,dx\\ &= \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{1}{1-s}\frac{\zeta'}{\zeta}(2s) - \frac{r_5(s)}{1-s} - \left(\frac{2^{1-s}}{(1-s)^2} - \frac{2^{1-s}\log 2}{1-s}\right)\\ &\quad - \frac{C_1 2^{1-s}}{1-s}\\ &= \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{1}{1-s}\frac{\zeta'}{\zeta}(2s) - \frac{r_5(s)}{1-s} - \frac{2^{1-s}}{(1-s)^2} + \frac{2^{1-s}\log 2}{1-s}\\ &\quad - \frac{2^{1-s}}{1-s}\left(C_0 + \sum_p \frac{\log p}{p(p-1)}\right)\end{aligned}$$

by Definition 1.2. If we now define

$$r_{\theta_r}(s) = -\frac{1}{1-s}\left(\frac{\zeta'}{\zeta}(2) + r_5(s) + 2^{1-s}\sum_p \frac{\log p}{p(p-1)}\right),$$

then the second identity in (3.11) holds; it remains to show that $r_{\theta_r}(s)$ is analytic for $\Re s > \frac{1}{3}$, for which it suffices to show that the quantity in parentheses vanishes at $s = 1$. But by the definition (3.12) of $r_5(s)$,

$$\frac{\zeta'}{\zeta}(2s) + r_5(s) = -\sum_{k=2}^\infty \sum_p \frac{\log p}{p^{ks}} = -\sum_p \log p \sum_{k=2}^\infty p^{-ks} = -\sum_p \frac{\log p}{p^s(p^s-1)},$$

and therefore

$$\frac{\zeta'}{\zeta}(2) + r_5(1) + 2^{1-1}\sum_p \frac{\log p}{p(p-1)} = \lim_{s\to 1}\left(\frac{\zeta'}{\zeta}(2s) + r_5(s)\right) + \sum_p \frac{\log p}{p(p-1)} = 0$$

as required. ∎

LEMMA 3.16. *For any complex number $s$ with $\Re s > 1$,*

$$\int_2^\infty \Pi_r(x)x^{-s}\,dx = \frac{1}{s-1}\log\zeta(s),$$

$$\int_2^\infty \Pi_r(x)x^{-s}\log x\,dx = -\frac{1}{s-1}\,\frac{\zeta'}{\zeta}(s) + \frac{1}{(s-1)^2}\log\zeta(s),$$

$$\int_2^\infty E^{\Pi_r}(x)x^{-s-1/2}\,dx = -\frac{1}{s-1}\,\frac{\zeta'}{\zeta}(s) + \frac{\log\big(\zeta(s)(s-1)\big)-1}{(s-1)^2} + r_{\Pi_r}(s)$$

*where $r_{\Pi_r}(s)$ is entire.*

*Proof.* Since

$$\Pi_r(x) = \sum_{n\le x}\frac{\Lambda(n)}{n\log n} \quad\text{and}\quad \log\zeta(s) = \sum_{n=1}^\infty \frac{\Lambda(n)}{\log n}n^{-s} \quad \text{for } \Re s > 1,$$

the first formula follows from equation (3.5). The second formula follows from the first by differentiating both sides with respect to $s$ and negating both sides. Then by Lemmas 3.9 and 3.4 combined with the second formula,

$$\begin{aligned}
&\int_2^\infty E^{\Pi_r}(x)x^{-s-1/2}\,dx\\
&\quad= \int_2^\infty \big(\Pi_r(x) - (\log\log x + C_0)\big)\sqrt{x}\log x\cdot x^{-s-1/2}\,dx\\
&\quad= \int_2^\infty \Pi_r(x)x^{-s}\log x\,dx - \int_2^\infty x^{-s}\log x\log\log x\,dx - \int_2^\infty C_0 x^{-s}\log x\,dx\\
&\quad= -\frac{1}{s-1}\,\frac{\zeta'}{\zeta}(s) + \frac{1}{(s-1)^2}\log\zeta(s)\\
&\qquad - \left(-\frac{\log(s-1)+C_0-1}{(s-1)^2} + r_3(s)\right) - C_0\left(\frac{2^{1-s}}{(1-s)^2} - \frac{2^{1-s}\log 2}{1-s}\right)\\
&\quad= -\frac{1}{s-1}\,\frac{\zeta'}{\zeta}(s) + \frac{\log\big(\zeta(s)(s-1)\big)-1}{(s-1)^2} + r_{\Pi_r}(s),
\end{aligned}$$

where one can verify that

$$r_{\Pi_r}(s) = C_0\left(\frac{1-2^{1-s}}{(1-s)^2} + \frac{2^{1-s}\log 2}{1-s}\right) - r_3(s)$$

has a removable singularity at $s = 1$. ∎

LEMMA 3.17. *For any complex number $s$ with $\Re s > 1$,*

$$\begin{aligned}\int_2^\infty \pi_r(x)x^{-s}\log x\,dx &= \frac{1}{(1-s)^2}\log\zeta(s) + \frac{1}{1-s}\frac{\zeta'}{\zeta}(s)\\ &\quad - \frac{1}{2(1-s)^2}\log\zeta(2s) - \frac{1}{1-s}\frac{\zeta'}{\zeta}(2s)\\ &\quad - \frac{C_0 - \frac{1}{2}\log 2 + C_2 2^{1-s}}{(1-s)^2} + \frac{C_2 2^{1-s}\log 2}{1-s}\\ &\quad + r_{\pi_r}(s) + r_3(s),\end{aligned}$$

(3.14)

$$\begin{aligned}\int_2^\infty E^{\pi_r}(x)x^{-s-1/2}\,dx &= \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) + \frac{\log(\zeta(s)(s-1)) - 1}{(1-s)^2}\\ &\quad - \frac{1}{2(1-s)^2}\log\frac{\zeta(2s)}{\zeta(2)} - \frac{1}{1-s}\frac{\zeta'}{\zeta}(2s) + r_{\pi_r}(s),\end{aligned}$$

*where $r_3(s)$ (from Lemma 3.9) is entire and $r_{\pi_r}(s)$ is analytic for $\Re s > \frac{1}{3}$.*

*Proof.* We begin by defining

$$\begin{aligned}r_6(s) &= \sum_p \frac{1}{p^s} - \left(\log\zeta(s) - \frac{1}{2}\log\zeta(2s)\right)\\ &= \sum_p \frac{1}{p^s} - \sum_{p^k}\frac{1}{kp^{ks}} + \frac{1}{2}\sum_{p^k}\frac{1}{kp^{2ks}}\\ &= -\sum_{\substack{k\ge 3\\ k\text{ odd}}}\sum_p \frac{1}{kp^{ks}},\end{aligned} \tag{3.15}$$

which converges to an analytic function for $\Re s > \frac{1}{3}$. Equation (3.5) then gives

$$\begin{aligned}\int_2^\infty \pi_r(x)x^{-s}\,dx &= \frac{1}{s-1}\sum_p \frac{1}{p^s}\\ &= \frac{1}{s-1}\log\zeta(s) - \frac{1}{2(s-1)}\log\zeta(2s) + \frac{r_6(s)}{s-1},\end{aligned}$$

and differentiating with respect to $s$ and negating both sides yields

$$\begin{aligned}\int_2^\infty \pi_r(x)x^{-s}\log x\,dx &= \frac{1}{(s-1)^2}\log\zeta(s) - \frac{1}{s-1}\frac{\zeta'}{\zeta}(s)\\ &\quad - \frac{1}{2(s-1)^2}\log\zeta(2s) + \frac{1}{s-1}\frac{\zeta'}{\zeta}(2s) + \frac{r_6(s)}{(s-1)^2} - \frac{r_6'(s)}{s-1}.\end{aligned}$$

By Lemmas 3.9 and 3.4, it follows that

$$\begin{aligned}
&\int_2^\infty E^{\pi_r}(x)x^{-s-1/2}\,dx = \int_2^\infty \big(\pi_r(x) - \log\log x + C_2\big)\sqrt{x}\cdot x^{-s-1/2}\,dx \\
&= \int_2^\infty \pi_r(x)x^{-s}\log x\,dx - \int_2^\infty x^{-s}\log x\log\log x\,dx + C_2\int_2^\infty x^{-s}\log x\,dx \\
&= \frac{\log\zeta(s)}{(1-s)^2} + \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{\log\zeta(2s)}{2(1-s)^2} - \frac{1}{1-s}\frac{\zeta'}{\zeta}(2s) + \frac{r_6(s)}{(1-s)^2} + \frac{r_6'(s)}{1-s} \\
&\quad - \left(-\frac{\log(s-1)+C_0-1}{(1-s)^2} + r_3(s)\right) + C_2\left(\frac{2^{1-s}}{(1-s)^2} - \frac{2^{1-s}\log 2}{1-s}\right) \\
&= \frac{\log\big(\zeta(s)(s-1)\big)-1}{(1-s)^2} + \frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{1}{2(1-s)^2}\log\frac{\zeta(2s)}{\zeta(2)} - \frac{1}{1-s}\frac{\zeta'}{\zeta}(2s) \\
&\quad + \frac{r_6(s)}{(1-s)^2} + \frac{r_6'(s)}{1-s} + \frac{C_0 - \frac12\log 2}{(1-s)^2} - r_3(s) + C_2\left(\frac{2^{1-s}}{(1-s)^2} - \frac{2^{1-s}\log 2}{1-s}\right).
\end{aligned}$$

If we now define

$$\begin{aligned}
r_{\pi_r}(s) = \frac{1}{(1-s)^2}\Big(&r_6(s) + C_0 - \frac12\log 2 + C_2 2^{1-s} \\
&+ (1-s)\big(r_6'(s) - C_2 2^{1-s}\log 2\big)\Big) - r_3(s),
\end{aligned}$$

then the identity (3.14) holds; it remains to show that $r_{\pi_r}(s)$ is analytic for $\Re s > \frac13$, for which it suffices to show that the quantity in parentheses has a double zero at $s=1$. First, plugging in $s=1$ yields the value

$$r_6(1) + C_0 - \frac12\log 2 + C_2 = r_6(1) - \frac12\log 2 + \sum_{k=2}^\infty\sum_p \frac{1}{kp^k},$$

whose vanishing follows from (3.15) written in the form

$$r_6(s) - \frac12\log\zeta(2s) = \sum_p \frac{1}{p^s} - \log\zeta(s) = -\sum_{k\ge 2}\sum_p \frac{1}{kp^{ks}}.$$

Second, taking the derivative of the expression in parentheses and plugging in $s=1$ yields

$$r_6'(1) - C_2 2^{1-1}\log 2 - r_6'(1) + C_2 2^{1-1}\log 2 = 0$$

as required. ∎

LEMMA 3.18. *Suppose $f \in \mathcal{B}\cup\mathcal{B}_r$ has the property that $E^{\Pi_r}(x) - E^f(x)$ is unbounded above and $E^{\pi_r}(x) - E^f(x)$ is unbounded below. Then $E^{\pi_\ell}(x) - E^f(x)$ is both unbounded above and unbounded below.*

*Proof.* By Lemma 3.3, there exists some $Z_0$ such that for any real $z \geq Z_0$, we have $E^{\Pi_r}(z) < E^{\pi_\ell}(z) < E^{\pi_r}(z)$ and thus

$$E^{\Pi_r}(z) - E^f(z) < E^{\pi_\ell}(z) - E^f(z) < E^{\pi_r}(z) - E^f(z).$$

The lemma follows immediately, since $E^{\pi_\ell}(x) - E^f(x)$ is eventually greater than a function that is unbounded above, and eventually less than a function that is unbounded below. ∎

*Proof of Theorem 1.6(b).* First, assume that RH is true, and choose any functions $f$ and $g$ in $\mathcal{B} \cup \mathcal{B}_r$ with $\beta_f < \beta_g$. Propositions 2.1 and 2.4 show that $E^g(x) - E^f(x) = \beta_g - \beta_f + o(1) \geq 1 + o(1)$, which immediately implies that $E^f(x) < E^g(x)$ when $x$ is sufficiently large.

Now, assume that RH is false. We first work out the special case $f = \theta$ and $g = \psi_r$, and afterwards we indicate how the proof extends to all other cases except those involving $g = \pi_\ell$. By Lemmas 3.11 and 3.14, for any real number $L$ and any complex number $s$ with $\Re s > 1$,

$$\begin{aligned}(3.16)\qquad &\int_2^\infty (E^\theta(x) - E^{\psi_r}(x) + L)x^{-s-1/2}\,dx \\ &\quad = \left(-\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s}\frac{\zeta'}{\zeta}(2s) + \frac{2^{1-s}}{1-s} + r_\theta(s)\right) \\ &\qquad - \left(\frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \frac{2^{1-s}}{(1-s)^2} + \frac{2^{1-s}(\log 2 - C_0)}{1-s}\right) + \frac{2^{1/2-s}L}{s-1/2}.\end{aligned}$$

We can verify that the right-hand side has a removable singularity at $s = 1$ and no real singularities for $s > \frac{1}{2}$. Moreover, the right-hand equals $\frac{1}{s(s-1)}\frac{\zeta'}{\zeta}(s)$ plus a function with no singularities in the half-plane $\Re s > \frac{1}{2}$ except at $s = 1$; in particular, since we are assuming that RH is false, the right-hand side has non-real singularities in the half-plane $\Re s > \frac{1}{2}$.

Since the analytic continuation of the left-hand side of equation (3.16) has non-real singularities to the right of all its real singularities, we conclude from Landau's theorem [12, Lemma 15.1] that the integrand cannot eventually be of constant sign. In particular, the difference $E^\theta(x) - E^{\psi_r}(x)$ is neither bounded above nor bounded below.

To make this proof work, the key properties we needed the right-hand side to satisfy are:

- it has a meromorphic continuation to $\Re s > \frac{1}{3}$,
- it has a removable singularity at $s = 1$ and no real singularities for $s > \frac{1}{2}$;
- it has non-removable singularities at the zeros of $\zeta(s)$.

If we choose any pair of distinct functions $f, g \in \mathcal{B} \cup \mathcal{B}_r \setminus \{\pi_\ell\}$, other than the five pairs in (1.2), we can use the lemmas in this section to verify that all three of these properties hold for $\int_2^\infty (E^f(x) - E^g(x) + L)x^{-s-1/2}\,dx$. (Indeed,

the assumption $\beta_f < \beta_g$ is not even needed for this part of the proof.) The same Landau's theorem argument then finishes the proof of all cases of the theorem.

One further instructive example is $f = \psi_r$ and $g = \Pi$, in which case the relevant terms in the Mellin transform of $E^{\psi_r}(x) - E^{\Pi}(x) + L$ are

$$\frac{1}{1-s}\frac{\zeta'}{\zeta}(s) - \left(-\frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s^2}\log\big(\zeta(s)(s-1)\big)\right) = \frac{1}{s(1-s)}\frac{\zeta'}{\zeta}(s) - \frac{1}{s^2}\log\big(\zeta(s)(s-1)\big).$$

Both terms have singularities at the zeros of $\zeta(s)$; the fact that the singularity of the difference is non-removable follows from the observation that the first term has poles at each zero while the second term has a logarithmic singularity at each zero.

For $f, g \in \mathcal{B} \cup \mathcal{B}_r$ other than the five pairs in (1.2), the above proof can be extended to pairs that include $\pi^\ell$ by the observation that it holds for all $f, g \in \mathcal{B} \cup \mathcal{B}_r \setminus \{\pi_\ell\}$ (except the aforementioned five pairs) and the application of Lemma 3.18 to every $f \in \mathcal{B} \cup \mathcal{B}_r$ except for $\Pi_r$, $\pi_\ell$, and $\pi_r$. ∎

**4. Fourier transforms of distributions and densities.** In this section, we record the Fourier transforms of the nine normalized error terms from Definition 1.2, both individually and jointly, as well as the Fourier transforms (characteristic functions) of the random-variable distributions $\mu_1$ and $\mu_2$ from Definitions 1.9 and 1.11. We then establish properties of these distributions that allow us to prove Theorems 1.12–1.14 and 1.17.

**4.1. Fourier transforms.** Our first lemma allows us to use Fourier transforms to identify when a random variable on a high-dimensional space is really an affine image of a random variable on a lower-dimensional subspace.

LEMMA 4.1. *Let $D \geq d \geq 1$ be integers. Let $\mathbf{X} = (X_1, \dots, X_D)$ be a random variable on $\mathbb{R}^D$ whose distribution $\mu$ has Fourier transform $\hat{\mu}(t_1, \dots, t_D)$ and let $\mathbf{Y} = (Y_1, \dots, Y_d)$ be a random variable on $\mathbb{R}^d$ whose distribution $\nu$ has Fourier transform $\hat{\nu}(t_1, \dots, t_d)$.*

*Suppose there exists a partition $A_1, \dots, A_d$ of $\{1, \dots, D\}$, and real vectors $\vec{c} = (c_1, \dots, c_D)$ and $\vec{r} = (r_1, \dots, r_D)$ with $r_k \neq 0$ for all $1 \leq k \leq D$, such that the identity*

$$\hat{\mu}(t_1, \dots, t_D) = \exp\Big(-i\sum_{k=1}^{D} c_k t_k\Big)\hat{\nu}\Big(\sum_{k\in A_1} r_k t_k, \dots, \sum_{k\in A_d} r_k t_k\Big)$$

*holds for all $(t_1, \dots, t_D) \in \mathbb{R}^D$. Then $\mathbf{X}$ has the same distribution as the*

*random variable*

$$\sum_{j=1}^{d} \sum_{k \in A_j} (r_k Y_j + c_k) \mathbf{e}_k, \tag{4.1}$$

*where* $\{\mathbf{e}_1, \dots, \mathbf{e}_D\}$ *is the standard basis for* $\mathbb{R}^D$.

*Proof.* By definition, the Fourier transform of the random variable (4.1) is

$$\begin{aligned}
&\mathbb{E} \exp\Big(-i \sum_{j=1}^{d} \sum_{k \in A_j} t_k (r_k Y_j + c_k)\Big) \\
&\qquad = \exp\Big(-i \sum_{k=1}^{D} c_k t_k\Big) \mathbb{E} \exp\Big(-i \sum_{j=1}^{d} \sum_{k \in A_j} t_k r_k Y_j\Big) \\
&\qquad = \exp\Big(-i \sum_{k=1}^{D} c_k t_k\Big) \hat{\nu}\Big(\sum_{k \in A_1} r_k t_k, \dots, \sum_{k \in A_d} r_k t_k\Big) \\
&\qquad = \hat{\mu}(t_1, \dots, t_D),
\end{aligned}$$

and therefore the two random variables have the same distribution. ∎

Next we give a criterion for the absolute continuity of a probability measure based on its Fourier transform.

Definition 4.2. The *Bessel function of the first kind* of order 0 is the even function from $\mathbb{R}$ to $\mathbb{R}$ defined by

$$J_0(x) = \frac{1}{\pi} \int_0^{\pi} e^{ix \cos\theta} \, d\theta = \sum_{n=0}^{\infty} \frac{(-1)^n}{4^n} \frac{x^{2n}}{n!^2}$$

(see [1, item 9.1.21] and [16, equation (1.71.1)]). The defining integral shows that $|J_0(x)| \le 1$, and in fact the stronger bound

$$|J_0(x)| \le \min\left\{1, \sqrt{\frac{2}{\pi |x|}}\right\} \tag{4.2}$$

holds for all $x \in \mathbb{R}$ (see [16, Theorem 7.31.2]).

Lemma 4.3. *Let* $\mu$ *be a probability measure on* $\mathbb{R}^d$ *with Fourier transform*

$$\hat{\mu}(\vec{t}\,) = \prod_{m=1}^{\infty} J_0(|\vec{r}_m \cdot \vec{t}\,|)$$

*for some sequence* $(\vec{r}_m)$ *of vectors in* $\mathbb{C}^d$. *The measure* $\mu$ *is absolutely continuous with respect to Lebesgue measure on* $\mathbb{R}^d$ *if there exist* $2d+1$ *disjoint sets* $S_1, \dots, S_{2d+1}$ *of indices such that for* $1 \le k \le 2d+1$,

$$\operatorname{Span}(\{\Re \vec{r}_m \colon m \in S_k\} \cup \{\Im \vec{r}_m \colon m \in S_k\}) = \mathbb{R}^d.$$

*In particular, $\mu$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$ if either*

(a) *$d = 1$ and at least three elements of the sequence $\vec{r}_m$ are nonzero, or*
(b) *$d = 2$ and $\{\Re\vec{r}_m, \Im\vec{r}_m\}$ is a basis for $\mathbb{R}^2$ for at least five elements of the sequence $(\vec{r}_m)$.*

*Proof.* Parts (a) and (b) are special cases of the general statement (simply take the $S_k$ to be singleton sets containing the relevant elements $r_m$). We prove the general statement using the fact (see [11, Lemma A.6(b)]) that a probability measure $\mu$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$ (which we denote by $\lambda_d$) if $\int_{\mathbb{R}^d} |\hat{\mu}(\vec{t})|\, d\lambda(\vec{t})$ converges.

By assumption, for each $k$ the set $\{\Re\vec{r}_m \colon m \in S_k\} \cup \{\Im\vec{r}_m \colon m \in S_k\}$ contains a basis $B_k$ for $\mathbb{R}^d$. Let $\mathbf{M}_k$ be the $d \times d$ matrix whose rows are the elements of $B_k$, so that $\mathbf{M}_k$ is invertible. On one hand, the elements of $\mathbf{M}_k\vec{t}$ are $\{\vec{b} \cdot \vec{t} \colon \vec{b} \in B_k\}$, and so $|\mathbf{M}_k\vec{t}|^2 = \sum_{\vec{b}\in B_k} (\vec{b} \cdot \vec{t})^2$; on the other hand, $|\mathbf{M}_k\vec{t}| \geq K_k|\vec{t}|$ for some constant $K_k$ (namely the reciprocal of the operator norm of $\mathbf{M}_k^{-1}$). It follows that there must be at least one $\vec{b} \in B_k$ such that $(\vec{b} \cdot \vec{t})^2 \geq K_k^2|\vec{t}|^2/d$, and thus at least one $m \in S_k$ such that $|\vec{r}_m \cdot \vec{t}| \geq K_k|\vec{t}|/\sqrt{d}$. Consequently, by the inequality (4.2),

$$|\hat{\mu}(\vec{t})| = \prod_{m=1}^{\infty} |J_0(|\vec{r}_m \cdot \vec{t}|)| \leq \prod_{k=1}^{2d+1} \prod_{m\in S_k} |J_0(|\vec{r}_m \cdot \vec{t}|)| \leq \prod_{k=1}^{2d+1} \min\left\{1, \sqrt{\frac{2\sqrt{d}}{\pi K_k|\vec{t}|}}\right\}$$
$$\leq \min\left\{1, K|\vec{t}|^{-d-1/2}\right\}$$

with $K = \left(\frac{2}{\pi}\right)^{d+1/2} d^{d/2+1/4} \min\{K_1, \ldots, K_{2d+1}\}^{-d-1/2}$. Finally, if $\mathcal{B}$ is the unit ball in $\mathbb{R}^d$ centered at the origin and $\mathcal{B}^C$ its complement, then

$$\int_{\mathbb{R}^d} |\hat{\mu}(\vec{t})|\, d\lambda(\vec{t}) \leq \int_{\mathcal{B}} 1\, d\lambda(\vec{t}) + \int_{\mathcal{B}^C} K|\vec{t}|^{-d-1/2}\, d\lambda(\vec{t}) < \infty,$$

which confirms that $\mu$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$. ∎

**4.2. Properties of the distribution $\mu_2$.** In this section, we formally connect the normalized error terms in Definition 1.2 to the distributions $\mu_1$ and $\mu_2$ from Definitions 1.9 and 1.11, as well as establish some useful properties of these distributions, starting with their Fourier transforms.

LEMMA 4.4. *The Fourier transform of $\mu_1$ is*

$$\hat{\mu}_1(t) = \prod_{m=1}^{\infty} J_0\left(\frac{2t}{\sqrt{\frac{1}{4} + \gamma_m^2}}\right).$$

*The Fourier transform of $\mu_2$ is*

$$\hat{\mu}_2(t_1,t_2)=\prod_{m=1}^{\infty}J_0\Big(\frac{|2\gamma_m(t_1+t_2)+i(t_1-t_2)|}{\frac{1}{4}+\gamma_m^2}\Big).$$

*Proof.* If $Z$ is a random variable uniformly distributed on the complex unit circle and $\vec{\alpha}\in\mathbb{C}^k$, then the Fourier transform of $2\Re(Z\vec{\alpha})$, for any real vector $\vec{t}$, is by definition

$$\mathbb{E}e^{-i\vec{t}\cdot 2\Re(Z\vec{\alpha})}=\mathbb{E}e^{-i\Re(2(\vec{t}\cdot\vec{\alpha})Z)}=\int_0^1\exp\big(-i\Re\big(2(\vec{t}\cdot\vec{\alpha})e^{2\pi i\theta}\big)\big)\,d\theta=J_0(2|\vec{t}\cdot\vec{\alpha}|),$$

where the last equality follows from a standard calculation (see for example [2, starting from (5.1)]). Given an independent collection of many such random variables, the Fourier transform of their sum is simply the product of their individual Fourier transforms. This establishes the lemma both for $\mu_1$, for which each $\vec{\alpha}$ is of the form $1/\big(\frac{1}{2}+i\gamma_m\big)\in\mathbb{C}^1$, and for $\mu_2$, for which each $\vec{\alpha}$ is of the form $\big(1/\big(\frac{1}{2}+i\gamma_m\big),1/\big(-\frac{1}{2}+i\gamma_m\big)\big)\in\mathbb{C}^2$. (In both cases, the square-summability of the series of coefficients implies both the almost-sure convergence of the overall random variable and the convergence of the infinite product that forms the Fourier transform.) ∎

Corollary 4.5. *Assume RH and LI. For every $f\in\mathcal{B}\cup\mathcal{B}_r$, the normalized error term $E^f(x)$ has a limiting logarithmic distribution equal to $\mu_1+\beta_f$.*

*Proof.* The claim follows immediately by applying [2, Theorems 1.4 and 1.9] to the nine explicit formulas given in Propositions 2.1 and 2.4, and confirming that in all cases the result matches the Fourier transform of $\mu_1$ computed in Lemma 4.4. ∎

Remark 4.6. Each of those nine explicit formulas is a sum of terms of the form $2x^{i\gamma}$ weighted by a coefficient. For the four standard functions in Proposition 2.1 that coefficient is $1/\rho=1/\big(\frac{1}{2}+i\gamma\big)$, while for the five reciprocal functions in Proposition 2.4 that coefficient is $1/(\rho-1)=1/\big(-\frac{1}{2}+i\gamma\big)$. The coincidence that $\big|1/\big(\frac{1}{2}+i\gamma\big)\big|=\big|1/\big(-\frac{1}{2}+i\gamma\big)\big|$ is what causes these nine limiting logarithmic distributions to be identical up to the bias factors. Furthermore, the fact that $1/\big(\frac{1}{2}+i\gamma\big)$ and $1/\big(-\frac{1}{2}+i\gamma\big)$ are very close (when $\gamma$ is large) but not equal is what causes the two coordinates of $\mu_2$ to be highly but not perfectly correlated.

Lemma 4.7. *The distribution $\mu_2$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^2$.*

*Proof.* Lemma 4.4 can be written as

$$\hat{\mu}_2(\vec{t})=\prod_{m=1}^{\infty}J_0\big(|\vec{r}_m\cdot\vec{t}\,|\big)\quad\text{where}\quad\vec{r}_m=\frac{2\gamma_m}{1/4+\gamma_m^2}(1,1)+\frac{i}{1/4+\gamma_m^2}(1,-1);$$

since the real and imaginary parts of each $\vec{r}_m$ form a basis for $\mathbb{R}^2$, the absolute continuity of $\mu_2$ follows from Lemma 4.3. ∎

Our last goal of this section is to determine the precise support of $\mu_2$, for which we need a preliminary lemma.

LEMMA 4.8. *Assume RH. For any real number $0 \le v < \frac{w}{2}$, there exists a set $M \subset \mathbb{N}$ of positive integers, containing only finitely many pairs of consecutive integers, such that $\sum_{m\in M} 1/\big(\frac14+\gamma_m^2\big) = v$.*

*Proof.* We prove that the greedy algorithm produces such a subset $M = \{m_1, m_2, \dots\}$: given $m_1, \dots, m_{k-1}$, let $m_k$ be the smallest integer exceeding $m_{k-1}$ such that $\sum_{j=1}^{k} 1/\big(\frac14+\gamma_{m_j}^2\big) \le v$. Since $\sum_{m\in\mathbb{N}} 1/\big(\frac14+\gamma_m^2\big) = \frac{w}{2}$ by (2.5), this greedy algorithm produces a finite or infinite set $M$ such that $\sum_{m\in M} 1/\big(\frac14+\gamma_m^2\big) = v$; the fact that equality holds when $M$ is infinite follows in a standard way from the fact that each $1/\big(\frac14+\gamma_m^2\big)$ is smaller than $\sum_{j=m+1}^{\infty} 1/\big(\frac14+\gamma_j^2\big)$, which itself follows from the fact that $\gamma_{m+2} < \gamma_m\sqrt{2}$ when $m$ is sufficiently large. The same fact shows that $M$ cannot contain every integer from some point onwards, since the greedy algorithm would have actually selected the last integer not in $M$.

To argue that $M$ contains only finitely many pairs of consecutive integers, it thus suffices to show that the set of integers $m$ with $m-1 \notin M$ and $m, m+1 \in M$ is bounded above. Indeed, for any such $m$, we must have

$$\frac{1}{\frac14+\gamma_m^2} + \frac{1}{\frac14+\gamma_{m+1}^2} \le t - \sum_{j\in M\cap\{1,\dots,m-1\}} \frac{1}{\frac14+\gamma_j^2} < \frac{1}{\frac14+\gamma_{m-1}^2}.$$

But the classic counting function for the zeros of $\zeta(s)$ implies that $\gamma_m = \frac{2\pi m}{\log m}\big(1+O\big(\frac{1}{\log m}\big)\big)$, and therefore we would have

$$\frac{\log^2 m}{4\pi^2 m^2}\left(1+O\left(\frac{1}{\log m}\right)\right) + \frac{\log^2(m+1)}{4\pi^2(m+1)^2}\left(1+O\left(\frac{1}{\log m}\right)\right) < \frac{\log^2(m-1)}{4\pi^2(m-1)^2}\left(1+O\left(\frac{1}{\log m}\right)\right),$$

which is false when $m$ is sufficiently large. ∎

LEMMA 4.9. *Assume RH. The support of the distribution $\mu_2$ is equal to the diagonal strip $\{(x,y)\in\mathbb{R}^2 : |x-y| \le w\}$, where $w$ is as in* (1.1).

*Proof.* By Definition 1.11, the absolute difference between the two components of $V_2$ is

$$\left|2\Re\sum_{m=1}^{\infty} Z_m\left(\frac{1}{\frac12+i\gamma_m} - \frac{1}{-\frac12+i\gamma_m}\right)\right| = \left|2\Re\sum_{m=1}^{\infty}\frac{Z_m}{\frac14+\gamma_m^2}\right| \le 2\sum_{m=1}^{\infty}\frac{1}{\frac14+\gamma_m^2} = \sum_{\rho}\frac{1}{\rho(1-\rho)} = w$$

by Lemma 2.2 (this is the same computation as in equation (2.5)). Thus, the support of $\mu_2$ is contained in the given diagonal strip.

To prove the converse, it suffices to show that the open diagonal strip is contained in the support, since the support is always closed. Any point $(x, y)$ in the open diagonal strip can be written as $(x, y) = (1, 1)u + (1, -1)v$ where $|v| < w/2$. By the symmetry of $\mu_2$ it suffices to consider $0 \le v < w/2$. By Lemma 4.8 we may choose $M \subset \mathbb{N}$, containing only finitely many pairs of consecutive integers, such that $\sum_{m\in M} 1/\big(\frac{1}{4} + \gamma_m^2\big) = v$.

We are now ready to assign values to the independent random variables $Z_m$ in Definition 1.11 that force $V_2 = (x, y)$. Note that

$$\begin{aligned} V_2 &= \sum_{m=1}^{\infty} \left( 2\Re \frac{Z_m}{\frac{1}{2} + i\gamma_m}, 2\Re \frac{Z_m}{-\frac{1}{2} + i\gamma_m} \right) \\ &= \Re \sum_{m=1}^{\infty} \left( -iZ_m \frac{2\gamma_m}{\frac{1}{4} + \gamma_m^2}(1, 1) + Z_m \frac{1}{\frac{1}{4} + \gamma_m^2}(1, -1) \right). \end{aligned}$$

Set $Z_m = 1$ for each $m \in M$, and set $Z_m = \pm i$ for each $m \in \mathbb{N} \setminus M$, so that

$$V_2 = (1, 1) \sum_{m\in\mathbb{N}\setminus M} \pm \frac{2\gamma_m}{\frac{1}{4} + \gamma_m^2} + (1, -1) \sum_{m\in M} \frac{1}{\frac{1}{4} + \gamma_m^2}.$$

The second sum equals $v$ by construction. Since $M$ contains only finitely many pairs of consecutive integers, $\mathbb{N} \setminus M$ is a set of integers of density at least $\frac{1}{2}$; therefore, the first sum diverges when all terms are positive. Consequently, by the Riemann rearrangement theorem, there is a choice of signs that makes that sum equal $u$, as required. ∎

**4.3. Joint distribution and densities.** We can now establish the joint limiting logarithmic distribution of all nine normalized error terms from Definition 1.2, from which we can deduce Theorems 1.12–1.14 and 1.17.

Proposition 4.10. *Assume RH and LI. Define the $\mathbb{R}^9$-valued function*

$$\begin{aligned} \mathbf{E}(x) = \big(&E^{\psi}(x), E^{\theta}(x), E^{\Pi}(x), E^{\pi}(x), \\ &E^{\psi_r}(x), E^{\theta_r}(x), E^{\Pi_r}(x), E^{\pi_r}(x), E^{\pi_\ell}(x)\big). \end{aligned}$$

*Then the limiting logarithmic distribution of $\mathbf{E}(x)$ exists and is the same as the distribution of the random variable*

$$(0, -1, 0, -1, 0, 1, 0, 1, 1) + Y_1(1, 1, 1, 1, 0, 0, 0, 0, 0) + Y_2(0, 0, 0, 0, 1, 1, 1, 1, 1)$$

*where $Y_1$ and $Y_2$ are the components of the $\mathbb{R}^2$-valued random variable $V_2$ from Definition* 1.11.

Note that the constant vector is exactly the list of bias terms for the nine functions in order.

*Proof of Proposition 4.10.* Given Propositions 2.1 and 2.4, we may apply [2, Theorems 1.4 and 1.9] to conclude that the limiting logarithmic distribution of $\mathbf{E}(x)$ exists and has Fourier transform

$$\exp\bigl(-i(-\xi_2-\xi_4+\xi_6+\xi_8+\xi_9)\bigr) \cdot \prod_{m=1}^{\infty} J_0\biggl(2\biggl|\frac{\xi_1+\xi_2+\xi_3+\xi_4}{\frac{1}{2}+i\gamma_m}+\frac{\xi_5+\xi_6+\xi_7+\xi_8+\xi_9}{-\frac{1}{2}+i\gamma_m}\biggr|\biggr).$$

In light of the formula for $\hat{\mu}_2$ from Lemma 4.4, the proposition now follows from an application of Lemma 4.1 with $D=9$ and $d=2$, $A_1=\{1,2,3,4\}$, $A_2=\{5,6,7,8,9\}$, $(c_1,\dots,c_9)=(0,-1,0,-1,0,1,0,1,1)$, and $(r_1,\dots,r_9)=(1,\dots,1)$. ∎

*Proof of Theorem 1.17.* All cases of the theorem follow immediately from Proposition 4.10.

As examples of part (a), if $f$ and $g$ are the standard functions $\psi$ and $\theta$, then the limiting logarithmic distribution of $\bigl(E^{\psi}(x),E^{\pi}(x)\bigr)$ is given by the first two coordinates of the limiting logarithmic distribution of $\mathbf{E}(x)$, which yields $(0,-1)+Y_1(1,1)$ where $Y_1$ is a random variable whose distribution is $\mu_1$. Similarly, if $f$ and $g$ are the reciprocal functions $\pi_r$ and $\pi_\ell$, then the limiting logarithmic distribution of $\bigl(E^{\pi_r}(x),E^{\pi_\ell}(x)\bigr)$ is given by the last two coordinates of the limiting logarithmic distribution of $\mathbf{E}(x)$, which yields $(1,1)+Y_2(1,1)$ where $Y_2$ is also a random variable whose distribution is $\mu_1$. In both cases, the statement of the theorem follows from the properties of $\mu_1$ noted in Definition 1.9.

On the other hand, an example of part (b) is $f=\psi$ and $g=\psi_r$, for which the limiting logarithmic distribution of $\bigl(E^{\psi}(x),E^{\psi_r}(x)\bigr)$ is given by the first and fifth coordinates of the limiting logarithmic distribution of $\mathbf{E}(x)$; this yields $(0,0)+Y_1(1,0)+Y_2(0,1)=(Y_1,Y_2)$, which by definition has distribution $\mu_2$. In this case, the statement of the theorem follows from Lemmas 4.7 and 4.9. ∎

*Proof of Theorem 1.12.* Assuming RH, note that Theorem 1.6 tells us that the hypothesis $\beta_f<\beta_g$ (which holds for all parts of this theorem) implies that $E^f(x)<E^g(x)$ for all sufficiently large $x$. In particular, Corollary 4.5 implies that

$$\delta(0<E^f<E^g)=\delta(0<E^f)=\Pr(0<V_1+\beta_f)=\Pr(-\beta_f<V_1),$$
$$\delta(E^f<E^g<0)=\delta(E^g<0)=\Pr(V_1+\beta_g<0)=\Pr(V_1<-\beta_g),$$

in the notation of Definition 1.9.

(a) When $\beta_f=-1$ and $\beta_g=0$, we have $\Pr(-\beta_f<V_1)=\eta_1$ and $\Pr(V_1<-\beta_g)=\frac{1}{2}$.

(b) When $\beta_f = 0$ and $\beta_g = 1$, we have $\Pr(-\beta_f < V_1) = \frac{1}{2}$ and $\Pr(V_1 < -\beta_g) = \eta_1$.
(c) When $\beta_f = -1$ and $\beta_g = 1$, we have $\Pr(-\beta_f < V_1) = \Pr(V_1 < -\beta_g) = \eta_1$.

In all cases, we then have $\delta(E^f < 0 < E^g) = 1 - \delta(0 < E^f) - \delta(E^g < 0)$ by the absolute continuity of $\mu_1$. ∎

*Proof of Theorem 1.13.* Consider the pairs of functions treated by the theorem: let

$$(f, g) \in \big\{(\pi, \theta),\ (\theta_r, \pi_r),\ (\theta_r, \pi_\ell),\ (\pi_r, \pi_\ell),\ (\psi, \Pi),\ (\psi_r, \Pi_r)\big\},$$

so that $\beta_f = \beta_g$, and $f$ and $g$ are either both standard functions or both reciprocal functions. In these cases, Theorem 1.4(a) implies that $E^f(x) = E^g(x) + o(1)$. Consequently, for every $\varepsilon > 0$,

$$\delta(E^f \le 0 \le E^g) < \delta(-\varepsilon < E^f \le 0) = \mu_1\big((-\varepsilon, 0]\big),$$

and therefore $\delta(E^f \le 0 \le E^g) = 0$ by the absolute continuity of $\mu_1$ (Lemma 4.7). By the same argument, $\delta(E^g \le 0 \le E^f) = 0$ as well. It follows that

$$\delta(0 < E^f \text{ and } 0 < E^g) = \delta(0 < E^f) = \Pr(0 < V_1 + \beta_f) = \Pr(-\beta_f < V_1),$$
$$\delta(E^f < 0 \text{ and } E^g < 0) = \delta(E^f < 0) = \Pr(V_1 + \beta_f < 0) = \Pr(V_1 < -\beta_f),$$

in the notation of Definition 1.9.

(a) When $(f, g) = (\pi, \theta)$, we have $\beta_f = \beta_g = -1$ and therefore $\Pr(-\beta_f < V_1) = \eta_1$ and $\Pr(V_1 < -\beta_f) = 1 - \eta_1$ by definition.
(b) When $(f, g) \in \big\{(\psi, \Pi),\ (\psi_r, \Pi_r)\big\}$, we have $\beta_f = \beta_g = 0$ and therefore $\Pr(-\beta_f < V_1) = \Pr(V_1 < -\beta_f) = \frac{1}{2}$ by symmetry.
(c) Finally, when $(f, g) \in \big\{(\theta_r, \pi_r),\ (\theta_r, \pi_\ell),\ (\pi_r, \pi_\ell)\big\}$, we have $\beta_f = \beta_g = 1$ and therefore $\Pr(-\beta_f < V_1) = 1 - \eta_1$ and $\Pr(V_1 < -\beta_f) = \eta_1$. ∎

*Proof of Theorem 1.14.* Since $f$ is a standard function and $g$ is a reciprocal function with $\beta_f = \beta_g = 0$, Theorem 1.17 gives that the limiting logarithmic distribution of $(E^f(x), E^g(x))$ is $\mu_2$. Definition 1.11 tells us that $\eta_2 = \delta(E^f < 0 < E^g) + \delta(E^g < 0 < E^f)$; but the two densities on the right-hand side are equal by symmetry, and thus each individual density equals $\frac{\eta_2}{2}$. Similarly,

$$\delta(E^f < E^g < 0) + \delta(E^g < E^f < 0) + \delta(0 < E^f < E^g) + \delta(0 < E^g < E^f) = 1 - \eta_2$$

by the absolute continuity of $\mu_2$; but the four densities on the left-hand side are all equal by the symmetries of $\mu_2$, and therefore each individual density equals $\frac{1-\eta_2}{4}$. ∎

We note for future use that this proof implies that the $\mu_2$-measure of the first quadrant is

$$\mu_2\big(\{(x,y)\in\mathbb{R}^2\colon x>0,\, y>0\}\big) = \delta(0<E^f<E^g)+\delta(0<E^g<E^f) = \frac{1-\eta_2}{2}. \tag{4.3}$$

**5. A formula for densities related to $\eta_2$.** It turns out that the calculation of the quantity $\eta_2$ from Definition 1.11 is equivalent to determining the mass that $\mu_2$ assigns to the first quadrant $Q_1=\{(x,y)\in\mathbb{R}^2\colon x>0,\, y>0\}$ (the precise relationship is given in equation (4.3) above). In this section, we follow the method of Feuerverger and the second author [6] to write this mass in terms of a principal value of a two-dimensional integral involving the Fourier transform $\hat{\mu}_2$.

**5.1. Well-behaved products of Bessel functions.** The concept of a "well-behaved function" appeared in [6] in connection with principal values of multivariate integrals. In this paper, we adopt the following slightly more general definition:

Definition 5.1. A function $f$ defined on $\mathbb{R}^n$ is *well-behaved* if it has continuous partial derivatives of all orders and there exist positive constants $\beta_1,\beta_2,\beta_3$ such that for every subset $\{j_1,\dots,j_k\}\subseteq\{1,\dots,n\}$ we have

$$\left|\frac{\partial f}{\partial x_{j_1}\dots\partial x_{j_k}}f(\vec{x})\right| \le \beta_1\exp(-\beta_2|\vec{x}|^{\beta_3})$$

where $|\cdot|_2$ is the Euclidean norm on $\vec{x}$ in $\mathbb{R}^n$.

(The definition in [6] is identical except that it fixes $\beta_3=1$.) The goal of this section is to show that the function $\hat{\mu}_2$ given in Lemma 4.4, and indeed a large class of functions of the same shape, are well-behaved.

Lemma 5.2. *Let $\vec{r}_m$ be a sequence of nonzero vectors in $\mathbb{C}^d$ with $\sum_{m=1}^{\infty}|\vec{r}_m|^2$ finite, and let $f\colon\mathbb{R}^d\to\mathbb{R}$ be the function*

$$f(\vec{t}) = \prod_{m=1}^{\infty} J_0(|\vec{r}_m\cdot\vec{t}\,|). \tag{5.1}$$

*Then $f$ is analytic on $\mathbb{R}^d$ and, in particular, has continuous partial derivatives of all orders.*

*Proof.* In this proof, $\vec{t}=(t_1,\dots,t_d)$ and $\vec{r}_m=(r_{m,a},\dots,r_{m,b})$. Define

$$s(\vec{t})^2=\sum_{a=1}^{d}\sum_{b=1}^{d}t_at_b \quad\text{and}\quad g(\vec{t})=\prod_{m=1}^{\infty}\exp\left(\frac{1}{4}|\vec{r}_m|^2s(\vec{t})^2\right).$$

Since $\sum_{m=1}^{\infty} |\vec{r}_m|^2$ converges, say to $R^2$, the infinite product defining $\vec{g}(t)$ can be written as

$$\begin{aligned} g(\vec{t}) &= \exp\left(\frac{1}{4}\sum_{m=1}^{\infty} |\vec{r}_m|^2 s(\vec{t})^2\right) \\ &= \exp\left(\frac{1}{4}R^2 s(\vec{t})^2\right) = \sum_{n=0}^{\infty} \frac{1}{4^n}\,\frac{R^{2n}(\sum_{a=1}^{d}\sum_{b=1}^{d} t_a t_b)^n}{n!}. \end{aligned}$$

The power series on the right-hand side converges absolutely for all $\vec{t} \in \mathbb{R}^d$, and hence the inner sum can be expanded by the multinomial theorem and rearranged into the form

$$g(t_1, \ldots, t_d) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_d=0}^{\infty} G_{n_1,\ldots,n_d} t_1^{n_1} \cdots t_d^{n_d}$$

for some nonnegative constants $G_{n_1,\ldots,n_d}$. In particular, $g(\vec{t})$ is an analytic function on $\mathbb{R}^d$.

For any positive integer $B$, define

$$\begin{aligned} f_B(\vec{t}) &= \prod_{m=1}^{B} J_0(|\vec{r}_m \cdot \vec{t}\,|) = \prod_{m=1}^{B}\sum_{n=0}^{\infty} \frac{(-1)^n}{4^n}\,\frac{(|\vec{r}_m \cdot \vec{t}\,|^2)^n}{n!^2}, \\ g_B(\vec{t}) &= \prod_{m=1}^{B} \exp\left(\frac{1}{4}|\vec{r}_m|^2 s(\vec{t})^2\right) = \prod_{m=1}^{B}\sum_{n=0}^{\infty} \frac{1}{4^n}\,\frac{(|\vec{r}_m|^2 s(\vec{t})^2)^n}{n!}, \end{aligned}$$

by the Maclaurin series for $J_0(x)$ in Definition 4.2 and the exponential function. Both $f_B(\vec{t})$ and $g_B(\vec{t})$ are certainly analytic on $\mathbb{R}^d$, and thus there exist real constants $F_{B;n_1,\ldots,n_d}$ and $G_{B,n_1,\ldots,n_d}$ such that

$$\begin{aligned} f_B(\vec{t}) &= \sum_{n_1=0}^{\infty} \cdots \sum_{n_d=0}^{\infty} F_{B;n_1,\ldots,n_d} t_1^{n_1} \cdots t_d^{n_d}, \\ g_B(\vec{t}) &= \sum_{n_1=0}^{\infty} \cdots \sum_{n_d=0}^{\infty} G_{B;n_1,\ldots,n_d} t_1^{n_1} \cdots t_d^{n_d}. \end{aligned}$$

Since $g_B(\vec{t})$ is a product of exponential functions, it is again clear that the $G_{B;n_1,\ldots,n_d}$ are nonnegative. Moreover, since

$$|\vec{r}_m \cdot \vec{t}\,|^2 = \sum_{a=1}^{d}\sum_{b=1}^{d} r_{m,a}\bar{r}_{m,b} t_a t_b \quad \text{and} \quad |\vec{r}_m|^2 s(\vec{t})^2 = \sum_{a=1}^{d}\sum_{b=1}^{d} |\vec{r}_m|^2 t_a t_b,$$

the inequality $|r_{m,a}\bar{r}_{m,b}| \le |\vec{r}_m|^2$ implies that $|F_{B;n_1,\ldots,n_d}| \le G_{B;n_1,\ldots,n_d}$ for all nonnegative integers $n_1, \ldots, n_d$. Since $\lim_{B\to\infty} g_B(\vec{t}) = g(\vec{t})$, the dominated convergence theorem implies that $\lim_{B\to\infty} f_B(\vec{t}) = f(\vec{t})$ as well, which shows that $f(\vec{t})$ is analytic on $\mathbb{R}^d$ as claimed. ∎

LEMMA 5.3. *Let $\vec{r}_m$ be a sequence of nonzero vectors in $\mathbb{C}^d$ with $\sum_{m=1}^{\infty} |\vec{r}_m|^2$ finite. For any finite set $M \subseteq \mathbb{N}$, define*

$$f(\vec{t}, M) = \prod_{m \in \mathbb{N} \setminus M} J_0(|\vec{r}_m \cdot \vec{t}|).$$

*Suppose there exist constants $\alpha_1, \alpha_2, R > 0$ and $\alpha_3 > 1$ such that for all $\vec{t} \in \mathbb{R}^d$ with $|\vec{t}| > R$, there exists a set $S(\vec{t}) \subseteq \mathbb{N}$ with $\#S(\vec{t}) \geq \alpha_1 |\vec{t}|^{\alpha_2}$ such that $\frac{\pi}{2}|\vec{r}_m \cdot \vec{t}| \geq \alpha_3$ for all $m \in S(\vec{t})$. Then there exist positive constants $\beta_1, \beta_2, \beta_3$, depending only on $\alpha_1, \alpha_2, \alpha_3, R$, and $\#M$, such that*

$$|f(\vec{t}, M)| \leq \beta_1 \exp(-\beta_2 |\vec{t}|^{\beta_3}) \quad \text{for all } \vec{t} \in \mathbb{R}^d.$$

*Proof.* Since $\sum_{m \in \mathbb{N} \setminus M}^{\infty} |\vec{r}_m|^2 \leq \sum_{m=1}^{\infty} |\vec{r}_m|^2 < \infty$, Lemma 5.2 implies that $f(\vec{t}, M)$ is well defined. Set $N = \#M$. By the inequality (4.2), when $|\vec{t}| > R$ we have

$$|f(\vec{t}, M)| \leq \prod_{m \in S(\vec{t}) \setminus M} \sqrt{\frac{2}{\pi |\vec{r}_m \cdot \vec{t}|}} \leq \prod_{m \in S(\vec{t}) \setminus M} \alpha_3^{-1/2} \leq \alpha_3^{-\alpha_1 (|\vec{t}|^{\alpha_2})/2 + N/2}.$$

Hence when $|\vec{t}| > \max\left\{R, \left(\frac{2N}{\alpha_1}\right)^{1/\alpha_2}\right\}$, we have $\frac{N}{2} < \frac{\alpha_1}{4} |\vec{t}|^{\alpha_2}$ and thus

$$|f(\vec{t}, M)| \leq \alpha_3^{-\alpha_1 |\vec{t}|^{\alpha_2}/4} = \exp(-\beta_2 |\vec{t}|^{\beta_3}),$$

where $\beta_2 = \frac{\alpha_1}{4} \log \alpha_3$ and $\beta_3 = \alpha_2$. On the other hand, $|f(\vec{t}, M)| \leq 1$ for all $\vec{t} \in \mathbb{R}^d$, and so when $|\vec{t}| \leq \max\{R, (2N/\alpha_1)^{1/\alpha_2}\}$ we have

$$|f(\vec{t}, M)| \leq \exp(-\beta_2 |\vec{t}|^{\beta_3}) \exp(\beta_2 \max\{R^{\beta_3}, 2N/\alpha_1\}).$$

Together these inequalities show that

$$|f(\vec{t}, M)| \leq \beta_1 \exp(-\beta_2 |\vec{t}|^{\beta_3})$$

for all $\vec{t} \in \mathbb{R}^d$, where $\beta_1 = 1 + \exp(\beta_2 \max\{R^{\beta_3}, 2N/\alpha_1\})$. ∎

DEFINITION 5.4. Define the entire function

$$\tilde{J}_0(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{4^n} \frac{x^n}{n!^2}.$$

Note that $\tilde{J}_0(x) = J_0(\sqrt{x})$ for $x \geq 0$.

LEMMA 5.5. *For every integer $n \geq 1$, we have $\tilde{J}_0^{(n)}(x) \ll_n 1$ for all $x \geq 0$.*

*Proof.* The bound $\tilde{J}_0^{(n)}(x) \ll_n 1$ holds for $0 \leq x \leq 1$ simply by continuity, so we may assume that $x > 1$. The inequalities $|J_0'(x)| \leq x/2$ and $|J_0^{(n)}(x)| \leq 1$ for $n \geq 2$ follow easily by differentiating under the integral sign in Definition 4.2 (as pointed out in [6, p. 540]). It is easy to prove by induction that there exist constants $\alpha_{n,k}$ such that

$$\begin{aligned}\tilde{J}_0^{(n)}(x) &= \sum_{k=1}^{n} \alpha_{n,k} x^{-n+k/2} J_0^{(k)}(\sqrt{x}) \\ &= \alpha_{n,1} x^{-n+1/2} J_0'(\sqrt{x}) + \sum_{k=2}^{n} \alpha_{n,k} x^{-n+k/2} J_0^{(k)}(\sqrt{x}).\end{aligned}$$

The lemma follows from the fact that each term is bounded for $x > 1$ and $n \geq 1$. ∎

LEMMA 5.6. *Let $\ell \geq 1$ and $i_1, \dots, i_\ell \in \{1, \dots, d\}$ be integers. For any $R > 0$,*

$$\frac{\partial^\ell J_0(|\vec{r} \cdot \vec{t}|)}{\partial t_{i_1} \cdots \partial t_{i_\ell}} \ll_{\ell,R} |\vec{r}|^2 \max\{1, |\vec{t}|^\ell\}$$

*uniformly for $\vec{t} = (t_1, \dots, t_d) \in \mathbb{R}^d$ and $\vec{r} \in \mathbb{C}^d$ with $|\vec{r}| \leq R$.*

*Proof.* Our strategy uses the identity $J_0(|\vec{r} \cdot \vec{t}|) = \tilde{J}_0\big(|\vec{r} \cdot \vec{t}|^2\big)$. (This identity reflects the observation that $J_0$ is an even function and therefore $J_0\big(|\vec{r} \cdot \vec{t}|\big)$ is still smooth despite the square root implicit in $|\vec{r} \cdot \vec{t}|$.) Write $\vec{r} = (r_1, \dots, r_d)$. To guide our intuition, we compute explicitly that for any $i, j \in \{1, \dots, d\}$,

$$\begin{aligned}\frac{\partial}{\partial t_i} \tilde{J}_0\big(|\vec{r} \cdot \vec{t}|^2\big) &= \tilde{J}_0'\big(|\vec{r} \cdot \vec{t}|^2\big) \frac{\partial}{\partial t_i} |\vec{r} \cdot \vec{t}|^2 \\ &= \tilde{J}_0'\big(|\vec{r} \cdot \vec{t}|^2\big) \Big(2\Re r_i \sum_{k=1}^{d} \Re r_k t_k + 2\Im r_i \sum_{k=1}^{d} \Im r_k t_k\Big), \\ \frac{\partial^2}{\partial t_i \partial t_j} \tilde{J}_0\big(|\vec{r} \cdot \vec{t}|^2\big) &= \tilde{J}_0''\big(|\vec{r} \cdot \vec{t}|^2\big) \Big(2\Re r_i \sum_{k=1}^{d} \Re r_k t_k + 2\Im r_i \sum_{k=1}^{d} \Im r_k t_k\Big)^2 \\ &\quad + \tilde{J}_0'\big(|\vec{r} \cdot \vec{t}|^2\big)(2\Re r_i \Re r_j + 2\Im r_i \Im r_j).\end{aligned}$$

In general, one can prove by induction on $\ell \geq 1$ that

$$\frac{\partial^\ell \tilde{J}_0\big(|\vec{r} \cdot \vec{t}|^2\big)}{\partial t_{i_1} \cdots \partial t_{i_\ell}} = \sum_{\ell/2 \leq h \leq \ell} P_{2h,2h-\ell}(\vec{r}, \vec{t}) \tilde{J}_0^{(h)}\big(|\vec{r} \cdot \vec{t}|^2\big),$$

where $P_{a,b}(\vec{r}, \vec{t})$ is a polynomial that is homogeneous of degree $a$ in the variables $\Re r_k$ and $\Im r_k$ and also homogeneous of degree $b$ in the variables $t_k$. It follows from Lemma 5.5 that

$$\frac{\partial^\ell \tilde{J}_0\big(|\vec{r} \cdot \vec{t}|^2\big)}{\partial t_{i_1} \cdots \partial t_{i_\ell}} \ll_\ell \sum_{\ell/2 \leq h \leq \ell} |\vec{r}|^{2h} |\vec{t}|^{2h-\ell} \cdot 1 \ll_{\ell,R} \sum_{\ell/2 \leq h \leq \ell} |\vec{r}|^2 |\vec{t}|^{2h-\ell},$$

which implies the statement of the lemma. ∎

LEMMA 5.7. *Let $\vec{r}_m$ be a sequence of nonzero vectors in $\mathbb{C}^d$ with $\sum_{m=1}^{\infty} |\vec{r}_m|^2$ finite, and let $f(\vec{t})$ be as in equation (5.1). Suppose there exist constants $\alpha_1, \alpha_2 > 0$ and $\alpha_3, \tau > 1$ such that for all $\vec{t} \in \mathbb{R}^d$ with $|\vec{t}| > \tau$, there exists a*

*set* $S(\vec{t}) \subseteq \mathbb{N}$ *with* $\#S(\vec{t}) \geq \alpha_1|\vec{t}|^{\alpha_2}$ *such that* $\frac{\pi}{2}|\vec{r}_m \cdot \vec{t}| \geq \alpha_3$ *for all* $m \in S(\vec{t})$. *Then* $f(\vec{t})$ *is well-behaved.*

*Proof.* By Lemma 5.2, $f(\vec{t})$ is analytic on $\mathbb{R}^d$ with continuous partial derivatives of all orders. It therefore suffices to establish the bound in Definition 5.1 on partial derivatives of $f(\vec{t})$; indeed, we will establish such a bound even when multiple derivatives with respect to the same variable are present. We note that the convergence of $\sum_{m=1}^{\infty} |\vec{r}_m|^2$ implies the existence of a real number $R > 0$ such that $|\vec{r}_m| \leq R$ for all $m \in \mathbb{N}$.

By repeated application of the product rule,

$$\frac{\partial^n f(\vec{t})}{\partial t_{j_1} \cdots \partial t_{j_n}} = \sum_{M=(m_1,\ldots,m_n)\in\mathbb{N}^n} \prod_{m\in\mathbb{N}\setminus M} J_0\big(|\vec{r}_m \cdot \vec{t}|\big) \prod_{m\in M} D_{M,m} J_0\big(|\vec{r}_m \cdot \vec{t}|\big),$$

where $D_{M,m}$ is the composition of the $\ell(m)$ operators $\frac{\partial}{\partial t_{j_i}}$ over all $1 \leq i \leq n$ such that $m_i = m$. (We are abusing notation slightly: $m \in M = (m_1, \ldots, m_n)$ means $m \in \{m_1, \ldots, m_n\}$.) By Lemma 5.3, there exist positive constants $\beta_1, \beta_2, \beta_3$ (depending only on $\alpha_1, \alpha_2, \alpha_3, R$, and $\#M \leq n$) such that the first product is at most $\beta_1 \exp(-\beta_2|\vec{t}|^{\beta_3})$ in absolute value for all $\vec{t} \in \mathbb{R}^d$. Therefore by Lemma 5.6, applied $\#M$ times with the sum of the relevant values of $\ell(m)$ equaling $n$,

$$\begin{aligned}
\frac{\partial^n f(\vec{t})}{\partial t_{j_1} \cdots \partial t_{j_n}} &\ll_{n,R} \sum_{M=(m_1,\ldots,m_n)\in\mathbb{N}^n} \exp(-\beta_2|\vec{t}|^{\beta_3}) \prod_{m\in M} |\vec{r}|^2 \max\{1, |\vec{t}|^{\ell(m)}\} \\
&= \exp(-\beta_2|\vec{t}|^{\beta_3}) \max\{1, |\vec{t}|^n\} \sum_{M=(m_1,\ldots,m_n)\in\mathbb{N}^n} \prod_{m\in M} |\vec{r}|^2 \\
&\leq \exp(-\beta_2|\vec{t}|^{\beta_3}) \max\{1, |\vec{t}|^n\} \sum_{L=1}^{n} \left\{ {n \atop L} \right\} \Big( \prod_{m=1}^{\infty} \big(1 + |\vec{r}|^2\big) \Big)^L,
\end{aligned}$$

where the Stirling number $\left\{ {n \atop k} \right\}$ is the number of partitions of $\{1, \ldots, n\}$ into $L$ subsets. Since $\sum_{m=1}^{\infty} |\vec{r}_m|^2 < \infty$, each of these infinite products converges, and therefore

$$\begin{aligned}
\frac{\partial^n f(\vec{t})}{\partial t_{j_1} \cdots \partial t_{j_n}} &\ll_{n,\{r_m\}} \exp(-\beta_2|\vec{t}|^{\beta_3}) \max\{1, |\vec{t}|^n\} \\
&\ll_{n,\beta_2,\beta_3,\{r_m\}} \exp(-\beta_2|\vec{t}|^{\beta_3/2}),
\end{aligned}$$

which completes the proof of the lemma. ∎

Proposition 5.8. *The function* $\hat{\mu}_2$ *is well-behaved.*

*Proof.* By Lemma 4.4, $\hat{\mu}_2(\vec{t}) = \prod_{m=1}^{\infty} J_0(|\vec{r}_m \cdot \vec{t}|)$ with

$$\vec{r}_m = \left( \frac{2\gamma_m + i}{\frac{1}{4} + \gamma_m^2}, \frac{2\gamma_m - i}{\frac{1}{4} + \gamma_m^2} \right),$$

where $\gamma_1, \gamma_2, \ldots$ is the ascending sequence of the positive ordinates of the zeros of the Riemann zeta function. We use the fact that $N(T)$, the number of indices $m$ such that $\gamma_m \le T$, is $\asymp T \log T$. In particular, the convergence $\sum_{m=1}^{\infty} |\vec{r}_m|^2$ follows by partial summation from

$$|\vec{r}_m|^2 = 2\frac{4\gamma_m^2 + 1}{(\frac{1}{4} + \gamma_m^2)^2} \ll \frac{1}{\gamma_m^2}.$$

Since $\gamma_1 \ge 1$, we see that for any $\vec{t} \in \mathbb{R}^2$,

$$\begin{aligned} |\vec{r}_m \cdot \vec{t}| &= \frac{\sqrt{4\gamma_m^2(t_1+t_2)^2 + (t_1-t_2)^2}}{\frac{1}{4} + \gamma_m^2} \\ &\ge \frac{\sqrt{(t_1+t_2)^2 + (t_1-t_2)^2}}{\frac{1}{4} + \gamma_m^2} \\ &= \frac{\sqrt{2}|\vec{t}|}{\frac{1}{4} + \gamma_m^2} \ge \frac{|\vec{t}|}{\gamma_m^2}. \end{aligned}$$

In particular, if we define $S(\vec{t}) = \{m \in \mathbb{N} \colon \gamma_m \le (\pi|\vec{t}|)^{1/2}\}$ for $\vec{t} \ne \vec{0}$, then $\frac{\pi}{2}|\vec{r}_m \cdot \vec{t}| \ge \frac{1}{2}$ for all $m \in S(\vec{t})$. The number of elements in $S(\vec{t})$ is $N((\pi|\vec{t}|)^{1/2}) \gg |\vec{t}|^{1/2} \log |\vec{t}| \gg |\vec{t}|^{1/2}$ for $|\vec{t}| \ge 2$. Therefore the proposition follows from Lemma 5.7 applied with $\alpha_2 = \alpha_3 = \frac{1}{2}$ and $\tau = 2$. ∎

**5.2. Integral formula for the densities.** We are interested in the mass that $\mu_2$ from Definition 1.11 assigns to the first quadrant. The following formula identifies this quantity in terms of the principal value of a two-dimensional integral (see [6, (2-30)] for a precise definition).

PROPOSITION 5.9. *If $Q_1 \subseteq \mathbb{R}^2$ is the first quadrant, then*

$$\mu_2(Q_1) = \frac{1}{4} - \frac{1}{4\pi^2}\,\mathrm{P.V.} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)}{uv}\,du\,dv.$$

*Proof.* By the Fourier inversion formula and the monotone convergence theorem,

$$\mu_2(Q_1) = \lim_{c \to 0^+} \frac{1}{4\pi^2} \int_0^\infty \int_0^\infty e^{-c(x+y)} \iint\limits_{\mathbb{R}^2} e^{i(ux+vy)} \hat{\mu}_2(u,v)\,du\,dv\,dx\,dy.$$

Hence, by Fubini's Theorem,

$$\begin{aligned} \mu_2(Q_1) &= \lim_{c \to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \hat{\mu}_2(u,v) \int_0^\infty \int_0^\infty e^{(uxi - cx) + (vyi - cy)}\,dx\,dy\,du\,dv \\ &= \lim_{c \to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)}{(c - iu)(c - iv)}\,du\,dv \end{aligned}$$

$$= \lim_{c\to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)(c+iu)(c+iv)}{(u^2+c^2)(v^2+c^2)}\,du\,dv$$
$$= \lim_{c\to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)(c^2-uv+ic(u+v))}{(u^2+c^2)(v^2+c^2)}\,du\,dv.$$

Therefore

$$\begin{aligned}\mu_2(Q_1) = &\lim_{c\to 0^+} \frac{c^2}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)}{(u^2+c^2)(v^2+c^2)}\,du\,dv\\ &+ \lim_{c\to 0^+} \frac{ic}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)(u+v)}{(u^2+c^2)(v^2+c^2)}\,du\,dv\\ &- \lim_{c\to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)uv}{(u^2+c^2)(v^2+c^2)}\,du\,dv.\end{aligned}$$

The second integral vanishes since $\hat{\mu}_2(u,v) = \hat{\mu}_2(-u,-v)$. Hence, by the change of variables $\alpha = \frac{u}{c}$, $\beta = \frac{v}{c}$ in the first integral,

$$\begin{aligned}\mu_2(Q_1) = &\lim_{c\to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(\alpha c,\beta c)}{(1+\alpha^2)(1+\beta^2)}\,d\alpha\,d\beta\\ &- \lim_{c\to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)uv}{(u^2+c^2)(v^2+c^2)}\,du\,dv.\end{aligned}$$

Since

$$|\mu_2(\alpha c,\beta c)| \le 1 = |\mu_2(0,0)|,$$

it follows by the dominated convergence theorem that

$$\begin{aligned}\mu_2(Q_1) &= \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{d\alpha\,d\beta}{(1+\alpha^2)(1+\beta^2)}\\ &\quad - \lim_{c\to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)uv}{(u^2+c^2)(v^2+c^2)}\,du\,dv\\ &= \frac{1}{4} - \lim_{c\to 0^+} \frac{1}{4\pi^2} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)uv}{(u^2+c^2)(v^2+c^2)}\,du\,dv.\end{aligned}$$

Since $\hat{\mu}_2$ is well-behaved by Proposition 5.8, the evaluation

$$\mu_2(Q_1) = \frac{1}{4} - \frac{1}{4\pi^2}\ \text{P.V.} \iint\limits_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)}{uv}\,du\,dv.$$

now follows from [6, Lemma 2.6] with $n = k = 2$. (That lemma was derived under the more restrictive definition of "well-behaved" where $\beta_3 = 1$, but the proof is essentially identical for any $\beta_3 > 0$.) ∎

**6. Calculation and error analysis.** We carry out the numerical calculation of $\eta_2$ in this section, with a rigorous error term. Both the numerical approximation itself and the error analysis used in this section closely follow the techniques used by Feuerverger and the second author in [6], which itself was inspired by the methods of Rubinstein and Sarnak [15]. To summarize, in Section 6.1 we rewrite the relevant principal value as a discrete sum over a lattice in $\mathbb{R}^2$. That sum is then truncated in Section 6.2 to a bounded portion of the lattice, and then the summand involving $\hat{\mu}_2$ is approximated in Section 6.3 using only the zeros of $\zeta(s)$ up to a fixed height. Each of these steps introduces errors, and the errors are explicitly bounded in the next three sections.

**6.1. Discretizing the integral.** Our first approximation uses

$$S(\varepsilon) = \varepsilon^2 \sum_{\substack{m,n\in\mathbb{Z}\\ m,n \text{ odd}}}\sum \frac{\hat{\mu}_2(n\varepsilon/2, m\varepsilon/2)}{(m\varepsilon/2)(n\varepsilon/2)}, \tag{6.1}$$

$$\mathrm{Err}_1 = \mathrm{P.V.}\iint_{\mathbb{R}^2} \frac{\hat{\mu}_2(u,v)}{uv}\,du\,dv - S(\varepsilon).$$

Note that the double integral appears in Proposition 5.9 for the mass $\mu_2$ assigned to the first quadrant.

Bounding $\mathrm{Err}_1$ requires tail bounds for the distribution $\mu_2$. As an initial step, we establish such a bound for the one-dimensional distribution $\mu_1$ that equals each component of $\mu_2$. We use the same method as Rubinstein and Sarnak [15], who actually treat $\mu_1$ itself as well as other related distributions; the fact that we record better numerical constants for $\mu_1$ specifically is because they were bounding several distributions simultaneously.

Lemma 6.1. *Assume RH. Let $X$ be a random variable with distribution $\mu_1$. For any real number $x \geq 0.14$,*

$$\Pr(X \geq x) \leq \exp(-1.875(x-0.14)^2).$$

*Proof.* By Definition 1.9,

$$X = \sum_{k=1}^{\infty} r_k \sin(2\pi U_k) \quad \text{where} \quad r_k = \frac{2}{\sqrt{1/4+\gamma_k^2}}.$$

Since $r_k \gg \frac{1}{\gamma_k} \gg \frac{\log k}{k}$ by the counting function for nontrivial zeros of $\zeta(s)$, the series $\sum_{k=1}^{\infty} r_k$ diverges; on the other hand, by Lemma 2.2, $\sum_{k=1}^{\infty} r_k^2 = \sum_{k=1}^{\infty} 4/(1/4+\gamma_k^2) = \sum_{\rho} 2/(\rho(1-\rho)) = 2w$ assuming RH. It follows from [6, Lemma 3.1] that for $x \geq 2r_1$,

$$\Pr(X \geq x) \leq \exp\left(\frac{-3(x-2r_1)^2}{32w}\right).$$

Since $w < 0.05$ and $r_1 < 0.07$, it follows that

$$\Pr(X \geq x) \leq \exp\left(\frac{-3(x-0.14)^2}{32 \cdot 0.05}\right)$$

for $x \geq 0.14$, as claimed. ∎

We will bootstrap this one-dimensional tail bound into a bound for the two-dimensional distribution $\mu_2$, using the following notation.

DEFINITION 6.2. For random variables $X, Y$ with joint distribution $\mu_2$, define $\bar{P}(u,v) = \Pr(X > u,\, Y > v)$, and set $\bar{P}_1(u) = \bar{P}(u, -\infty)$ and $\bar{P}_2(v) = \bar{P}(-\infty, v)$. Also define

$$P^\star(u,v) = \bar{P}(2\pi u, 2\pi v) - \bar{P}_1(2\pi u)\bar{P}_2(2\pi v),$$
$$Q(u,v) = P^\star(u,v) + P^\star(-u,v) + P^\star(u,-v) + P^\star(-u,-v).$$

LEMMA 6.3. *For all $u, v \in \mathbb{R}$,*

$$\begin{aligned} &Q(u,v) \\ &= 2\Pr(X > 2\pi u, Y > 2\pi v) + 2\Pr(X > -2\pi u, Y > 2\pi v) - 2\Pr(Y > 2\pi v) \\ &= 2\Pr(X > 2\pi u, Y > 2\pi v) + 2\Pr(X > 2\pi u, Y > -2\pi v) - 2\Pr(X > 2\pi u). \end{aligned}$$

*Proof.* By definition,

$$\begin{aligned} Q(u,v) = {} & \Pr(X > 2\pi u, Y > 2\pi v) - \Pr(X > 2\pi u)\Pr(Y > 2\pi v) \\ & + \Pr(X > -2\pi u, Y > 2\pi v) - \Pr(X > -2\pi u)\Pr(Y > 2\pi v) \\ & + \Pr(X > 2\pi u, Y > -2\pi v) - \Pr(X > 2\pi u)\Pr(Y > -2\pi v) \\ & + \Pr(X > -2\pi u, Y > -2\pi v) - \Pr(X > -2\pi u)\Pr(Y > -2\pi v). \end{aligned}$$

By the symmetry and absolute continuity of $\mu_1$,

$$\Pr(X > -2\pi u) = 1 - \Pr(X < -2\pi u) = 1 - \Pr(X > 2\pi u),$$
$$\Pr(Y > -2\pi v) = 1 - \Pr(Y < -2\pi v) = 1 - \Pr(Y > 2\pi v).$$

Thus

$$\begin{aligned} Q(u,v) = {} & \Pr(X > 2\pi u, Y > 2\pi v) + \Pr(X > -2\pi u, Y > 2\pi v) \\ & + \Pr(X > 2\pi u, Y > -2\pi v) + \Pr(X > -2\pi u, Y > -2\pi v) - 1. \end{aligned}$$

Then by Lemmas 4.7 and Definition 1.11,

$$\begin{aligned} &\Pr(X > -2\pi u, Y > 2\pi v) \\ &\quad = 1 - \Pr(X < -2\pi u) - \Pr(Y < 2\pi v) + \Pr(X < -2\pi u, Y < 2\pi v) \\ &\quad = 1 - \Pr(X > 2\pi u) - \Pr(Y < 2\pi v) + \Pr(X > 2\pi u, Y > -2\pi v), \\ &\Pr(X > 2\pi u, Y > -2\pi v) \\ &\quad = 1 - \Pr(X < 2\pi u) - \Pr(Y < -2\pi v) + \Pr(X < 2\pi u, Y < -2\pi v) \\ &\quad = 1 - \Pr(X < 2\pi u) - \Pr(Y > 2\pi v) + \Pr(X > -2\pi u, Y > 2\pi v), \end{aligned}$$

$$\begin{aligned}\Pr(X > &-2\pi u, Y > -2\pi v)\\ &= 1 - \Pr(X < -2\pi u) - \Pr(Y < -2\pi v) + \Pr(X < -2\pi u, Y < 2\pi v)\\ &= 1 - \Pr(X > 2\pi u) - \Pr(Y > 2\pi v) + \Pr(X > 2\pi u, Y > 2\pi v).\end{aligned}$$

Thus, $Q(u,v)$ can be simplified even further as

$$\begin{aligned}Q(u,v) &= 2\Pr(X > 2\pi u, Y > 2\pi v) + 2\Pr(X > -2\pi u, Y > 2\pi v)\\ &\quad + 1 - \Pr(X < 2\pi u) - \Pr(Y > 2\pi v)\\ &\quad + 1 - \Pr(X > 2\pi u) - \Pr(Y > 2\pi v) - 1\\ &= 2\Pr(X > 2\pi u, Y > 2\pi v) + 2\Pr(X > -2\pi u, Y > 2\pi v)\\ &\quad - 2\Pr(Y > 2\pi v),\end{aligned}$$

or alternatively as

$$\begin{aligned}Q(u,v) &= 2\Pr(X > 2\pi u, Y > 2\pi v) + 2\Pr(X > 2\pi u, Y > -2\pi v)\\ &\quad + 1 - \Pr(X > 2\pi u) - \Pr(Y < 2\pi v)\\ &\quad + 1 - \Pr(X > 2\pi u) - \Pr(Y > 2\pi v) - 1\\ &= 2\Pr(X > 2\pi u, Y > 2\pi v) + 2\Pr(X > 2\pi u, Y > -2\pi v)\\ &\quad - 2\Pr(X > 2\pi u). \ \blacksquare\end{aligned}$$

LEMMA 6.4. *For all* $u, v \in \mathbb{R}$,

$$|Q(u,v)| \le 2\min\big\{\Pr(X > 2\pi u), \Pr(Y > 2\pi v)\big\}.$$

*Proof.* The trivial inequalities $0 \le \Pr(X > \pm 2\pi u, Y > 2\pi v) \le \Pr(Y > 2\pi v)$ imply that

$$\begin{aligned}-2\Pr(Y > 2\pi v) &= 0 + 0 - 2\Pr(Y > 2\pi v)\\ &\le 2\Pr(X > 2\pi u, Y > 2\pi v) + 2\Pr(X > -2\pi u, Y > 2\pi v)\\ &\quad - 2\Pr(Y > 2\pi v)\\ &\le 2\Pr(Y > 2\pi v) + 2\Pr(Y > 2\pi v) - 2\Pr(Y > 2\pi v)\\ &= 2\Pr(Y > 2\pi v);\end{aligned}$$

similarly, the inequalities $0 \le \Pr(X > 2\pi u, Y > \pm 2\pi v) \le \Pr(X > 2\pi u)$ imply that

$$\begin{aligned}-2\Pr(X > 2\pi u) &= 0 + 0 - 2\Pr(X > 2\pi u)\\ &\le 2\Pr(X > 2\pi u, Y > 2\pi v) + 2\Pr(X > 2\pi u, Y > -2\pi v)\\ &\quad - 2\Pr(X > 2\pi u)\\ &\le 2\Pr(X > 2\pi u) + 2\Pr(X > 2\pi u) - 2\Pr(X > 2\pi u)\\ &= 2\Pr(X > 2\pi u).\end{aligned}$$

The claim now follows directly from Lemma 6.3. ∎

LEMMA 6.5. *We have*

$$|\mathrm{Err}_1| \le 4\pi^2 \sum_{\substack{(\kappa,\lambda)\in\mathbb{Z}_{\ge 0}\\(\kappa,\lambda)\ne(0,0)}} \left|Q\left(\frac{\kappa}{\varepsilon},\frac{\lambda}{\varepsilon}\right)\right|.$$

*Proof.* This assertion is established by the calculations leading to [6, (3-7)]. Those calculations involved a different two-dimensional distribution, but the calculations apply to any (sufficiently nice) two-dimensional distribution, including $\mu_2$. ∎

LEMMA 6.6. *Assume RH. If* $0 < \varepsilon \le 13$ *then*

$$|\mathrm{Err}_1| \le 96\pi^2 \exp\left(-1.875\left(\frac{2\pi}{\varepsilon} - 0.14\right)^2\right).$$

*Proof.* By Lemmas 6.5, 6.4, and 6.1,

$$\begin{aligned}
|\mathrm{Err}_1| &\le 4\pi^2 \sum_{\substack{(\kappa,\lambda)\in\mathbb{Z}_{\ge 0}\\(\kappa,\lambda)\ne(0,0)}} \left|Q\left(\frac{\kappa}{\varepsilon},\frac{\lambda}{\varepsilon}\right)\right| \\
&\le 4\pi^2 \sum_{\substack{(\kappa,\lambda)\in\mathbb{Z}_{\ge 0}\\(\kappa,\lambda)\ne(0,0)}} 2\min\left\{\Pr\left(X > 2\pi\frac{\kappa}{\varepsilon}\right), \Pr\left(Y > 2\pi\frac{\lambda}{\varepsilon}\right)\right\} \\
&\le 8\pi^2 \sum_{\substack{(\kappa,\lambda)\in\mathbb{Z}_{\ge 0}\\(\kappa,\lambda)\ne(0,0)}} \min\left\{\exp\left(-1.875\left(\frac{2\pi\kappa}{\varepsilon} - 0.14\right)^2\right),\right. \\
&\qquad\qquad \left.\exp\left(-1.875\left(\frac{2\pi\lambda}{\varepsilon} - 0.14\right)^2\right)\right\}.
\end{aligned}$$

The inequalities $0 < \varepsilon \le 13$ imply $\frac{2\pi}{\varepsilon} - 0.14 > 0.14$, which means that $\left(\frac{2\pi k}{\varepsilon} - 0.14\right)^2$ is an increasing function of the nonnegative integer variable $k$. Consequently,

$$\begin{aligned}
|\mathrm{Err}_1| &\le 8\pi^2 \sum_{\substack{(\kappa,\lambda)\in\mathbb{Z}_{\ge 0}\\(\kappa,\lambda)\ne(0,0)}} \exp\left(-1.875\left(\frac{2\pi}{\varepsilon}\max\{\kappa,\lambda\} - 0.14\right)^2\right) \\
&\le 16\pi^2 \sum_{\kappa=0}^{\infty} \sum_{\lambda=\max\{1,\kappa\}}^{\infty} \exp\left(-1.875\left(\frac{2\pi}{\varepsilon}\lambda - 0.14\right)^2\right).
\end{aligned}$$

One can check that each term in the inner sum is at most half the previous term when $0 < \varepsilon \le 13$, and therefore

$$|\mathrm{Err}_1| \le 16\pi^2\left(2\exp\left(-1.875\left(\frac{2\pi}{\varepsilon} - 0.14\right)^2\right)\right.$$

$$+\sum_{\kappa=1}^{\infty} 2\exp\left(-1.875\left(\frac{2\pi}{\varepsilon}\kappa - 0.14\right)^2\right)$$
$$\leq 96\pi^2 \exp\left(-1.875\left(\frac{2\pi}{\varepsilon} - 0.14\right)^2\right)$$

as claimed. ∎

**6.2. Truncating the range of summation.** The next step, as in [6], is to compute the error caused by truncating the infinite sum $S(\varepsilon)$ from equation (6.1) to a finite region. In that work, they truncated their sum to run over a bounded square aligned with the coordinate axes. For us, however, the extremely different decay rates of $\mu_2$ in the two diagonal directions means that the same method would necessarily include many points in which the summand is very small, slowing the numerical computations significantly. Instead, we will truncate $S(\varepsilon)$ to run over the rectangular region

(6.2)
$$\mathcal{R}(C_x, C_y) = \left\{(x,y)\colon \left|(x,y)\cdot\left(\tfrac{1}{\sqrt{2}},\tfrac{1}{\sqrt{2}}\right)\right| \leq C_x,\ \left|(x,y)\cdot\left(-\tfrac{1}{\sqrt{2}},\tfrac{1}{\sqrt{2}}\right)\right| \leq C_y\right\}$$

which is rotated with respect to the $xy$-axes. More precisely, for any $C_x, C_y > 0$ we define

(6.3)
$$S(\varepsilon, C_x, C_y) = \varepsilon^2 \sum_{\substack{m,n \text{ odd} \\ (n\varepsilon/2, m\varepsilon/2)\in\mathcal{R}(C_x,C_y)}}\sum \frac{\hat{\mu}_2(n\varepsilon/2, m\varepsilon/2)}{(m\varepsilon/2)(n\varepsilon/2)},$$
$$\mathrm{Err}_2 = S(\varepsilon) - S(\varepsilon, C_x, C_y).$$

By partitioning the square lattice into two complementary diagonal lattices, and exploiting symmetries of $\hat{\mu}_2(x,y)$, we can establish more convenient expressions for $S(\varepsilon)$ and this new truncation.

Definition 6.7. For any $\varepsilon > 0$, define

$$\mathfrak{a}_\varepsilon(k,\ell) = \tfrac{\varepsilon}{2}(1,1) + \varepsilon k(1,1) + \varepsilon\ell(-1,1) = \varepsilon\left(\tfrac{1}{2} + k - \ell, \tfrac{1}{2} + k + \ell\right),$$
$$\mathfrak{b}_\varepsilon(k,\ell) = \tfrac{\varepsilon}{2}(-1,1) + \varepsilon k(1,1) + \varepsilon\ell(-1,1) = \varepsilon\left(-\tfrac{1}{2} + k - \ell, \tfrac{1}{2} + k + \ell\right).$$

Further, for any $\varepsilon, C_x, C_y > 0$ define $\tilde{C}_x = C_x/(\varepsilon\sqrt{2})$ and $\tilde{C}_y = C_y/(\varepsilon\sqrt{2})$.

Lemma 6.8. *For any $\varepsilon > 0$,*

$$S(\varepsilon) = \frac{2}{\varepsilon^2}\sum_{k=0}^{\infty}\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,0))}{(k+1/2)^2} - \frac{2}{\varepsilon^2}\sum_{\ell=0}^{\infty}\frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(0,\ell))}{(\ell+1/2)^2}$$
$$+ \frac{4}{\varepsilon^2}\sum_{k=0}^{\infty}\sum_{\ell=1}^{\infty}\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2 - \ell^2} + \frac{4}{\varepsilon^2}\sum_{k=1}^{\infty}\sum_{\ell=0}^{\infty}\frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell))}{k^2 - (\ell+1/2)^2}.$$

*Further, for any* $\varepsilon, C_x, C_y > 0$,

$$\begin{aligned} S(\varepsilon, C_x, C_y) = \frac{2}{\varepsilon^2} \sum_{0\le k\le \tilde{C}_x - \frac{1}{2}} \frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,0))}{(k+1/2)^2} - \frac{2}{\varepsilon^2} \sum_{0\le \ell\le \tilde{C}_y - \frac{1}{2}} \frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(0,\ell))}{(\ell+1/2)^2} \\ + \frac{4}{\varepsilon^2} \sum_{0\le k\le \tilde{C}_x - \frac{1}{2}} \sum_{1\le \ell\le \tilde{C}_y} \frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2 - \ell^2} \\ + \frac{4}{\varepsilon^2} \sum_{1\le k\le \tilde{C}_x} \sum_{0\le \ell\le \tilde{C}_y - \frac{1}{2}} \frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell))}{k^2 - (\ell+1/2)^2}. \end{aligned}$$

*Proof.* Define the function $s(x,y) = \frac{\hat{\mu}_2(x,y)}{xy}$ and the lattice

$$\mathfrak{L} = \{(n\varepsilon/2, m\varepsilon/2) \colon n, m \text{ are odd integers}\},$$

so that

$$S(\varepsilon) = \sum_{(x,y)\in\mathfrak{L}} s(x,y).$$

From Lemma 4.4 we see that $s(x,y) = s(y,x) = s(-y,-x) = s(-x,-y)$, and thus

$$\begin{aligned} S(\varepsilon) &= \sum_{(x,y)\in\mathfrak{L}} s(x,y) \\ &= \sum_{\substack{(x,y)\in\mathfrak{L}\\ x=y>0}} \big(s(x,y) + s(-x,-y)\big) + \sum_{\substack{(x,y)\in\mathfrak{L}\\ -x=y>0}} \big(s(x,y) + s(y,x)\big) \\ &\quad + \sum_{\substack{(x,y)\in\mathfrak{L}\\ y>|x|}} \big(s(x,y) + s(-y,-x) + s(y,x) + s(-x,-y)\big) \\ &= 2 \sum_{\substack{(x,x)\in\mathfrak{L}\\ x>0}} s(x,x) + 2 \sum_{\substack{(x,-x)\in\mathfrak{L}\\ y>0}} s(x,-x) + 4 \sum_{\substack{(x,y)\in\mathfrak{L}\\ y>|x|}} s(x,y). \end{aligned}$$

It is easy to check that $\mathfrak{L}$ is the disjoint union of the two sublattices

$$\begin{aligned} \mathfrak{A} &= \{\mathfrak{a}_\varepsilon(k,\ell) \colon k,\ell\in\mathbb{Z}\} \\ &= \left\{\mathfrak{a}_\varepsilon\left(\frac{m+n}{4} - \frac{1}{2}, \frac{m-n}{4}\right) \colon n, m \text{ are odd integers, } n \equiv m \pmod 4\right\} \\ &= \{(n\varepsilon/2, m\varepsilon/2) \colon n, m \text{ are odd integers, } n \equiv m \pmod 4\}, \\ \mathfrak{B} &= \{\mathfrak{b}_\varepsilon(k,\ell) \colon k,\ell\in\mathbb{Z}\} \\ &= \left\{\mathfrak{b}_\varepsilon\left(\frac{m+n}{4}, \frac{m-n}{4} - \frac{1}{2}\right) \colon n, m \text{ are odd integers, } n \not\equiv m \pmod 4\right\} \\ &= \{(n\varepsilon/2, m\varepsilon/2) \colon n, m \text{ are odd integers, } n \not\equiv m \pmod 4\}. \end{aligned}$$

Consequently, in this notation, we have

$$\begin{aligned} S(\varepsilon) &= 2\sum_{\substack{(x,x)\in\mathfrak{L}\\ x>0}} s(x,x) + 2\sum_{\substack{(x,-x)\in\mathfrak{L}\\ y>0}} s(x,-x) + 4\sum_{\substack{(x,y)\in\mathfrak{L}\\ y>|x|}} s(x,y) \\ &= 2\sum_{k=0}^{\infty} s\big(\mathfrak{a}_\varepsilon(k,0)\big) - 2\sum_{\ell=0}^{\infty} s\big(\mathfrak{b}_\varepsilon(0,\ell)\big) \\ &\quad + 4\Bigg(\sum_{k=0}^{\infty}\sum_{\ell=1}^{\infty} s\big(\mathfrak{a}_\varepsilon(k,\ell)\big) + \sum_{k=1}^{\infty}\sum_{\ell=0}^{\infty} s\big(\mathfrak{b}_\varepsilon(k,\ell)\big)\Bigg), \end{aligned}$$

which is equivalent to the first assertion of the lemma. The bounds for the sums in the second assertion follow immediately from the definitions of $\mathfrak{a}_\varepsilon$, $\mathfrak{b}_\varepsilon$, and $\mathcal{R}(C_x, C_y)$. ∎

Lemma 6.9. *For any $\varepsilon, C_x, C_y > 0$,*

$$\begin{aligned} |\operatorname{Err}_2| &\le 2\sum_{k>\tilde{C}_x-\frac12} \left|\frac{\hat\mu_2(\mathfrak{a}_\varepsilon(k,0))}{(k+1/2)^2}\right| + 2\sum_{\ell>\tilde{C}_y-\frac12} \left|\frac{\hat\mu_2(\mathfrak{b}_\varepsilon(0,\ell))}{(\ell+1/2)^2}\right| \\ &\quad + 4\sum_{k>\tilde{C}_x-\frac12}\sum_{\ell=1}^{\infty} \left|\frac{\hat\mu_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\right| + 4\sum_{k>\tilde{C}_x}\sum_{\ell=0}^{\infty} \left|\frac{\hat\mu_2(\mathfrak{b}_\varepsilon(k,\ell))}{k^2-(\ell+1/2)^2}\right| \\ &\quad + 4\sum_{0\le k\le\tilde{C}_x-\frac12}\sum_{\ell>\tilde{C}_y} \left|\frac{\hat\mu_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\right| + 4\sum_{1\le k\le\tilde{C}_x}\sum_{\ell>\tilde{C}_y-\frac12} \left|\frac{\hat\mu_2(\mathfrak{b}_\varepsilon(k,\ell))}{k^2-(\ell+1/2)^2}\right| \end{aligned}$$

*in the notation of Definition* 6.7.

*Proof.* Since $\operatorname{Err}_2 = S(\varepsilon) - S(\varepsilon, C_x, C_y)$, this bound follows immediately from Lemma 6.8 and the triangle inequality. ∎

Definition 6.10. For any $J, K \in \mathbb{N}$, define

$$\Xi_J = \prod_{j=1}^{J} \sqrt{\frac{1/4+\gamma_j^2}{\gamma_j}} \quad\text{and}\quad \Upsilon_K = \prod_{j=1}^{K} \sqrt{1/4+\gamma_j^2}.$$

The following lemma encodes a key observation that will be used repeatedly.

Lemma 6.11. *For any $\varepsilon, C_x, C_y > 0$ and $J, K \in \mathbb{N}$,*

$$2\sum_{k>\tilde{C}_x-\frac12} \left|\frac{\hat\mu_2(\mathfrak{a}_\varepsilon(k,0))}{(k+1/2)^2}\right| \le 2\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\left(\frac{\varepsilon\sqrt{2}}{C_x}\right)\left(\frac{\varepsilon\sqrt{2}}{C_x}+\frac{2}{2+J}\right),$$

$$2\sum_{\ell>\tilde{C}_y-\frac12} \left|\frac{\hat\mu_2(\mathfrak{b}_\varepsilon(0,\ell))}{(\ell+1/2)^2}\right| \le 2\Upsilon_K 2^{K/4}\pi^{-K/2}C_y^{-K/2}\left(\frac{\varepsilon\sqrt{2}}{C_y}\right)\left(\frac{\varepsilon\sqrt{2}}{C_y}+\frac{2}{2+K}\right).$$

*Proof.* For any $k \geq 0$, we have $\mathfrak{a}_\varepsilon(k,0) = \varepsilon(k+\frac{1}{2}, k+\frac{1}{2})$ by Definition 6.7. Lemma 4.4 then gives

$$\big|\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,0))\big| = \prod_{m=1}^{\infty} \left| J_0\Big(\frac{2\gamma_m \varepsilon(2k+1)}{1/4+\gamma_m^2}\Big)\right| \leq \prod_{m=1}^{J} \left| J_0\Big(\frac{2\gamma_m \varepsilon(2k+1)}{1/4+\gamma_m^2}\Big)\right|$$

by inequality (4.2), which further implies

$$\big|\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,0))\big| \leq \prod_{m=1}^{J} \sqrt{\frac{1}{\pi}\,\frac{1/4+\gamma_m^2}{\gamma_m \varepsilon(2k+1)}} = \Xi_J(\pi\varepsilon(2k+1))^{-J/2}. \tag{6.4}$$

Therefore

$$\begin{aligned}
2 \sum_{k>\tilde{C}_x-\frac{1}{2}} & \left| \frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,0))}{(k+1/2)^2} \right| \\
&\leq 2\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2} \sum_{k>\tilde{C}_x-\frac{1}{2}} (k+1/2)^{-2-J/2} \\
&\leq 2\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\left(\left(\frac{\varepsilon\sqrt{2}}{C_x}\right)^{2+J/2} + \int_{\tilde{C}_x}^{\infty} u^{-2-J/2}\,du\right) \\
&\leq 2\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\left(\left(\frac{\varepsilon\sqrt{2}}{C_x}\right)^{2+J/2} + \frac{2}{2+J}\left(\frac{\varepsilon\sqrt{2}}{C_x}\right)^{1+J/2}\right) \\
&\leq 2\Xi_J 2^{-J/4}(\pi)^{-J/2} C_x^{-J/2}\left(\left(\frac{\varepsilon\sqrt{2}}{C_x}\right)^{2} + \frac{2}{2+J}\left(\frac{\varepsilon\sqrt{2}}{C_x}\right)\right).
\end{aligned}$$

This establishes the first inequality, and the second follows from a similar calculation. ∎

Definition 6.12. For any $J \in \mathbb{N}$, define

$$\tilde{\mathbf{A}}_J(k,\ell) = \prod_{j=1}^{J} \frac{\sqrt{1/4+\gamma_j^2}}{\big(\ell^2 + 4\gamma_j^2(k+1/2)^2\big)^{1/4}},$$

$$\tilde{\mathbf{B}}_J(k,\ell) = \prod_{j=1}^{J} \frac{\sqrt{1/4+\gamma_j^2}}{\big((\ell+1/2)^2 + 4\gamma_j^2 k^2\big)^{1/4}}.$$

The derivation of the inequality (6.4) generalizes immediately to show that

$$\begin{aligned}
\big|\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))\big| &\leq (\pi\varepsilon)^{-J/2}\tilde{\mathbf{A}}_J(k,\ell), \\
\big|\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell))\big| &\leq (\pi\varepsilon)^{-J/2}\tilde{\mathbf{B}}_J(k,\ell),
\end{aligned}$$

for any integers $k, \ell \geq 0$.

LEMMA 6.13. *For any $\varepsilon > 0$ and $C_x \geq 5\varepsilon\sqrt{2}$ and $J \in \mathbb{N}$,*

$$4 \sum_{k > \tilde{C}_x - \frac{1}{2}} \sum_{1 \leq \ell \leq k} \left| \frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2 - \ell^2} \right|$$

$$\leq 4\Xi_J 2^{-J/4} \pi^{-J/2} C_x^{-J/2} \bigg( \bigg( \frac{\varepsilon\sqrt{2}}{C_x} \bigg) \bigg( 1 + \frac{1}{4\tilde{C}_x - 1} + \frac{1}{2} \log \bigg( \frac{4C_x}{\varepsilon\sqrt{2}} - 1 \bigg) \bigg)$$

$$+ \frac{2}{J} \bigg( 1 + \frac{1}{J} \bigg) \bigg( 1 + \frac{1}{4\tilde{C}_x - 1} \bigg) + \frac{1}{J} \log \bigg( \frac{4C_x}{\varepsilon\sqrt{2}} - 1 \bigg) \bigg),$$

$$4 \sum_{k > \tilde{C}_x} \sum_{0 \leq \ell \leq k-1} \left| \frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell))}{k^2 - (\ell+1/2)^2} \right|$$

$$\leq 4\Xi_J 2^{-J/4} \pi^{-J/2} C_x^{-J/2} \bigg( \bigg( \frac{\varepsilon\sqrt{2}}{C_x} \bigg) \bigg( 1 + \frac{1}{4\tilde{C}_x - 1} + \frac{1}{2} \log \bigg( \frac{4C_x}{\varepsilon\sqrt{2}} - 1 \bigg) \bigg)$$

$$+ \frac{2}{J} \bigg( 1 + \frac{1}{J} \bigg) \bigg( 1 + \frac{1}{4\tilde{C}_x - 1} \bigg) + \frac{1}{J} \log \bigg( \frac{4C_x}{\varepsilon\sqrt{2}} - 1 \bigg) \bigg).$$

*Proof.* Observe for all $k \geq 0$ and $J \in \mathbb{N}$,

$$\tilde{\mathbf{A}}_J(k,\ell) = \Xi_J (k+1/2)^{-J/2} \prod_{j=1}^{J} \sqrt{\frac{\gamma_j}{\sqrt{(\frac{\ell}{k+\frac{1}{2}})^2 + 4\gamma_j^2}}} \leq \Xi_J 2^{-J/2} (k+1/2)^{-J/2}.$$

Consequently,

$$4 \sum_{k > \tilde{C}_x - \frac{1}{2}} \sum_{1 \leq \ell \leq k} \left| \frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2 - \ell^2} \right|$$

$$\leq 4\Xi_J 2^{-J/2} (\pi\varepsilon)^{-J/2} \sum_{k > \tilde{C}_x - \frac{1}{2}} \sum_{1 \leq \ell \leq k} \frac{(k+1/2)^{-J/2}}{(k+1/2)^2 - \ell^2}$$

$$\leq 4\Xi_J 2^{-J/2} (\pi\varepsilon)^{-J/2}$$

$$\cdot \sum_{k > \tilde{C}_x - \frac{1}{2}} (k+1/2)^{-J/2} \bigg( \frac{1}{k+\frac{1}{4}} + \int_1^k \frac{du}{(k+1/2)^2 - u^2} \bigg)$$

$$\leq 4\Xi_J 2^{-J/2} (\pi\varepsilon)^{-J/2}$$

$$\cdot \sum_{k > \tilde{C}_x - \frac{1}{2}} (k+1/2)^{-J/2} \bigg( \frac{1}{k+\frac{1}{4}} + \frac{1}{2k+1} \int_1^k \frac{1}{k+\frac{1}{2}+u} + \frac{1}{k+\frac{1}{2}-u} \, du \bigg)$$

$$\leq 4\Xi_J 2^{-J/2} (\pi\varepsilon)^{-J/2}$$

$$\cdot \sum_{k > \tilde{C}_x - \frac{1}{2}} (k+1/2)^{-J/2} \bigg( \frac{1}{k+\frac{1}{4}} + \frac{1}{2k+1} \log \bigg( 2(2k+1/2) \frac{k-\frac{1}{2}}{k+\frac{3}{2}} \bigg) \bigg)$$

$$\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2} \cdot \sum_{k>\tilde{C}_x-\frac{1}{2}} (k+1/2)^{-J/2}\bigg(\frac{1}{k+1/4}+\frac{1}{2k+1}\log(4k+1)\bigg).$$

As $k+1/2>\tilde{C}_x$ and $u/(u-1/4)$ is a decreasing function of $u$,

$$\bigg(\frac{1}{k+1/4}+\frac{1}{2k+1}\log(4k+1)\bigg) \le \frac{1}{k+1/2}\bigg(\frac{\tilde{C}_x}{\tilde{C}_x-1/4}\bigg(1+\frac{1}{J}\bigg)+\frac{1}{2}\log(4k+1)-\frac{k+1/2}{J(k+1/4)}\bigg)$$

for $C_x>\varepsilon\sqrt{2}$. Hence

$$4\sum_{k>\tilde{C}_x-\frac{1}{2}}\sum_{1\le\ell\le k}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg| \le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2} \cdot \sum_{k>\tilde{C}_x-\frac{1}{2}} (k+1/2)^{-1-J/2}\bigg(\frac{\tilde{C}_x}{\tilde{C}_x-\frac{1}{4}}\bigg(1+\frac{1}{J}\bigg)+\frac{1}{2}\log(4k+1)-\frac{k+\frac{1}{2}}{J(k+\frac{1}{4})}\bigg).$$

By observing that

$$\int u^{-1-J/2}\bigg(\frac{1}{2}\log(4u-1)-\frac{u}{J(u-1/4)}\bigg)\,du=-\frac{1}{J}u^{-J/2}\log(4u-1)+C,$$

and that the integrand can be easily verified to be decreasing when $u\ge\tilde{C}_x\ge 5$ and $J\in\mathbb{N}$, we see that

$$\begin{aligned}
&4\sum_{k>\tilde{C}_x-\frac{1}{2}}\sum_{1\le\ell\le k}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg| \\
&\quad\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)^{1+J/2}\bigg(\frac{\tilde{C}_x}{\tilde{C}_x-\frac{1}{4}}+\frac{1}{2}\log(4\tilde{C}_x-1)\bigg) \\
&\qquad+\int_{\tilde{C}_x}^{\infty} u^{-1-J/2}\frac{\tilde{C}_x}{\tilde{C}_x-\frac{1}{4}}\bigg(1+\frac{1}{J}\bigg)+\frac{1}{2}\log(4u-1)-\frac{u}{J(u-1/4)}\,du\bigg) \\
&\quad\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)^{1+J/2}\bigg(1+\frac{1}{4\tilde{C}_x-1}+\frac{1}{2}\log(4\tilde{C}_x-1)\bigg) \\
&\qquad+\frac{2}{J}\bigg(1+\frac{1}{4\tilde{C}_x-1}\bigg)\bigg(1+\frac{1}{J}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)^{J/2} \\
&\qquad+\frac{1}{J}\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)^{J/2}\log(4\tilde{C}_x-1)\bigg)
\end{aligned}$$

$$\le 4\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\bigg(1+\frac{1}{4\tilde{C}_x-1}+\frac{1}{2}\log(4\tilde{C}_x-1)\bigg)$$
$$+\frac{2}{J}\bigg(1+\frac{1}{4\tilde{C}_x-1}\bigg)\bigg(1+\frac{1}{J}\bigg)+\frac{1}{J}\log(4\tilde{C}_x-1)\bigg).$$

This establishes the first inequality, and the second follows from a similar calculation. ∎

LEMMA 6.14. *For any $\varepsilon>0$ and $C_y>C_x>0$ and $K\in\mathbb{N}$,*

$$4\sum_{0\le k\le\tilde{C}_x-\frac{1}{2}}\sum_{\ell>\tilde{C}_y}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg|$$
$$\le 4\Upsilon_K 2^{K/4}\pi^{-K/2}C_y^{-K/2}\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_y}+\frac{2}{2+K}\bigg)$$
$$\cdot\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}+\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg)\bigg),$$
$$4\sum_{1\le k\le\tilde{C}_x}\sum_{\ell>\tilde{C}_y-\frac{1}{2}}\bigg|\frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell))}{k^2-(\ell+1/2)^2}\bigg|$$
$$\le 4\Upsilon_K 2^{K/4}\pi^{-K/2}C_y^{-K/2}\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_y}+\frac{2}{2+K}\bigg)$$
$$\cdot\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}+\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg)\bigg).$$

*Proof.* By the inequalities in Definition 6.12,

$$\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg|\le(\pi\varepsilon)^{-K/2}\frac{\tilde{\mathbf{A}}_K(k,\ell)}{|(k+1/2)^2-\ell^2|}.$$

Observe that for $\ell>0$,

$$\tilde{\mathbf{A}}_K(k,\ell)=\Upsilon_K\prod_{j=1}^{K}(\ell^2+4\gamma_j^2(k+1/2)^2)^{-1/4}$$
$$=\Upsilon_K\ell^{-K/2}\prod_{j=1}^{K}(1+4\gamma_j^2(k+1/2)^2\ell^{-2})^{-1/4}\le\Upsilon_K\ell^{-K/2}.$$

Hence, for $C_y>C_x>0$,

$$4\sum_{0\le k\le\tilde{C}_x-\frac{1}{2}}\sum_{\ell>\tilde{C}_y}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg|$$
$$\le 4\Upsilon_K(\pi\varepsilon)^{-K/2}\sum_{0\le k\le\tilde{C}_x-\frac{1}{2}}\sum_{\ell>\tilde{C}_y}\ell^{-K/2}(\ell^2-(k+1/2)^2)^{-1}$$

$$\le 4\Upsilon_K(\pi\varepsilon)^{-K/2}\sum_{\ell>\tilde{C}_y}\ell^{-K/2}\bigg(\frac{1}{\ell^2-\frac{C_x^2}{2\varepsilon^2}}+\int_{1/2}^{\frac{C_x}{\varepsilon\sqrt{2}}}\frac{du}{\ell^2-u^2}\bigg)$$

$$\le 4\Upsilon_K(\pi\varepsilon)^{-K/2}\sum_{\ell>\tilde{C}_y}\ell^{-K/2}\bigg(\frac{1}{\ell^2-\frac{C_x^2}{2\varepsilon^2}}+\frac{1}{2\ell}\int_{1/2}^{\frac{C_x}{\varepsilon\sqrt{2}}}\bigg(\frac{1}{\ell+u}+\frac{1}{\ell-u}\bigg)\,du\bigg)$$

$$\le 4\Upsilon_K(\pi\varepsilon)^{-K/2}\sum_{\ell>\tilde{C}_y}\ell^{-K/2}\bigg(\frac{1}{\ell^2-\frac{C_x^2}{2\varepsilon^2}}+\frac{1}{2\ell}\log\bigg(\frac{(\ell+\tilde{C}_x)(\ell-\frac{1}{2})}{(\ell-\tilde{C}_x)(\ell+\frac{1}{2})}\bigg)\bigg)$$

By the Maclaurin series expansion of $\log(1+x)$ as an alternating series,

$$\log\bigg(\frac{(\ell+\tilde{C}_x)(\ell-\frac{1}{2})}{(\ell-\tilde{C}_x)(\ell+\frac{1}{2})}\bigg)\le\log\bigg(\frac{\ell+\tilde{C}_x}{\ell-\tilde{C}_x}\bigg)=\log\bigg(1+\frac{2\tilde{C}_x}{\ell-\tilde{C}_x}\bigg)\le\frac{2\tilde{C}_x}{\ell-\tilde{C}_x},$$

and hence

$$4\sum_{0\le k\le\tilde{C}_x-\frac{1}{2}}\sum_{\ell>\tilde{C}_y}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg|$$
$$\le 4\Upsilon_K(\pi\varepsilon)^{-K/2}\sum_{\ell>\tilde{C}_y}\ell^{-K/2}\bigg(\frac{1}{\ell^2-\frac{C_x^2}{2\varepsilon^2}}+\frac{\tilde{C}_x}{\ell^2-(\tilde{C}_x)\ell}\bigg)\bigg).$$

For $\ell>\tilde{C}_y$ we observe that

$$\frac{1}{\ell^2}\bigg(\frac{\ell^2}{\ell^2-\frac{C_x^2}{2\varepsilon^2}}\bigg)<\frac{1}{\ell^2}\bigg(\frac{C_y^2}{C_y^2-C_x^2}\bigg)=\frac{1}{\ell^2}\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}\bigg),$$

and likewise

$$\frac{1}{\ell^2}\bigg(\frac{\tilde{C}_x\ell^2}{\ell^2-(\tilde{C}_x)\ell}\bigg)<\frac{1}{\ell^2}\tilde{C}_x\bigg(\frac{C_y^2}{C_y^2-C_xC_y}\bigg)=\frac{1}{\ell^2}\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg).$$

Therefore

$$4\sum_{0\le k\le\tilde{C}_x-\frac{1}{2}}\sum_{\ell>\tilde{C}_y}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg|$$
$$\le 4\Upsilon_K(\pi\varepsilon)^{-K/2}\sum_{\ell>\tilde{C}_y}\ell^{-2-K/2}\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}+\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg)\bigg)$$
$$\le 4\Upsilon_K(\pi\varepsilon)^{-K/2}\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)^{2+K/2}+\int_{\tilde{C}_y}^{\infty}u^{-2-K/2}\,du\bigg)$$
$$\cdot\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}+\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg)\bigg)$$

$$\le 4\Upsilon_K(\pi\varepsilon)^{-K/2}\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)^{2+K/2}+\frac{2}{2+K}\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)^{1+K/2}\bigg)$$
$$\cdot\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}+\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg)\bigg)$$
$$\le 4\Upsilon_K 2^{K/4}\pi^{-K/2}C_y^{-K/2}\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)^{2}+\frac{2}{2+K}\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)\bigg)$$
$$\cdot\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}+\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg)\bigg).$$

This completes the proof of the first inequality, and the second is established analogously. ∎

LEMMA 6.15. *For any real numbers $\varepsilon>0$ and $C_x\ge 5\varepsilon\sqrt{2}$ and any $J\in\mathbb{N}$,*

$$4\sum_{k>\tilde{C}_x-\frac{1}{2}}\sum_{\ell\ge k+1}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg|$$
$$\le 4\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\bigg\{\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\bigg(1+\frac{1}{2}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)+\frac{1}{4J}\,\frac{\varepsilon\sqrt{2}}{C_x}\bigg)$$
$$+\frac{2}{J}\bigg(1+\frac{1}{J}\bigg)+\frac{1}{J}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)\bigg\},$$
$$4\sum_{k>\tilde{C}_x}\sum_{\ell\ge k}\bigg|\frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell))}{k^2-(\ell+1/2)^2}\bigg|$$
$$\le 4\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\bigg\{\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\bigg(1+\frac{1}{2}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)+\frac{1}{4J}\,\frac{\varepsilon\sqrt{2}}{C_x}\bigg)$$
$$+\frac{2}{J}\bigg(1+\frac{1}{J}\bigg)+\frac{1}{J}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)\bigg\}.$$

*Proof.* For all $k\ge 0$ and $J\in\mathbb{N}$,

$$\tilde{\mathbf{A}}_J(k,\ell)=\Xi_J(k+1/2)^{-J/2}\prod_{j=1}^{J}\sqrt{\frac{\gamma_j}{\sqrt{(\frac{\ell}{k+\frac{1}{2}})^2+4\gamma_j^2}}}\le\Xi_J 2^{-J/2}(k+1/2)^{-J/2}.$$

Consequently,

$$4\sum_{k>\tilde{C}_x-\frac{1}{2}}\sum_{\ell\ge k+1}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg|$$
$$\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\sum_{k>\tilde{C}_x-\frac{1}{2}}\sum_{\ell\ge k+1}\frac{(k+1/2)^{-J/2}}{\ell^2-(k+1/2)^2}$$
$$\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}$$

$$\begin{aligned}
&\cdot \sum_{k>\tilde{C}_x-\frac{1}{2}} (k+1/2)^{-J/2}\bigg(\frac{1}{(k+1)^2-(k+1/2)^2}+\int_{k+1}^{\infty}\frac{du}{u^2-(k+1/2)^2}\bigg)\\
&\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\\
&\quad\cdot \sum_{k>\tilde{C}_x-\frac{1}{2}} (k+1/2)^{-J/2}\bigg(\frac{1}{k+\frac{3}{4}}+\frac{1}{2k+1}\int_{k+1}^{\infty}\frac{1}{u-k-\frac{1}{2}}-\frac{1}{u+k+\frac{1}{2}}\,du\bigg)\\
&\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\sum_{k>\tilde{C}_x-\frac{1}{2}} (k+1/2)^{-J/2}\bigg(\frac{1}{k+\frac{3}{4}}+\frac{\log(4k+3)}{2k+1}\bigg)\\
&\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\\
&\quad\cdot \sum_{k>\tilde{C}_x-\frac{1}{2}} (k+1/2)^{-1-J/2}\bigg(\frac{k+\frac{1}{2}}{k+\frac{3}{4}}(1+1/J-1/J)+\frac{1}{2}\log(4k+3)\bigg)\\
&\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\\
&\quad\cdot \sum_{k>\tilde{C}_x-\frac{1}{2}} (k+1/2)^{-1-J/2}\bigg(1+\frac{1}{J}+\frac{1}{2}\log(4k+3)-\frac{k+\frac{1}{2}}{J(k+\frac{3}{4})}\bigg).
\end{aligned}$$

Observe that

$$\int u^{-1-J/2}\bigg(\frac{1}{2}\log(4u+1)-\frac{u}{J(u+1/4)}\bigg)\,du=-\frac{1}{J}u^{-J/2}\log(4u+1)+C,$$

and it is easy to verify that the integrand is decreasing for any $u\ge\tilde{C}_x\ge 5$ and $J\in\mathbb{N}$, so

$$\begin{aligned}
&4\sum_{k>\tilde{C}_x-\frac{1}{2}}\sum_{\ell\ge k+1}\bigg|\frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell))}{(k+1/2)^2-\ell^2}\bigg|\\
&\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\\
&\quad\cdot\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)^{1+J/2}\bigg(1+\frac{1}{J}+\frac{1}{2}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)-\frac{\tilde{C}_x}{J(\tilde{C}_x+\frac{1}{4})}\bigg)\\
&\quad+\int_{\tilde{C}_x}^{\infty}u^{-1-J/2}\bigg(1+\frac{1}{J}+\frac{1}{2}\log(4u+1)-\frac{u}{J(u+\frac{1}{4})}\bigg)\,du\bigg)\\
&\le 4\Xi_J 2^{-J/2}(\pi\varepsilon)^{-J/2}\\
&\quad\cdot\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)^{1+J/2}\bigg(1+\frac{1}{J}+\frac{1}{2}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)-\frac{\tilde{C}_x}{J(\tilde{C}_x+\frac{1}{4})}\bigg)\\
&\quad+\frac{2}{J}\bigg(1+\frac{1}{J}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)^{J/2}+\frac{1}{J}\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)^{J/2}\log\bigg(\frac{4Cx}{\varepsilon\sqrt{2}}+1\bigg)\bigg)
\end{aligned}$$

$$\leq 4\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\bigg(\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\bigg(1+\frac{1}{2}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)+\frac{1}{J(\frac{4C_x}{\varepsilon\sqrt{2}}+1)}\bigg)$$
$$+\frac{2}{J}\bigg(1+\frac{1}{J}\bigg)+\frac{1}{J}\log\bigg(\frac{4Cx}{\varepsilon\sqrt{2}}+1\bigg)\bigg).$$

The first inequality of the lemma follows upon noting that $1/J\big(\frac{4C_x}{\varepsilon\sqrt{2}}+1\big)\leq \varepsilon\sqrt{2}/4JC_x$. The second inequality is established analogously. ∎

LEMMA 6.16. *For any $\varepsilon>0$ and $C_y>C_x\geq 5\varepsilon\sqrt{2}$ and $J,K\in\mathbb{N}$,*

$$\begin{aligned}|\mathrm{Err}_2|\leq{}&\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\bigg[2\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_x}+\frac{2}{2+J}\bigg)\\
&+8\bigg\{\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\bigg(1+\frac{1}{4\tilde{C}_x-1}+\frac{1}{2}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}-1\bigg)\bigg)\\
&+\frac{2}{J}\bigg(1+\frac{1}{J}\bigg)\bigg(1+\frac{1}{4\tilde{C}_x-1}\bigg)+\frac{1}{J}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}-1\bigg)\bigg\}\\
&+8\bigg\{\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\bigg(1+\frac{1}{2}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)+\frac{1}{4J}\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\\
&+\frac{2}{J}\bigg(1+\frac{1}{J}\bigg)+\frac{1}{J}\log\bigg(\frac{4C_x}{\varepsilon\sqrt{2}}+1\bigg)\bigg\}\bigg]\\
&+\Upsilon_K 2^{K/4}\pi^{-K/2}C_y^{-K/2}\bigg[2\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_y}+\frac{2}{2+K}\bigg)\\
&+8\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_y}+\frac{2}{2+K}\bigg)\\
&\cdot\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}+\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg)\bigg)\bigg].\end{aligned}$$

*Proof.* Combining Lemmas 6.9, 6.11, and 6.13–6.15 yields

$$\begin{aligned}|\mathrm{Err}_2|\leq{}&2\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\bigg(\frac{\varepsilon\sqrt{2}}{C_x}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_x}+\frac{2}{2+J}\bigg)\\
&+2\Upsilon_K 2^{K/4}\pi^{-K/2}C_y^{-K/2}\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_y}+\frac{2}{2+K}\bigg)\\
&+8\Upsilon_K 2^{K/4}\pi^{-K/2}C_y^{-K/2}\bigg(\frac{\varepsilon\sqrt{2}}{C_y}\bigg)\bigg(\frac{\varepsilon\sqrt{2}}{C_y}+\frac{2}{2+K}\bigg)\\
&\cdot\bigg(1+\frac{C_x^2}{C_y^2-C_x^2}+\tilde{C}_x\bigg(1+\frac{C_x}{C_y-C_x}\bigg)\bigg)\end{aligned}$$

$$
\begin{aligned}
&+ 8\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\bigg(\Big(\frac{\varepsilon\sqrt{2}}{C_x}\Big)\Big(1+\frac{1}{4\tilde{C}_x-1}+\frac{1}{2}\log\Big(\frac{4C_x}{\varepsilon\sqrt{2}}-1\Big)\Big)\\
&+\frac{2}{J}\Big(1+\frac{1}{J}\Big)\Big(1+\frac{1}{4\tilde{C}_x-1}\Big)+\frac{1}{J}\log\Big(\frac{4C_x}{\varepsilon\sqrt{2}}-1\Big)\bigg)\\
&+ 8\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}\bigg(\Big(\frac{\varepsilon\sqrt{2}}{C_x}\Big)\Big(1+\frac{1}{2}\log\Big(\frac{4C_x}{\varepsilon\sqrt{2}}+1\Big)+\frac{1}{4J}\,\frac{\varepsilon\sqrt{2}}{C_x}\Big)\\
&+\frac{2}{J}\Big(1+\frac{1}{J}\Big)+\frac{1}{J}\log\Big(\frac{4C_x}{\varepsilon\sqrt{2}}+1\Big)\bigg),
\end{aligned}
$$

which implies the statement of the lemma. ∎

In the interest of making the optimization step tractable, we will only minimize the exponential terms $\Xi_J 2^{-J/4}\pi^{-J/2}C_x^{-J/2}$ and $\Upsilon_K 2^{K/4}\pi^{-K/2}C_y^{-K/2}$ for a given $C_x$ or $C_y$; in light of Definition 6.10, this means choosing the maximal $J$ and $K$ such that

$$
\sqrt{\frac{1/4+\gamma_J^2}{\gamma_J}} \le 2^{1/4}\pi^{1/2}C_x^{1/2} \quad\text{and}\quad \sqrt{1/4+\gamma_K^2} \le 2^{-1/4}\pi^{1/2}C_y^{1/2}. \tag{6.5}
$$

**6.3. Truncating the infinite product.** For the final step, the function $\hat{\mu}_2$ will be approximated by a finite product of Bessel functions and a correction factor.

In the calculations of the error term, certain constants must be known to arbitrary precision. The following lemmas provide the formulas used to calculate them.

LEMMA 6.17. *For any integer $n \ge 2$,*

$$
\sum_\rho \frac{1}{\rho^n} = 1 - \frac{1}{(n-1)!}\,\frac{d^n}{ds^n}\log\zeta(s)\Big|_{s=0} - \Big(-\frac{1}{2}\Big)^n\zeta(n).
$$

*In particular, $\sum_\rho \frac{1}{\rho^2} = B_2$ and $\sum_\rho \frac{1}{\rho^4} = B_4$, where we define*

$$
\begin{aligned}
B_2 &= 1 - \frac{\pi^2}{24} + 2\zeta''(0) + \log^2(2\pi),\\
B_4 &= 1 - \frac{\pi^4}{1440} - \frac{1}{6}\big(-2\zeta^{(4)}(0) + 8\zeta^{(3)}(0)\log(2\pi)\\
&\qquad - 12\zeta^{(2)}(0)(\zeta^{(2)}(0) + 2\log^2(2\pi)) - 6\log^4(2\pi)\big).
\end{aligned}
$$

*Proof.* By [12, Corollary 10.14], there exists a constant $B$ such that

$$
\frac{d}{ds}\log\zeta(s) = B + \frac{1}{2}\log\pi - \frac{1}{s-1} - \frac{1}{2}\,\frac{\Gamma'}{\Gamma}(s/2+1) + \sum_\rho\Big(\frac{1}{s-\rho}+\frac{1}{\rho}\Big).
$$

Consequently, for $n \geq 2$,

$$\frac{d^n}{ds^n}\log\zeta(s) = \frac{(-1)^n(n-1)!}{(s-1)^n} - \frac{1}{2}\frac{d^{n-1}}{ds^{n-1}}\frac{\Gamma'}{\Gamma}(s/2+1) + \sum_\rho\left(\frac{(-1)^{n-1}(n-1)!}{(s-\rho)^n}\right).$$

Furthermore, by [12, Equation C.11, p. 522],

$$\frac{\Gamma'}{\Gamma}(a) = -\frac{1}{a} - C_0 - \sum_{m=1}^{\infty}\left(\frac{1}{a+m} - \frac{1}{m}\right),$$

and therefore

$$\frac{d}{da}\frac{\Gamma'}{\Gamma}(a) = \frac{1}{a^2} + \sum_{m=1}^{\infty}\frac{1}{(a+m)^2} = \zeta(2,a),$$

where $\zeta(z,a)$ is the Hurwitz zeta-function. Since $\frac{d}{da}\zeta(z,a) = -z\zeta(z+1,a)$, we see that for $n \geq 2$,

$$\frac{d^{n-1}}{da^{n-1}}\frac{\Gamma'}{\Gamma}(a) = (-1)^n(n-1)!\zeta(n,a).$$

Consequently, by the chain rule,

$$\frac{d^n}{ds^n}\log\zeta(s) = \frac{(-1)^n(n-1)!}{(s-1)^n} - \left(-\frac{1}{2}\right)^n(n-1)!\zeta(n,s/2+1) + \sum_\rho\frac{(-1)^{n-1}(n-1)!}{(s-\rho)^n}.$$

The first assertion of the lemma follows by evaluating at $s=0$ (noting that $\zeta(n,1) = \zeta(n)$) and rearranging.

The specific formulas for $\sum_\rho \frac{1}{\rho^2}$ and $\sum_\rho \frac{1}{\rho^4}$ can be obtained by repeated differentiation of $\frac{\zeta'}{\zeta}(s)$ and the known values

$$\zeta(0) = -\frac{1}{2} \text{ and } \zeta'(0) = -\frac{1}{2}\log(2\pi) \quad [12, (10.11) \text{ and } (10.14)]. \blacksquare$$

LEMMA 6.18. *Assuming RH,*

$$\sum_{\gamma>0}\frac{\gamma^2}{(1/4+\gamma^2)^4} = \frac{B_1}{2} + \frac{B_2}{2} - \frac{B_4}{4} \approx 1.43512\ldots\cdot 10^{-7},$$

$$\sum_{\gamma>0}\frac{\gamma^2}{(1/4+\gamma^2)^2} = \frac{B_1 - B_2}{4} \approx 0.0230864\ldots,$$

$$\sum_{\gamma>0}\frac{1}{(1/4+\gamma^2)^2} = B_1 + B_2 \approx 3.71006\ldots\cdot 10^{-5},$$

*where* $B_1 = C_0 + 2 - \log(4\pi)$ *and* $B_2$ *and* $B_4$ *are defined in Lemma* 6.17.

*Proof.* For any complex number $\rho \neq 0, 1$,

$$\frac{-(\rho-1/2)^2}{\rho^4(1-\rho)^4} = \frac{1}{\rho(1-\rho)} + \frac{1/2}{\rho^2} + \frac{1/2}{(1-\rho)^2} + \frac{-1/4}{\rho^4} + \frac{-1/4}{(1-\rho)^4},$$
$$\frac{-(\rho-1/2)^2}{\rho^2(1-\rho)^2} = \frac{1/2}{\rho(1-\rho)} + \frac{-1/4}{\rho^2} + \frac{-1/4}{(1-\rho)^2},$$
$$\frac{1}{\rho^2(1-\rho)^2} = \frac{2}{\rho(1-\rho)} + \frac{1}{\rho^2} + \frac{1}{(1-\rho)^2}.$$

The functional equation for $\zeta(s)$ ensures that its nontrivial zeros come in pairs $\rho$, $1-\rho$, and $\frac{1}{\rho} + \frac{1}{1-\rho} = \frac{1}{\rho(1-\rho)}$, and therefore

$$\sum_\rho \frac{-(\rho-1/2)^2}{\rho^4(1-\rho)^4} = \sum_\rho \frac{1}{\rho(1-\rho)} + \sum_\rho \frac{1}{\rho^2} - \frac{1}{2}\sum_\rho \frac{1}{\rho^4} = B_1 + B_2 - \frac{B_4}{2},$$
$$\sum_\rho \frac{-(\rho-1/2)^2}{\rho^2(1-\rho)^2} = \frac{1}{2}\sum_\rho \frac{1}{\rho(1-\rho)} - \frac{1}{2}\sum_\rho \frac{1}{\rho^2} = \frac{B_1}{2} - \frac{B_2}{2},$$
$$\sum_\rho \frac{1}{\rho^2(1-\rho)^2} = 2\sum_\rho \frac{1}{\rho(1-\rho)} + 2\sum_\rho \frac{1}{\rho^2} = 2B_1 + 2B_2$$

unconditionally by [12, Corollary 10.14] and Lemma 6.17. If we now assume RH, we see that

$$\sum_{\gamma>0} \frac{\gamma^2}{(1/4+\gamma^2)^4} = \frac{1}{2}\sum_\rho \frac{-(\rho-1/2)^2}{\rho^4(1-\rho)^4} = \frac{B_1}{2} + \frac{B_2}{2} - \frac{B_4}{4},$$
$$\sum_{\gamma>0} \frac{\gamma^2}{(1/4+\gamma^2)^2} = \frac{1}{2}\sum_\rho \frac{-(\rho-1/2)^2}{\rho^2(1-\rho)^2} = \frac{B_1 - B_2}{4},$$
$$\sum_{\gamma>0} \frac{1}{(1/4+\gamma^2)^2} = \frac{1}{2}\sum_\rho \frac{1}{\rho^2(1-\rho)^2} = B_1 + B_2,$$

as claimed. ∎

Definition 6.19. For $T > 0$, define the truncated products

$$F_T(u,v) = \prod_{0<\gamma\le T} J_0\Big(\frac{\sqrt{\gamma^2u^2+v^2}}{1/4+\gamma^2}\Big) \text{ and } \tilde{F}_T(u,v) = \prod_{\gamma>T} J_0\Big(\frac{\sqrt{\gamma^2u^2+v^2}}{1/4+\gamma^2}\Big),$$

and define

$$G_T(u,v) = F_T(u,v)\big(1 - P_Tu^2 - Q_Tv^2 + (P_TQ_T + R_T)u^2v^2\big)$$

where

$$P_T = \frac{1}{4}\sum_{\gamma>T} \frac{\gamma^2}{(1/4+\gamma^2)^2}, \quad Q_T = \frac{1}{4}\sum_{\gamma>T} \frac{1}{(1/4+\gamma^2)^2}, \quad R_T = \frac{1}{32}\sum_{\gamma>T} \frac{\gamma^2}{(1/4+\gamma^2)^4}$$

are the tails of the infinite sums from Lemma 6.18. We then define

$$
\begin{aligned}
S(\varepsilon, C_x, C_y, T) = 2 &\sum_{0\le k\le \tilde{C}_x-\frac{1}{2}} \frac{G_T(4\varepsilon(k+1/2),0)}{(k+1/2)^2} \\
&- 2 \sum_{0\le \ell\le \tilde{C}_y-\frac{1}{2}} \frac{G_T(0,2\varepsilon(\ell+1/2))}{(\ell+1/2)^2} \\
&+ 4 \sum_{0\le k\le \tilde{C}_x-\frac{1}{2}} \sum_{1\le \ell\le \tilde{C}_y} \frac{G_T(4\varepsilon(k+1/2),2\varepsilon\ell)}{(k+1/2)^2-\ell^2} \\
&+ 4 \sum_{1\le k\le \tilde{C}_x} \sum_{0\le \ell\le \tilde{C}_y-\frac{1}{2}} \frac{G_T(4\varepsilon k,2\varepsilon(\ell+1/2))}{k^2-(\ell+1/2)^2}
\end{aligned} \tag{6.6}
$$

and the third error term

$$
\mathrm{Err}_3 = S(\varepsilon, C_x, C_y) - S(\varepsilon, C_x, C_y, T). \tag{6.7}
$$

For use in our upper bounds, we also define

$$
\begin{aligned}
\Delta_T(u,v) = &\frac{1}{2(1-P_Tu^2)(1-Q_Tv^2)} \\
&\cdot \left( \left(1-\frac{P_Tu^2}{2}\right) Q_T^2v^4 + P_T^2u^4\left(1-\frac{Q_Tv^2}{2}\right) - \frac{P_T^2u^4Q_T^2v^4}{2} \right).
\end{aligned}
$$

Recall the quantities $\mathfrak{a}_\varepsilon(k,\ell)$, $\mathfrak{b}_\varepsilon(k,\ell)$, $\tilde{C}_x$, and $\tilde{C}_y$ from Definition 6.7.

LEMMA 6.20. *Let $T > 0$, and choose any $0 < C_x < 1/\sqrt{8P_T}$ and $0 < C_y < 1/\sqrt{2Q_T}$.*

(a) *For any $0 \le k \le \tilde{C}_x - \frac{1}{2}$ and $0 \le \ell \le \tilde{C}_y$,*

$$
\begin{aligned}
&|\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell)) - G_T(4\varepsilon(k+1/2),2\varepsilon\ell)| \\
&\qquad \le |F_T(4\varepsilon(k+1/2),2\varepsilon\ell)|\Delta_T(4\varepsilon(k+1/2),2\varepsilon\ell).
\end{aligned}
$$

(b) *For any $0 \le k \le \tilde{C}_x$ and $0 \le \ell \le \tilde{C}_y - \frac{1}{2}$,*

$$
\begin{aligned}
&|\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell)) - G_T(4\varepsilon k,2\varepsilon(\ell+1/2))| \\
&\qquad \le |F_T(4\varepsilon k,2\varepsilon(\ell+1/2))|\Delta_T(4\varepsilon k,2\varepsilon(\ell+1/2)).
\end{aligned}
$$

*Proof.* By the Maclaurin expansion in Definition 4.2,

$$
\tilde{F}_T(u,v) = \prod_{\gamma>T} \sum_{m=0}^{\infty} \frac{(-1)^m}{(m!)^2} \left( \frac{1}{4} \cdot \frac{\gamma^2u^2+v^2}{(1/4+\gamma^2)^2} \right)^m = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} b_{m,n}u^{2m}v^{2n}
$$

where

$$
b_{0,0} = 1, \quad b_{1,0} = -P_T, \quad b_{0,1} = -Q_T, \quad b_{1,1} = P_TQ_T + R_T.
$$

Furthermore, since

$$\begin{aligned}\prod_{\gamma>T}\sum_{m=0}^{\infty}\frac{1}{m!}\left(\frac{1}{4}\cdot\frac{\gamma^2u^2+v^2}{(1/4+\gamma^2)^2}\right)^m &= \prod_{\gamma>T}\exp\left(\frac{1}{4}\cdot\frac{\gamma^2u^2+v^2}{(1/4+\gamma^2)^2}\right)\\ &= \exp(P_Tu^2)\exp(Q_Tv^2)\\ &= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{P_T^mQ_T^n}{m!n!}u^{2m}v^{2n},\end{aligned}$$

we see that

$$|b_{m,n}|\le\frac{P_T^mQ_T^n}{m!n!}.$$

It follows that whenever $P_Tu^2<1$ and $Q_Tv^2<1$,

$$\begin{aligned}\Big|\sum_{m=0}^{1}\sum_{n=2}^{\infty}b_{m,n}u^{2m}v^{2n}\Big| &\le \sum_{m=0}^{1}\sum_{n=2}^{\infty}\frac{P_T^mQ_T^n}{m!n!}u^{2m}v^{2n}\\ &\le \sum_{m=0}^{1}\frac{P_T^m}{m!}u^{2m}\frac{Q_T^2v^4}{(2!)(1-Q_Tv^2)}\\ &= \left(1+\frac{P_Tu^2}{2}\right)\cdot\frac{Q_T^2v^4}{2(1-Q_Tv^2)}\\ &\le \frac{1-\frac{P_Tu^2}{2}-\frac{P_T^2u^4}{2}}{1-P_Tu^2}\cdot\frac{Q_T^2v^4}{2(1-Q_Tv^2)}.\end{aligned}$$

Similarly, under the same assumptions,

$$\begin{aligned}\Big|\sum_{m=2}^{\infty}\sum_{n=0}^{1}b_{m,n}u^{2m}v^{2n}\Big| &\le \sum_{m=2}^{\infty}\sum_{n=0}^{1}\frac{P_T^mQ_T^n}{m!n!}u^{2m}v^{2n}\\ &\le \sum_{m=2}^{\infty}\frac{P_T^m}{m!}u^{2m}\left(1+\frac{Q_Tu^2}{2}\right)\\ &\le \frac{P_T^2u^4}{2(1-P_Tu^2)}\cdot\frac{1-\frac{Q_Tv^2}{2}-\frac{Q_T^2v^4}{2}}{1-Q_Tv^2},\\ \Big|\sum_{m=2}^{\infty}\sum_{n=2}^{\infty}b_{m,n}u^{2m}v^{2n}\Big| &\le \sum_{m=2}^{\infty}\sum_{n=2}^{\infty}\frac{P_T^mQ_T^n}{m!n!}u^{2m}v^{2n}\\ &\le \sum_{m=2}^{\infty}\frac{P_T^m}{m!}u^{2m}\frac{Q_T^2v^4}{2(1-Q_Tv^2)}\\ &\le \frac{P_T^2u^4}{2(1-P_Tu^2)}\cdot\frac{Q_T^2v^4}{2(1-Q_Tv^2)}.\end{aligned}$$

These inequalities combine to show that

$$\left|\tilde{F}_T(u,v) - (1 - P_T u^2 - Q_T v^2 + (P_T Q_T + R_T)u^2 v^2)\right| \le \Delta_T(u,v)$$

whenever $P_T u^2 < 1$ and $Q_T v^2 < 1$. Finally, since

$$\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell)) = F_T(4\varepsilon(k+1/2), 2\varepsilon\ell)\tilde{F}_T(4\varepsilon(k+1/2), 2\varepsilon\ell),$$
$$\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell)) = F_T(4\varepsilon k, 2(\ell+1/2))\tilde{F}_T(4\varepsilon k, 2\varepsilon(\ell+1/2)),$$

the lemma follows by setting $(u,v) = (4\varepsilon(k+1/2), 2\varepsilon\ell)$ and $(u,v) = (4\varepsilon k, 2\varepsilon(\ell+1/2))$. ∎

LEMMA 6.21. *If $P_T < \frac{1}{8C_x^2}$ and $Q_T < \frac{1}{2C_y^2}$, then*

$$\begin{aligned}
|\mathrm{Err}_3| \le\ & 2 \sum_{0 \le k \le \tilde{C}_x - \frac{1}{2}} \frac{|F_T(4\varepsilon(k+1/2), 0)|}{(k+1/2)^2} \Delta_T(4\varepsilon(k+1/2), 0) \\
& + 2 \sum_{0 \le \ell \le \tilde{C}_y - \frac{1}{2}} \frac{|F_T(0, 2\varepsilon(\ell+1/2))|}{(\ell+1/2)^2} \Delta_T(0, 2\varepsilon(\ell+1/2)) \\
& + 4 \sum_{0 \le k \le \tilde{C}_x - \frac{1}{2}} \sum_{1 \le \ell \le \tilde{C}_y} \frac{|F_T(4\varepsilon(k+1/2), 2\varepsilon\ell)|}{|(k+1/2)^2 - \ell^2|} \Delta_T(4\varepsilon(k+1/2), 2\varepsilon\ell) \\
& + 4 \sum_{1 \le k \le \tilde{C}_x} \sum_{0 \le \ell \le \tilde{C}_y - \frac{1}{2}} \frac{|F_T(4\varepsilon k, 2\varepsilon(\ell+1/2))|}{|k^2 - (\ell+1/2)^2|} \Delta_T(4\varepsilon k, 2\varepsilon(\ell+1/2)),
\end{aligned}$$

*with $\Delta_T(u,v)$ as in Definition* 6.19.

*Proof.* By Lemma 6.8,

$$\begin{aligned}
\mathrm{Err}_3 &= S(\varepsilon, C_x, C_y) - S(\varepsilon, C_x, C_y, T) \\
&= 2 \sum_{0 \le k \le \tilde{C}_x - \frac{1}{2}} \frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,0)) - G_T(4\varepsilon(k+1/2), 0)}{(k+1/2)^2} \\
&\quad - 2 \sum_{0 \le \ell \le \tilde{C}_y - \frac{1}{2}} \frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(0,\ell)) - G_T(0, 2\varepsilon(\ell+1/2))}{(\ell+1/2)^2} \\
&\quad + 4 \sum_{0 \le k \le \tilde{C}_x - \frac{1}{2}} \sum_{1 \le \ell \le \tilde{C}_y} \frac{\hat{\mu}_2(\mathfrak{a}_\varepsilon(k,\ell) - G_T(4\varepsilon(k+1/2), 2\varepsilon\ell)}{(k+1/2)^2 - \ell^2} \\
&\quad + 4 \sum_{1 \le k \le \tilde{C}_x} \sum_{0 \le \ell \le \tilde{C}_y - \frac{1}{2}} \frac{\hat{\mu}_2(\mathfrak{b}_\varepsilon(k,\ell)) - G_T(4\varepsilon k, 2\varepsilon(\ell+1/2))}{k^2 - (\ell+1/2)^2}.
\end{aligned}$$

Since $P_T < \frac{1}{8C_x^2}$ and $Q_T < \frac{1}{2C_y^2}$, the rest follows by the triangle inequality and Lemma 6.20. ∎

**6.4. Result of calculation and error term estimation.** Assuming RH, Proposition 5.9 and equations (6.1), (6.3), and (6.7) tell us that

$$\mu_2(Q_1) = \frac{1}{4} - \frac{1}{4\pi^2}\big(S(\varepsilon, C_x, C_y, T) + \mathrm{Err}_1 + \mathrm{Err}_2 + \mathrm{Err}_3\big).$$

We coded the definition of $S(\varepsilon, C_x, C_y, T)$ and of these three error terms in C using the PARI library [17]. Smaller values of $\varepsilon$, and larger values of $C_x$, $C_y$, and $T$, result in smaller error terms at the expense of longer calculation time. We ended up choosing

$$\varepsilon = 1.5, \quad C_x = 30, \quad C_y = 2500, \quad T = 7500. \tag{6.8}$$

With these parameters we found that

$$S(\varepsilon, C_x, C_y, T) \approx -9.60218,$$

and therefore

$$|\mu_2(Q_1) - 0.493226| \le \frac{1}{4\pi^2}(|\mathrm{Err}_1| + |\mathrm{Err}_2| + |\mathrm{Err}_3|).$$

Under the same parameters, Lemma 6.6 gives

$$|\mathrm{Err}_1| \le 4.2458 \cdot 10^{-11};$$

also, the inequalities (6.5) result in the choices $J = 44$ and $K = 18$, in which case Lemma 6.16 gives

$$|\mathrm{Err}_2| \le 8.1525 \cdot 10^{-6}.$$

Finally, our calculations show that $P_{7500} \approx 4.2891 \cdot 10^{-5}$, $Q_{7500} \approx 2.3321 \cdot 10^{-13}$, and $R_{7500} \approx 3.0537 \cdot 10^{-22}$, and therefore Lemma 6.21 gives

$$|\mathrm{Err}_3| \le 3.5443 \cdot 10^{-4}.$$

These error bounds combine to yield

$$|\mu_2(Q_1) - 0.493226| \le 0.00001.$$

Since $\mu_2(Q_1) = \frac{1-\eta_2}{2}$ by (4.3), we deduce that

$$|\eta_2 - 0.013548| \le 0.00002.$$

**Appendix A. Graphs illustrating the main theorems.** We exhibit graphs showing every possible case of the densities from Theorems 1.12–1.14 and of the limiting logarithmic distributions from Theorem 1.17. The graphs are determined solely by whether each function is standard or reciprocal and by the bias factors of the two functions.

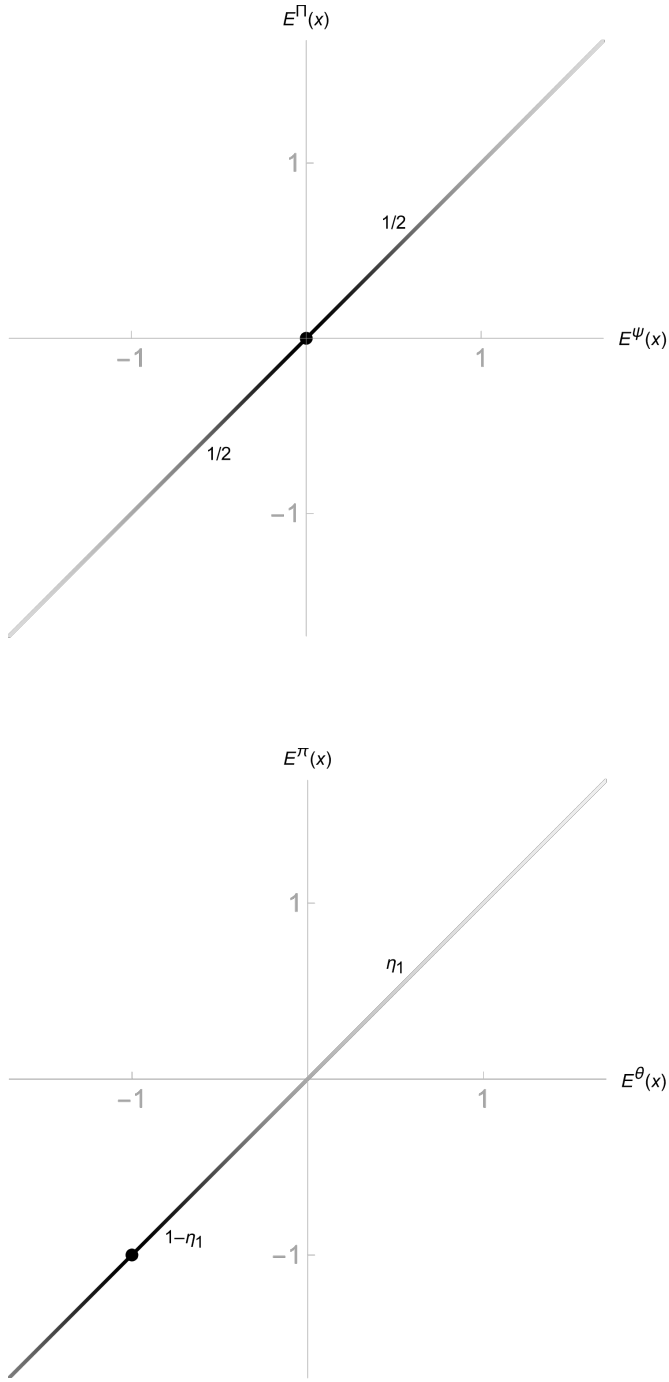



Fig. 2. Two pairs of standard functions with equal bias factors within each pair. Each joint distribution is singular with support equal to the line $y = x$ (perfect correlation).

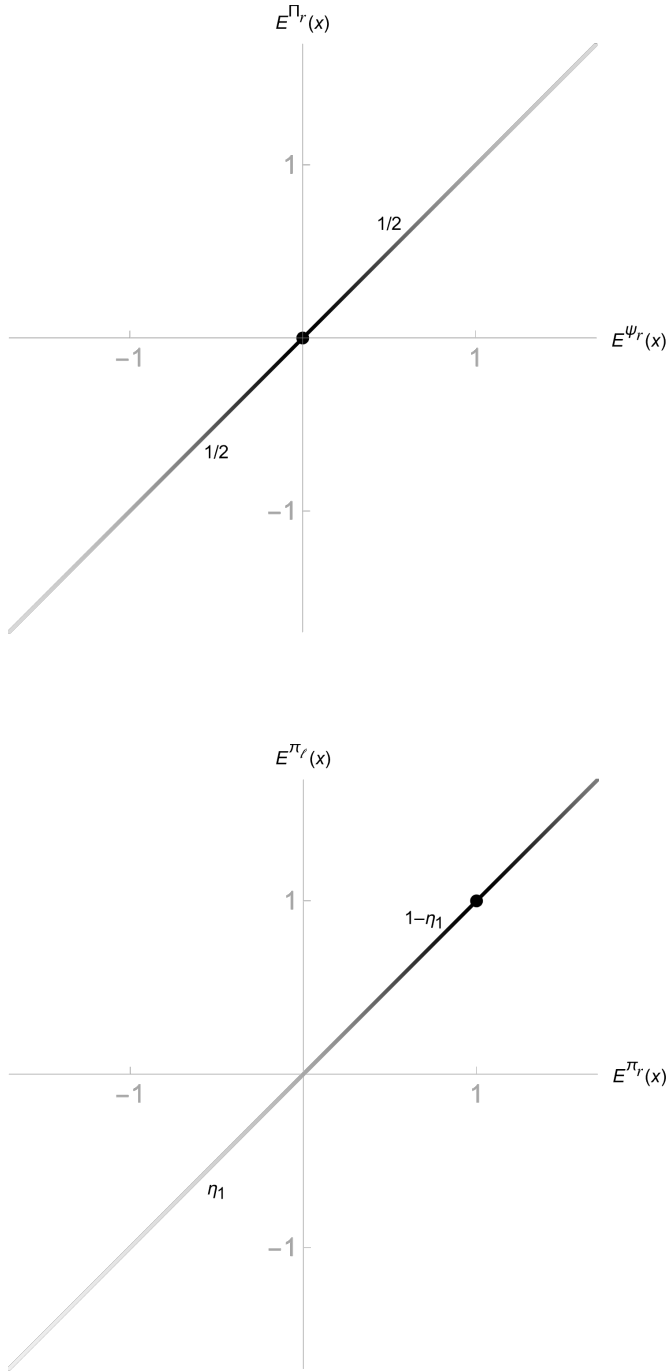



Fig. 3. Two pairs of reciprocal functions with equal bias factors within each pair. Each joint distribution is singular with support equal to the line $y = x$ (perfect correlation).

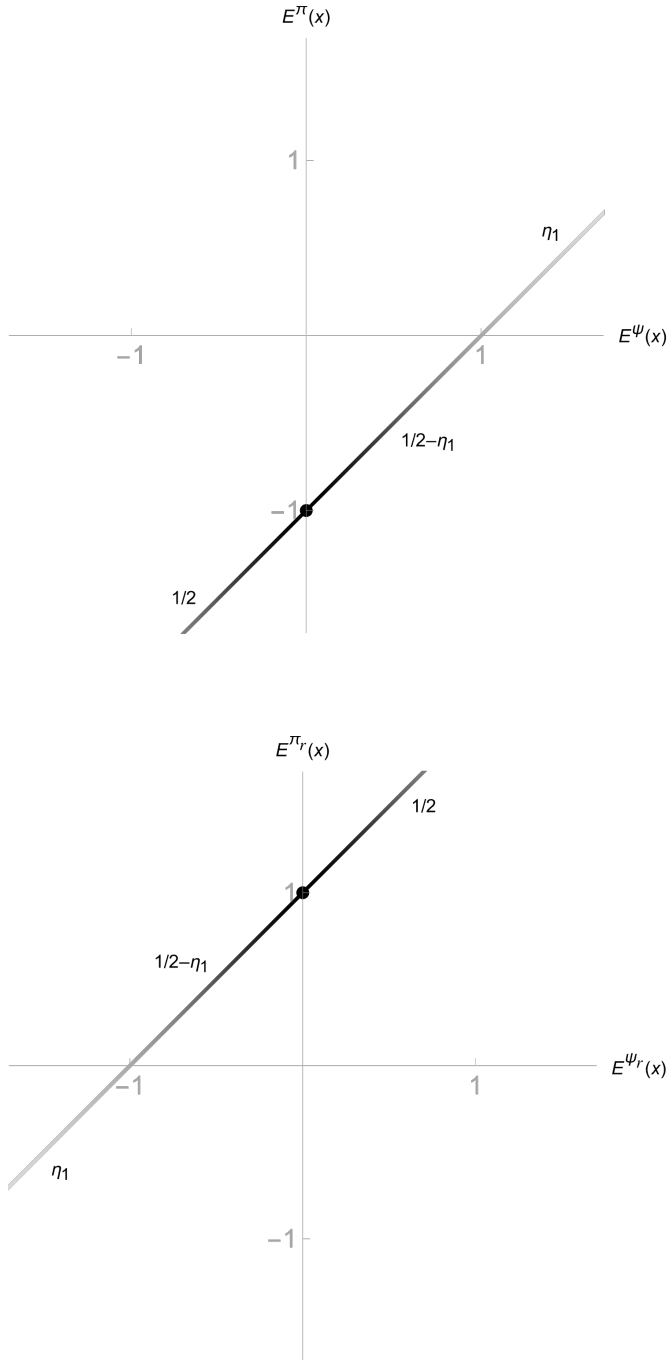



Fig. 4. A pair of standard functions and a pair of reciprocal functions, with unequal bias factors within each pair. Each joint distribution is singular with support equal to a line $y = x \pm 1$ (perfect correlation up to the bias factors).

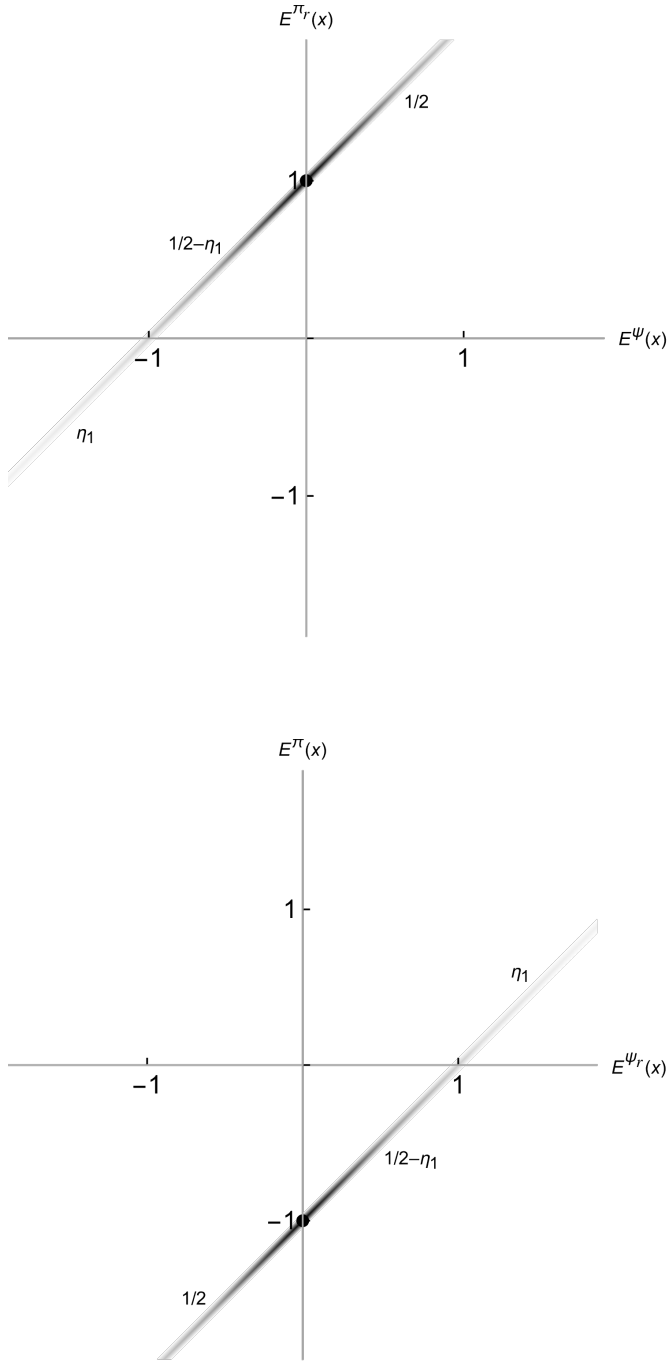



Fig. 5. Two pairs consisting of one standard function and one reciprocal function, with unequal bias factors within each pair. Each joint distribution is absolutely continuous and supported on a thin diagonal strip of (horizontal or vertical) width $2w$.

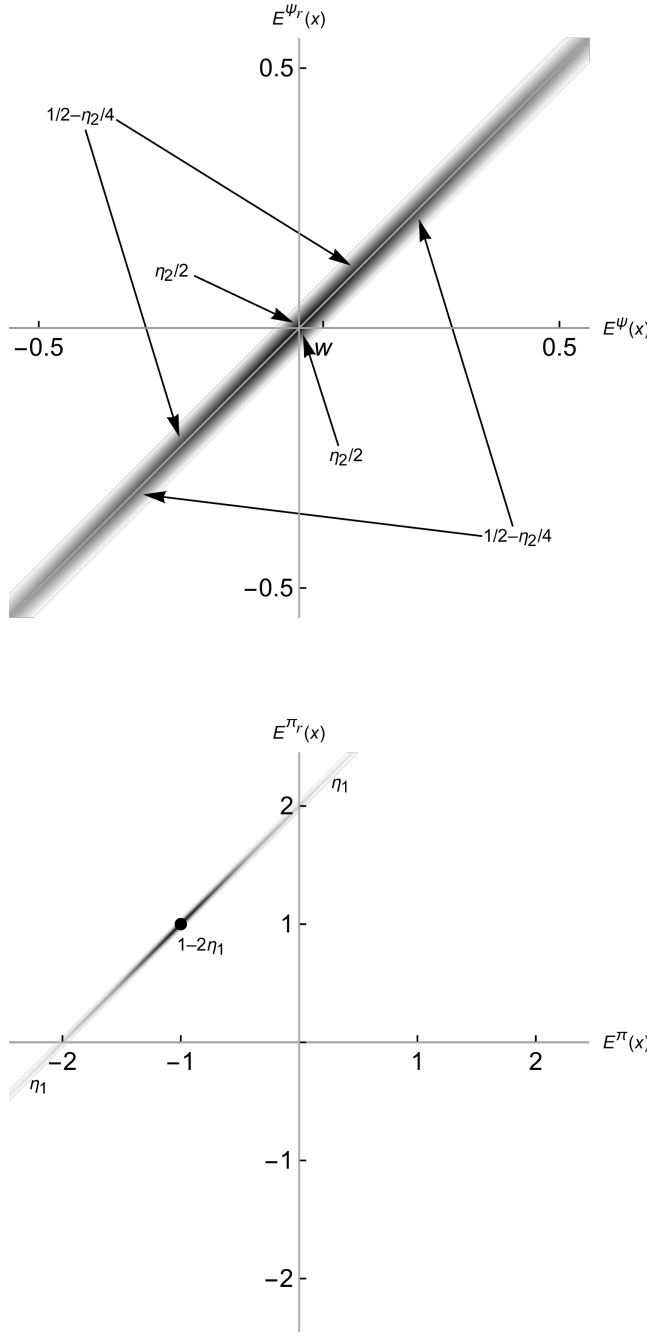



Fig. 6. Two pairs consisting of one standard function and one reciprocal function. Each joint distribution is absolutely continuous and supported on a thin diagonal strip of (horizontal or vertical) width $2w$. In the top graph, both bias factors equal 0; in the bottom graph, the bias factors are nonzero with opposite signs.

**Acknowledgements.** The authors thank the anonymous referees for their helpful comments. The second author was supported in part by a Natural Sciences and Research Council of Canada Discovery Grant.

**References**

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series 55, U. S. Government Printing Office, Washington, DC, 1964.

[2] A. Akbary, N. Ng, and M. Shahabi, *Limiting distributions of the classical error terms of prime number theory*, Quart. J. Math. 65 (2014), 743–780.

[3] C. Bays and R. H. Hudson, *A new bound for the smallest $x$ with $\pi(x) > \mathrm{li}(x)$*, Math. Comp. 69 (2000), 1285–1296.

[4] J. Büthe, *On the first sign change in Mertens' theorem*, Acta Arith. 171 (2015), 183–195.

[5] P. Dusart, *Inégalités explicites pour $\psi(X)$, $\theta(X)$, $\pi(X)$ et les nombres premiers*, C. R. Math. Acad. Sci. Soc. R. Canada 21 (1999), 53–59.

[6] A. Feuerverger and G. Martin, *Biases in the Shanks–Rényi prime number race*, Experiment. Math. 9 (2000), 535–570.

[7] J. Kaczorowski, *A note on primes in reciprocal arithmetic progressions*, Funct. Approx. Comment. Math. 25 (1997), 151–156.

[8] Y. Lamzouri, *A bias in Mertens' product formula*, Int. J. Number Theory 12 (2016), 97–109.

[9] J. P. S. Lay, *Sign changes in Mertens' first and second theorems*, arXiv:1505.03589v1 (2015).

[10] J. E. Littlewood, *Sur la distribution des nombres premiers*, C. R. Acad. Sci. Paris 158 (1914), 1869–1872.

[11] G. Martin and N. Ng, *Inclusive prime number races*, Trans. Amer. Math. Soc. 373 (2020), 3561–3607.

[12] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge Stud. Adv. Math. 97, Cambridge Univ. Press, 2006.

[13] D. J. Platt and T. S. Trudgian, *On the first sign change of $\theta(x) - x$*, Math. Comp. 85 (2016), 1539–1547.

[14] G. Robin, *Sur l'ordre maximum de la fonction somme des diviseurs*, in: Seminar on Number Theory, Paris 1981–82, Progr. Math. 38, Birkhäuser Boston, Boston, MA, 1983, 233–244.

[15] M. O. Rubinstein and P. Sarnak, *Chebyshev's bias*, Experiment. Math. 3 (1994), 173–197.

[16] G. Szegő, *Orthogonal Polynomials*, 4th ed., Amer. Math. Soc. Colloq. Publ. 23, Amer. Math. Soc., Providence, RI, 1975.

[17] The PARI Group, *PARI/GP version 2.17.3*, 2025, https://pari.math.u-bordeaux.fr/.

[18] A. Wintner, *On the asymptotic distribution of the remainder term of the prime-number theorem*, Amer. J. Math. 57 (1935), 534–538.

Shubhrajit Bhattacharya
Department of Mathematics
University of Chicago
Chicago, IL 60637, USA
E-mail: shubhrajit@uchicago.edu

Greg Martin, Reginald M. Simpson
Department of Mathematics
University of British Columbia
Vancouver, BC, V6T 1Z2, Canada
E-mail: gerg@math.ubc.ca, rs@mailc.net