Exact formulas for partial sums of the Möbius function expressed by partial sums weighted by the Liouville lambda function

Maxie Dion Schmidt
Georgia Institute of Technology
School of Mathematics

Abstract

The Mertens function, $M(x) := \sum_{n \leq x} \mu(n)$, is defined as the summatory function of the classical Möbius function. The Dirichlet inverse function $g(n) := (\omega + \mathbb{1})^{-1}(n)$ is defined in terms of the shifted strongly additive function $\omega(n)$ that counts the number of distinct prime factors of n without multiplicity. The Dirichlet generating function (DGF) of $g(n)$ is $\zeta(s)^{-1}(1+P(s))^{-1}$ for $\operatorname{Re}(s) > 1$ where $P(s) = \sum_p p^{-s}$ is the prime zeta function. We study the distribution of the unsigned functions $|g(n)|$ with DGF $\zeta(2s)^{-1}(1-P(s))^{-1}$ and $C_{\Omega}(n)$ with DGF $(1-P(s))^{-1}$ for $\operatorname{Re}(s) > 1$. We prove formulas for the average order and variance of $\log C_{\Omega}(n)$ and prove a central limit theorem for the distribution of its values over $n \leq x$ as $x \to \infty$. Discrete convolutions of the partial sums of $g(n)$ with the prime counting function provide new exact formulas for $M(x)$ that are sums of the Liouville function weighted by the unsigned summands $|g(n)|$.

Keywords and Phrases: *Möbius function; Mertens function; Liouville lambda function; prime omega function; Dirichlet inverse; prime zeta function; inversion of generalized convolutions.*
Math Subject Classifications (2010): *11N37; 11A25; 11N60; and 11N64.*

Article Index

1 Introduction **3**
1.1 Definitions 3
1.2 Statements of the main results 4

2 The function $C_{\Omega}(n)$ **5**
2.1 Definitions 5
2.2 Logarithmic variance 5
2.3 Remarks 6

3 The function $g(n)$ **7**
3.1 Definitions 7
3.2 Signedness 7
3.3 Relations to the function $C_{\Omega}(n)$ 8

4 The distribution of the function $C_{\Omega}(n)$ **9**
4.1 Proof of Theorem 1.8 9
4.2 Motivation 11

5 Applications to the Mertens function **11**
5.1 Proofs of the new formulas 12
5.2 Discrete plots and numerical experiments 13
5.3 Local cancellation in the formulas involving the partial sums of $g(n)$ 14

6 Conclusions **15**
6.1 Summary 15
6.2 Discussion of the new results 15

Acknowledgements **16**

References **16**

Appendices on supplementary material

A The distributions of $\omega(n)$ and $\Omega(n)$ **18**

B The upper incomplete gamma function **19**

C Inversion of partial sums of Dirichlet convolutions **21**

D The proof of Theorem 1.6 **22**

1 Introduction

1.1 Definitions

For integers $n \geq 2$, we define the strongly and completely additive functions, respectively, that count the number of prime divisors of n by

$$\omega(n) = \sum_{p|n} 1, \text{ and } \Omega(n) = \sum_{p^\alpha || n} \alpha.$$

We adopt the convention that the functions $\omega(1) = \Omega(1) = 0$. The Möbius function is defined as the multiplicative function [23, A008683]

$$\mu(n) = \begin{cases} 1, & \text{if } n = 1; \\ (-1)^{\omega(n)}, & \text{if } n \geq 2 \text{ and } \omega(n) = \Omega(n) \text{ (i.e., if } n \text{ is squarefree);} \\ 0, & \text{otherwise.} \end{cases}$$

The Mertens function is defined by the partial sums [23, A002321]

$$M(x) = \sum_{n \leq x} \mu(n), \text{ for } x \geq 1. \tag{1.1}$$

The Liouville lambda function is the completely multiplicative function defined for all $n \geq 1$ by $\lambda(n) := (-1)^{\Omega(n)}$ [23, A008836]. The partial sums of this function are defined by [23, A002819]

$$L(x) := \sum_{n \leq x} \lambda(n), \text{ for } x \geq 1. \tag{1.2}$$

Definition 1.1. For any arithmetic functions f and h, we define their Dirichlet convolution at n by the divisor sum

$$(f * h)(n) := \sum_{d|n} f(d) h\left(\frac{n}{d}\right), \text{ for } n \geq 1.$$

The arithmetic function f has a unique inverse with respect to Dirichlet convolution, denoted by f^{-1}, if and only if $f(1) \neq 0$. When it exists, the Dirichlet inverse of f satisfies $(f * f^{-1})(n) = (f^{-1} * f)(n) = \delta_{n,1}$.

We define the Dirichlet inverse function [23, A341444]

$$g(n) := (\omega + \mathbb{1})^{-1}(n), \text{ for } n \geq 1. \tag{1.3}$$

Th inverse function in equation (1.3) is computed recursively by applying the formula [1, §2.7]

$$g(n) = \begin{cases} 1, & \text{if } n = 1; \\ -\sum\limits_{\substack{d|n \\ d>1}} (\omega(d) + 1)\, g\left(\frac{n}{d}\right), & \text{if } n \geq 2. \end{cases}$$

The function $|g(n)| = \lambda(n) g(n)$ denotes the absolute value of $g(n)$ (see Proposition 3.3). The summatory function of $g(n)$ is defined as follows [23, A341472]:

$$G(x) := \sum_{n \leq x} g(n) = \sum_{n \leq x} \lambda(n) |g(n)|, \text{ for } x \geq 1. \tag{1.4}$$

1.2 Statements of the main results

Definition 1.2. Let the partial sums of the Dirichlet convolution $r * h$ be defined by the function

$$S_{r*h}(x) := \sum_{n \leq x} \sum_{d|n} r(d) h\left(\frac{n}{d}\right), \text{ for } x \geq 1.$$

The next theorem is proved by matrix methods in Appendix C.

Theorem 1.3. *Let $r, h : \mathbb{Z}^+ \to \mathbb{C}$ be any arithmetic functions such that $r(1) \neq 0$. Suppose that $R(x) := \sum_{n \leq x} r(n)$, $H(x) := \sum_{n \leq x} h(n)$, and that $R^{-1}(x) := \sum_{n \leq x} r^{-1}(n)$ for $x \geq 1$. The following formulas hold for all integers $x \geq 1$:*

$$\begin{aligned} S_{r*h}(x) &= \sum_{d=1}^{x} r(d) \times H\left(\left\lfloor \frac{x}{d} \right\rfloor\right) \\ S_{r*h}(x) &= \sum_{k=1}^{x} H(k) \times \left(R\left(\left\lfloor \frac{x}{k} \right\rfloor\right) - R\left(\left\lfloor \frac{x}{k+1} \right\rfloor\right)\right). \end{aligned}$$

Moreover, for any $x \geq 1$

$$\begin{aligned} H(x) &= \sum_{j=1}^{x} S_{r*h}(j) \times \left(R^{-1}\left(\left\lfloor \frac{x}{j} \right\rfloor\right) - R^{-1}\left(\left\lfloor \frac{x}{j+1} \right\rfloor\right)\right) \\ &= \sum_{k=1}^{x} r^{-1}(k) \times S_{r*h}(x). \end{aligned}$$

For integers $x \geq 1$, the function $\pi(x) := \sum_{p \leq x} 1$ is the classical prime counting function [23, A000720]. We find new exact formulas for $M(x)$ by applying Theorem 1.3 to the expansion of the partial sums (see Remark 3.2)

$$\pi(x) + 1 = \sum_{n \leq x} \sum_{d|n} (\omega(d) + 1) \mu\left(\frac{n}{d}\right), \text{ for } x \geq 1.$$

Theorem 1.4. *For all $x \geq 1$*

$$M(x) = G(x) + \sum_{1 \leq k \leq x} |g(k)| \pi\left(\left\lfloor \frac{x}{k} \right\rfloor\right) \lambda(k), \tag{1.5a}$$

$$M(x) = G(x) + \sum_{1 \leq k \leq \frac{x}{2}} \left(\pi\left(\left\lfloor \frac{x}{k} \right\rfloor\right) - \pi\left(\left\lfloor \frac{x}{k+1} \right\rfloor\right)\right) G(k), \tag{1.5b}$$

$$M(x) = G(x) + \sum_{p \leq x} G\left(\left\lfloor \frac{x}{p} \right\rfloor\right). \tag{1.5c}$$

The auxiliary unsigned function $C_\Omega(n)$ studied by Fröberg [11] has an exact formula given by

$$C_\Omega(n) = \begin{cases} 1, & \text{if } n = 1; \\ (\Omega(n))! \times \prod_{p^\alpha || n} \frac{1}{\alpha!}, & \text{if } n \geq 2. \end{cases} \tag{1.6}$$

Proposition 1.5. *For all $n \geq 1$*

$$|g(n)| = \sum_{d|n} \mu^2\left(\frac{n}{d}\right) C_\Omega(d). \tag{1.7}$$

Theorem 1.6. *As $n \to \infty$*

$$\frac{1}{n} \times \sum_{k \leq n} \log C_\Omega(k) = (\log\log n)(\log\log\log n)\left(1 + O\left(\frac{1}{(\log\log n)^{\frac{1}{3}}}\right)\right).$$

A proof of Theorem 1.6 is given in Appendix D.

Definition 1.7. The cumulative density function of the standard normal distribution is

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \times \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt, \text{ for } z \in (-\infty, \infty).$$

Theorem 1.8. *For $x \geq 19$, let $\mu_x, \sigma_x := (\log\log x)(\log\log\log x)$. For any $z \in (-\infty, \infty)$*

$$\lim_{x\to\infty} \frac{1}{x} \times \#\left\{19 \leq n \leq x : \frac{\log C_\Omega(n) - \mu_x}{\sigma_x} \leq z\right\} = \Phi(z).$$

2 The function $C_\Omega(n)$

In this section, we define the function $C_\Omega(n)$ and explore its properties. The function $C_\Omega(n)$ is key to understanding the unsigned inverse sequence $|g(n)|$ in terms of equation (1.7).

2.1 Definitions

Definition 2.1. We define the following bivariate sequence for integers $n \geq 1$ and $k \geq 0$:

$$C_k(n) := \begin{cases} \varepsilon(n), & \text{if } k = 0; \\ \sum\limits_{d|n} \omega(d) C_{k-1}\left(\frac{n}{d}\right), & \text{if } k \geq 1. \end{cases} \tag{2.1}$$

Using the notation for iterated convolution in Bateman and Diamond [3, Def. 2.3; §2], we have $C_0(n) \equiv \omega^{*0}(n)$ and $C_k(n) \equiv \omega^{*k}(n)$ for integers $n, k \geq 1$. The special case of (2.1) where $k := \Omega(n)$ occurs frequently in the next sections of the article. To avoid cumbersome notation when referring to this common function variant, we suppress the duplicate index n by writing $C_\Omega(n) := C_{\Omega(n)}(n)$ [23, A008480].

Remark 2.2. By recursively expanding the definition of $C_k(n)$ at any fixed $n \geq 2$, we see that we can form a chain of at most $\Omega(n)$ iterated (or nested) divisor sums by unfolding the definition of (2.1) inductively. We also see that at fixed n, the function $C_k(n)$ is non-zero only possibly for $1 \leq k \leq \Omega(n)$ when $n \geq 2$. By equation (1.6) we have that $C_\Omega(n) \leq (\Omega(n))!$ for all $n \geq 1$ with equality precisely at the squarefree integers so that $(\Omega(n))! = (\omega(n))!$ if and only if $\mu^2(n) = 1$.

2.2 Logarithmic variance

Definition 2.3. For any integers $x \geq 1$, we define the expectation (or mean value) of the function $\log C_\Omega(n)$ on the integers $1 \leq n \leq x$ by

$$\mathbb{E}\left[\log C_\Omega(x)\right] := \frac{1}{x} \times \sum_{n \leq x} \log C_\Omega(n).$$

The *variance* of this function is given by the centralized second moments

$$\mathrm{Var}\left[\log C_\Omega(x)\right] := \frac{1}{x} \times \sum_{n \leq x} \left(\log C_\Omega(n) - \mathbb{E}\left[\log C_\Omega(x)\right]\right)^2.$$

Proposition 2.4. *For $n > e^e$, the variance of the function $\log C_\Omega(n)$ is*

$$\sqrt{\mathrm{Var}\left[\log C_\Omega(n)\right]} \sim (\log\log n)(\log\log\log n), \text{ as } n \to \infty.$$

Proof. Suppose that $n \geq 16$. We have that for all $n \geq 1$

$$S_{2,\Omega}(n) := \sum_{k \leq n} \log^2 C_\Omega(k) - \left(\sum_{k \leq n} \log C_\Omega(k)\right)^2 = \sum_{1 \leq j < k \leq n} 2 \log C_\Omega(j) \log C_\Omega(k). \tag{2.2}$$

For integers $n \geq 1$, we define the sums

$$E_\Omega(n) := \frac{1}{n} \times \sum_{k \leq n} \log C_\Omega(k), \text{ and } V_\Omega^2(n) := \frac{1}{n} \times \sum_{k \leq n} \log^2 C_\Omega(k)$$

The expansion on the right-hand-side of (2.2) is rewritten as

$$\frac{S_{2,\Omega}(n)}{n^2} = V_\Omega^2(n) - E_\Omega^2(n) = \sum_{1 \leq j < n} 2 \log C_\Omega(j) \left(E_\Omega(n) - E_\Omega(j) \right). \tag{2.3}$$

Equation (2.3) implies that as $n \to \infty$

$$\begin{aligned} V_\Omega^2(n) &\sim 3E_\Omega^2(n) - 2(\log\log n)^2(\log\log\log n)^2 + I_A(n), \\ &\sim (\log\log n)^2(\log\log\log n)^2 + I_A(n). \end{aligned} \tag{2.4}$$

The integral term in the last equations is defined by

$$I_A(x) := 2 \times \int_{e^e}^x t(\log\log t)^2(\log\log\log t)^2 dt.$$

For $x > e^e$, we can integrate exactly to find

$$\int_{e^e}^x \frac{(\log\log t)^2(\log\log\log t)^2}{t(\log t)} dt \sim \frac{1}{3}(\log\log x)^3(\log\log\log x)^3, \text{ as } x \to \infty.$$

The mean value theorem shows that there is a bounded constant $c \equiv c(x) \in [e^e, x]$ such that

$$I_A(x) \sim \frac{2}{3} c(x) \log c(x) (\log\log x)^3 (\log\log\log x)^3.$$

For $x, y \in [0, \infty)$, the function $W(y)$ denotes the principal branch of the multi-valued Lambert W-function on the non-negative reals defined such that $x = W(y)$ if and only if $xe^x = y$. We can differentiate the previous equation and discard the lower order terms to solve for the main term of $c(x)$ as $x \to \infty$ to see that

$$c(x) \ll \frac{\log\log\log\log x}{W(\log\log\log\log x)} \ll \frac{\log\log\log\log x}{\log\log\log\log\log x}.$$

The last equation implies that $I_A(x) = o\left(E_\Omega(x)\right)$. The conclusion then follows from equation (2.4). $\square$

2.3 Remarks

Asymptotic formulae for the moments of the unscaled function $C_\Omega(n)$ on the positive integers $n \leq x$ as $x \to \infty$ is required to evaluate the average order (and higher-order moment statistics) of $|g(n)|$. The sign weight of $\lambda(n)$ on $g(n) = \lambda(n)|g(n)|$ in partial sums of this function is an unknown in bounding $M(x)$ by the formulas in Theorem 1.4. Proofs to evaluate the centralized moments of the former function are not immediate using the elementary methods used to establish Theorem 1.6 and Proposition 2.4.

Let the parameters $k \in \mathbb{Z}^+$ and $z \in \mathbb{C}$ subject to $1 \leq k \leq R \log\log x$ and $|z| \leq M$ for some bounded finite $0 < M < +\infty$ when $R \in (1, 2)$ be fixed. An approach to the average order of $C_\Omega(n)$ invokes the Selberg-Delange method [26, §II.6.1] [18, §7.4] in evaluating the partial sums

$$\sum_{\substack{n \leq x \\ \Omega(n) = k}} \frac{(-1)^{\omega(n)} C_\Omega(n) z^{2\Omega(n)}}{(\Omega(n))!}; \text{ and } \sum_{\substack{n \leq x \\ \Omega(n) = k}} \frac{(-1)^{\omega(n)} C_\Omega(n)}{(\Omega(n))!}$$

We can extract the coefficients of $z^{2\Omega(n)}$ from the expansions of the DGF

$$\sum_{n \geq 1} \frac{C_\Omega(n)}{(\Omega(n))!} \cdot \frac{(-1)^{\omega(n)} z^{\Omega(n)}}{n^s} = \prod_p \left(1 + \sum_{r \geq 1} \frac{z^{\Omega(p^r)}}{r! p^{rs}} \right)^{-1} = \exp\left(-zP(s)\right), \text{ for } \mathrm{Re}(s) > 1.$$

A proof of the corresponding first moment formula for the function $(-1)^{\omega(n)}C_\Omega(n)$ requires long and technical arguments that are non-trivial extensions of the model proofs in [18, 26]. Integration by parts and the mean value theorem applied to the signed sums in Lemma B.3 yield exact asymptotic formulae for the partial sums of the function $C_\Omega(n)$ over the $n \leq x$ such that $\Omega(n) = k$ when $1 \leq k \leq R \log\log x$.

3 The function $g(n)$

3.1 Definitions

Definition 3.1. For integers $n \geq 1$, we define the Dirichlet inverse function taken with respect to the operation of Dirichlet convolution to be

$$g(n) = (\omega + \mathbb{1})^{-1}(n), \text{ for } n \geq 1.$$

The function $|g(n)|$ denotes the unsigned inverse function, or equivalently the absolute value of $g(n)$, for integers $n \geq 1$.

Remark 3.2. Let $\chi_{\mathbb{P}}(n)$ denote the characteristic function of the primes, suppose that $\varepsilon(n) = \delta_{n,1}$ be the multiplicative identity with respect to Dirichlet convolution, the function $\mathbb{1}(n)$ be identically equal to one for all $n \geq 1$. We can see using elementary methods that

$$\chi_{\mathbb{P}} + \varepsilon = (\omega + \mathbb{1}) * \mu. \tag{3.1}$$

The result in (3.1) follows by Möbius inversion since $\mu * \mathbb{1} = \varepsilon$ and

$$\omega(n) = \sum_{d|n} \chi_{\mathbb{P}}(d), \text{ for } n \geq 1.$$

We recall the following statement of the inversion theorem of summatory functions for any Dirichlet invertible arithmetic function $\alpha(n)$ proved in [1, §2.14]:

$$G(x) = \sum_{n \leq x} \alpha(n) F\left(\frac{x}{n}\right) \implies F(x) = \sum_{n \leq x} \alpha^{-1}(n) G\left(\frac{x}{n}\right), \text{ for } x \geq 1. \tag{3.2}$$

Hence, to express the new formulas for $M(x)$ we may consider the inversion of the right-hand-side of the partial sums

$$\pi(x) + 1 = \sum_{n \leq x} (\chi_{\mathbb{P}} + \varepsilon)(n) = \sum_{n \leq x} \sum_{d|n} (\omega(d) + 1)\mu\left(\frac{n}{d}\right), \text{ for } x \geq 1.$$

3.2 Signedness

Proposition 3.3. *The sign of the function $g(n)$ is $\lambda(n)$ for all $n \geq 1$.*

Proof. The series $D_f(s) := \sum_{n \geq 1} f(n) n^{-s}$ defines the Dirichlet generating function (DGF) of any arithmetic function f which is convergent for all $s \in \mathbb{C}$ satisfying $\operatorname{Re}(s) > \sigma_f$ where σ_f is the abscissa of convergence of the series. Recall that $D_{\mathbb{1}}(s) = \zeta(s)$, $D_\mu(s) = \zeta(s)^{-1}$ and $D_\omega(s) = P(s)\zeta(s)$ for $\operatorname{Re}(s) > 1$. Whenever $f(1) \neq 0$ the DGF of $f^{-1}(n)$ is $D_f(s)^{-1}$. By equation (3.1) we have

$$D_{(\omega+1)^{-1}}(s) = \frac{1}{\zeta(s)(1 + P(s))}, \text{ for } \operatorname{Re}(s) > 1. \tag{3.3}$$

It follows that $(\omega + \mathbb{1})^{-1}(n) = (h^{-1} * \mu)(n)$ for $h := \chi_{\mathbb{P}} + \varepsilon$. We first show that $\operatorname{sgn}(h^{-1}) = \lambda$ which implies that $\operatorname{sgn}(h^{-1} * \mu) = \lambda$.

We recover exactly that [11, cf. §2]

$$h^{-1}(n) = \begin{cases} 1, & \text{if } n = 1; \\ \lambda(n)(\Omega(n))! \times \prod_{p^\alpha \| n} \frac{1}{\alpha!}, & \text{if } n \geq 2. \end{cases}$$

In particular, by expanding the DGF of h^{-1} formally in powers of $P(s)$, where $|P(s)| < 1$ whenever $\operatorname{Re}(s) \geq 2$, we count that

$$\begin{aligned} \frac{1}{1+P(s)} &= \sum_{n\geq 1} \frac{h^{-1}(n)}{n^s} = \sum_{k\geq 0} (-1)^k P(s)^k \\ &= 1 + \sum_{\substack{n\geq 2 \\ n=p_1^{\alpha_1} p_2^{\alpha_2} \times \cdots \times p_k^{\alpha_k}}} \frac{(-1)^{\alpha_1+\alpha_2+\cdots+\alpha_k}}{n^s} \times \binom{\alpha_1+\alpha_2+\cdots+\alpha_k}{\alpha_1, \alpha_2, \ldots, \alpha_k} \\ &= 1 + \sum_{\substack{n\geq 2 \\ n=p_1^{\alpha_1} p_2^{\alpha_2} \times \cdots \times p_k^{\alpha_k}}} \frac{\lambda(n)}{n^s} \times \binom{\Omega(n)}{\alpha_1, \alpha_2, \ldots, \alpha_k}. \end{aligned} \tag{3.4}$$

Since λ is completely multiplicative we have that $\lambda\left(\frac{n}{d}\right)\lambda(d) = \lambda(n)$ for all divisors $d|n$ when $n \geq 1$. We also have that $\mu(n) = \lambda(n)$ whenever n is squarefree so that

$$g(n) = (h^{-1} * \mu)(n) = \lambda(n) \times \sum_{d|n} \mu^2\left(\frac{n}{d}\right) |h^{-1}(n)|, \text{ for } n \geq 1. \qquad \square$$

The function $|h^{-1}(n)|$ from the last proof identically matches values of $C_\Omega(n)$ at all $n \geq 1$. The proof shows that the sequence $\lambda(n)C_\Omega(n)$ has DGF of $(1+P(s))^{-1}$ for all $\operatorname{Re}(s) > 1$. We can easily extend the last proof to see that $C_\Omega(n)$ has DGF $(1-P(s))^{-1}$ for all $\operatorname{Re}(s) > 1$.

3.3 Relations to the function $C_\Omega(n)$

Lemma 3.4. *For all* $n \geq 1$

$$g(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) \lambda(d) C_\Omega(d).$$

Proof. We expand the recurrence relation for the Dirichlet inverse with $g(1) = 1$ as

$$g(n) = -\sum_{\substack{d|n \\ d>1}} (\omega(d)+1) g\left(\frac{n}{d}\right) \quad \Longrightarrow \quad (g * 1)(n) = -(\omega * g)(n). \tag{3.5}$$

For $1 \leq m \leq \Omega(n)$, we can inductively expand the implication on the right-hand-side of (3.5) in the form of $(g * 1)(n) = F_m(n)$ where $F_m(n) := (-1)^m (C_m(-) * g)(n)$ as

$$F_m(n) = -\begin{cases} (\omega * g)(n), & m = 1; \\ \sum_{\substack{d|n \\ d>1}} F_{m-1}(d) \times \sum_{\substack{r|\frac{n}{d} \\ r>1}} \omega(r) g\left(\frac{n}{dr}\right), & 2 \leq m \leq \Omega(n); \\ 0, & \text{otherwise.} \end{cases}$$

When $n \geq 2$ and $m := \Omega(n)$, i.e., with the expansions in the previous equation taken to a maximal depth, we obtain

$$(g * 1)(n) = \lambda(n) C_\Omega(n). \tag{3.6}$$

The formula follows from equation (3.6) by Möbius inversion. $\square$

Proof of Proposition 1.5. The result follows from Lemma 3.4, Proposition 3.3 and the complete multiplicativity of $\lambda(n)$. Since $\mu(n)$ is non-zero only at squarefree integers and since at any squarefree $d \geq 1$ we have $\mu(d) = (-1)^{\omega(d)} = \lambda(d)$, we have

$$\begin{aligned} |g(n)| &= \lambda(n) \times \sum_{d|n} \mu\left(\frac{n}{d}\right) \lambda(d) C_{\Omega}(d) \\ &= \lambda(n^2) \times \sum_{d|n} \mu^2\left(\frac{n}{d}\right) C_{\Omega}(d). \end{aligned}$$

The leading term $\lambda(n^2) = 1$ for all $n \geq 1$ since the number of distinct prime factors (counting multiplicity) of any square integer is even. □

Remark 3.5. We have the following remarks on consequences of Corollary 1.5:

- Whenever $n \geq 1$ is squarefree
$$|g(n)| = \sum_{d|n} C_{\Omega}(d). \tag{3.7a}$$
Since all divisors of a squarefree integer are squarefree, for all squarefree integers $n \geq 1$, we have that
$$|g(n)| = \sum_{m=0}^{\omega(n)} \binom{\omega(n)}{m} \times m!. \tag{3.7b}$$
- The formula in (1.7) shows that the DGF of the unsigned inverse function $|g(n)|$ is given by the meromorphic function $\zeta(2s)^{-1}(1 - P(s))^{-1}$ for all $s \in \mathbb{C}$ with $\operatorname{Re}(s) > 1$. This DGF has a pole to the right of the line at $\operatorname{Re}(s) = 1$ at the unique real $\sigma \approx 1.39943$ such that $P(\sigma) = 1$ along the reals $\sigma > 1$.

4 The distribution of the function $C_{\Omega}(n)$

In this section, we prove a central limit theorem for the function $\log C_{\Omega}(n)$. The relations between $|g(n)|$ and $C_{\Omega}(n)$ proved in the last section are suggestive of applications of the result in Theorem 1.8 to bounding the partial sums of $g(n)$.

4.1 Proof of Theorem 1.8

Proof of Theorem 1.8. We outline the next steps to complete the proof of this result:

- Given a fixed $x \geq 1$, we select another integer $N \equiv N(x)$ uniformly at random from $\{1, 2, \ldots, x\}$. For each prime p we define
$$C_p^{(x)} := \begin{cases} 0, & p \nmid N(x); \\ \alpha, & p^{\alpha} \| N(x), \text{ for some } \alpha \geq 1. \end{cases}$$
For integers $k \geq 1$ and primes p, we have limiting convergence in distribution of $C_p^{(x)} \overset{d}{\Longrightarrow} Z_p$ where Z_p is a geometric random variable with parameter p^{-1} so that [2, §1.2]
$$\lim_{x \to \infty} \mathbb{P}\left[C_p^{(x)} = k\right] = \left(1 - \frac{1}{p}\right)\left(\frac{1}{p}\right)^k.$$
- For $n \geq 1$, we use equation (1.6) and Binet's log-gamma formula [22, §5.9(i)] to show that
$$\begin{aligned} \log C_{\Omega}(n) &= \log(\Omega(n))! - \sum_{\substack{p^{\alpha} \| n \\ \alpha \geq 2}} \log(\alpha!) \\ &= \Omega(n) \log \Omega(n) - \sum_{\substack{p^{\alpha} \| n \\ \alpha \geq 2}} \alpha \log(1 + \alpha) + O(\Omega(n)). \end{aligned} \tag{4.1}$$
Since $\Omega(n) = 1$ only for n within a subset of the positive integers with asymptotic density of zero (i.e., on the primes), it suffices to restrict our considerations to the $n \geq 2$ such that $\Omega(n) \geq 2$.

- We write the expansion from equation (4.1) as the difference $\log C_\Omega(N(x)) := \Theta_{N(x)} - A_{N(x)} + O(1)$ where

$$\Theta_{N(x)} := \Omega(N(x)) \log \Omega(N(x)),$$
$$A_{N(x)} := \sum_{p \leq x} C_p^{(x)} \log C_p^{(x)} \times \mathbb{1}_{\{C_p^{(x)} \geq 2\}}(p).$$

We can show that as $x \to \infty$

$$\mathbb{E}[A_{N(x)}] \ll \sum_{p \leq x} \mathbb{E}\left[C_p^{(x)} \log C_p^{(x)}\right] \times \mathbb{P}\left[C_p^{(x)} \geq 2\right] = o\left(\mathbb{E}[\Theta_{N(x)}]\right).$$

Analogous bounds can be proved to relate the variance of these two random variables as $x \to \infty$.
- Let $\mu_x := \mathbb{E}[\log C_\Omega(x)]$ and $\sigma_x^2 := \mathrm{Var}[\log C_\Omega(x)]$ be defined as in Definition 2.3. For $1 \leq n \leq x$, let the indicator random variable $\chi_{x,n}$ be defined as follows: $\chi_{x,n} := \mathbb{1}_{\{N(x)=n\}}$. For $x \geq 1$, let the random variables

$$S_x := \sum_{1 \leq n \leq x} \log C_\Omega(n) \chi_{x,n}.$$

We can calculate that for all integers $1 \leq n \leq x$

$$\mathbb{E}[S_x] = \mu_x; \mathrm{Var}[S_x] = \sigma_x^2; \text{ and } \hat{\mu}_{x,n} := \mathbb{E}[\log C_\Omega(n) \cdot \chi_{x,n}] = \frac{1}{x} \times \log C_\Omega(n).$$

- For fixed $\varepsilon > 0$ and large x, let

$$\widetilde{E}_\Omega(\varepsilon, x) := \frac{1}{\sigma_x^2} \times \sum_{1 \leq n \leq x} \mathbb{E}\left[(\log C_\Omega(n)\chi_{x,n} - \hat{\mu}_{x,n})^2 \times \mathbb{1}_{\{|\log C_\Omega(n)\chi_{x,n} - \hat{\mu}_{x,n}| > \varepsilon\sigma_x\}}\right].$$

The Lindeberg condition is satisfied when the following is true for every $\varepsilon > 0$:

$$\lim_{x \to \infty} \widetilde{E}_\Omega(\varepsilon, x) = 0. \tag{4.2}$$

Whenever equation (4.2) holds for all $\varepsilon > 0$, we can apply the Lindeberg central limit theorem (CLT). This CLT result shows using Theorem 1.6 and Proposition 2.4 that we have convergence in distribution to a standard normal random variable as follows [5, §27]:

$$\lim_{x \to \infty} \mathbb{P}\left[\frac{S_x - \mu_x}{\sigma_x} \leq z\right] = \Phi(z), \text{ for } z \in (-\infty, \infty). \tag{4.3}$$

- The analog of the Erdős-Kac theorem for the function $\Omega(n)$ is given by [18, Thm. 7.21; §7.4]

$$\frac{1}{x} \times \#\left\{3 \leq n \leq x : \frac{\Omega(n) - \log\log x}{\sqrt{\log\log x}} \leq z\right\} = \Phi(z) + O\left(\frac{1}{\sqrt{\log\log x}}\right), \text{ for } z \in (-\infty, \infty).$$

Therefore, for any $1 \leq k \leq \log_2(x)$

$$\mathbb{P}[\Omega(N(x)) = k] = \frac{e^{-\frac{(k - \log\log x)^2}{2\log\log x}}}{\sqrt{2\pi}} + o(1), \text{ as } x \to \infty.$$

As $x \to \infty$, the condition $k \log k > (1+\varepsilon)\mu_x$ is true when $k > \frac{(1+\varepsilon)\mu_x}{W((1+\varepsilon)\mu_x)} \sim (1+\varepsilon)\log\log x$ as $x \to \infty$. The inequality $k \log k \geq \left(k + \frac{1}{2}\right)\log(1+k) - k$ is satisfied for any real $k > 1.06975$.

- For fixed $\varepsilon > 0$, we have that

$$\begin{aligned}\widetilde{E}_\Omega(\varepsilon, x) &\ll \frac{1}{\sigma_x^2} \times \mathbb{E}\left[(S_x - \mu_x)^2 \times \mathbb{1}_{\{|S_x - \mu_x| > \varepsilon\sigma_x\}}\right] \\ &\ll \frac{1}{\sigma_x^2} \times \int_0^{\log\log x} (t - \mu_x)^2 \mathbb{P}\left[\lfloor t \rfloor \leq S_x < \lceil t \rceil\right] \times \mathbb{1}_{\{|t - \mu_x| > \varepsilon\sigma_x\}}(t) dt \\ &\ll \frac{1}{\sigma_x^2} \times \sum_{1 \leq k \leq \log_2(x)} (\log(k!) - \mu_x)^2 \, \mathbb{P}\left[\Omega(N(x)) = k\right] \times \mathbb{1}_{\{|\log(k!) - \mu_x| > \varepsilon\sigma_x\}}(k) \\ &\ll \frac{1}{\sigma_x^2 \sqrt{\log\log x}} \times \int_2^{\log_2(x)} (t \log t - \mu_x)^2 \, e^{-\frac{(t - \log\log x)^2}{2\log\log x}} \mathbb{1}_{\{t \log t > (1+\varepsilon)\sigma_x\}}(t) \qquad \left(\text{Set } v := \tfrac{t - \log\log x}{\sqrt{\log\log x}}\right) \\ &\ll \frac{(\log\log x)(\log\log\log x)^2}{\sigma_x^2} \times \int_2^\infty v^2 e^{-\frac{v^2}{2}} dv \xrightarrow{x \to \infty} 0.\end{aligned}$$

The last equations show that we obtain (4.3) from the Lindeberg CLT. □

4.2 Motivation

Remark 4.1. For $n \geq 2$, let the function $\mathcal{E}[n] := (\alpha_1, \ldots, \alpha_r)$ denote the unordered partition of exponents (r-tuple) for which $\omega(n) = r$ and $n = p_1^{\alpha_1} \times \cdots \times p_r^{\alpha_r}$ is the factorization of n into powers of distinct primes. For any $n_1, n_2 \geq 2$

$$\mathcal{E}[n_1] = \mathcal{E}[n_2] \implies C_\Omega(n_1) = C_\Omega(n_2) \text{ and } g(n_1) = g(n_2). \tag{4.4}$$

This property shows that there is a deeper structure connected to the prime divisors of the positive integers $n \geq 2$ underneath these functions than exists behind an arbitrary integer sequence. On the other hand, since the multiplicative function $\mu^2(n)$ and the strongly additive functions $\omega(n)$ and $\Omega(n)$ satisfy their analog to equation (4.4) for our auxiliary functions, this property alone is insufficient to predict the CLT proved in the last theorem.

More intuition about why the distribution of (the logarithm of) $C_\Omega(n)$ is governed by a limiting probabilistic model is heuristic:

Remark 4.2. By definition the function $C_\Omega(n)$ is identified with the $\Omega(n)$-fold Dirichlet convolution of the strongly additive $\omega(n)$ with itself via Definition 2.1. This perspective provides more insight into why we should expect to find a limiting distribution associated with the distinct values of $C_\Omega(n)$ over $n \leq x$ (pointwise) and of $\log C_\Omega(n)$ over $n \leq x$ (smoothly via Theorem 1.8). In particular, we associate the tendency of $\omega(n)$ towards its average order with the Erdős-Kac theorem (cf. Appendix A)

$$\frac{1}{x} \times \#\left\{n \leq x : \frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \leq z\right\} = \Phi(z) + o(1), \text{ for } z \in (-\infty, \infty), \text{ as } x \to \infty.$$

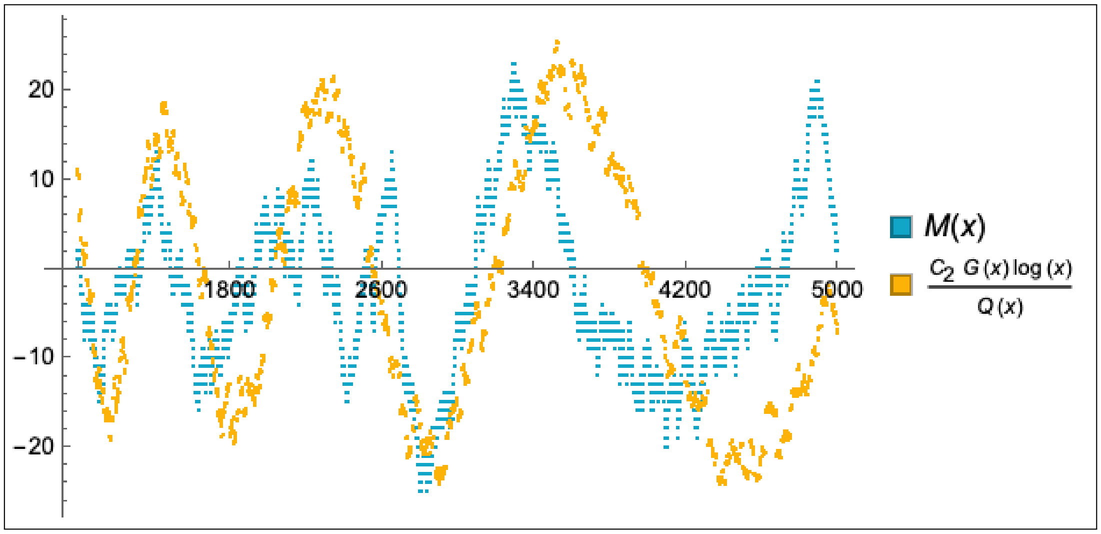
In the sense that multiple (Dirichlet) convolutions reflect a qualitative smoothing operation on average, the CLT statement underlying the distribution of $\omega(n)$ given by the Erdős-Kac theorem should predict a limiting distribution. Incidentally, equation (1.6) shows that the normalized function $\frac{C_\Omega(n)}{(\Omega(n))!}$ is multiplicative (cf. [8]).

5 Applications to the Mertens function

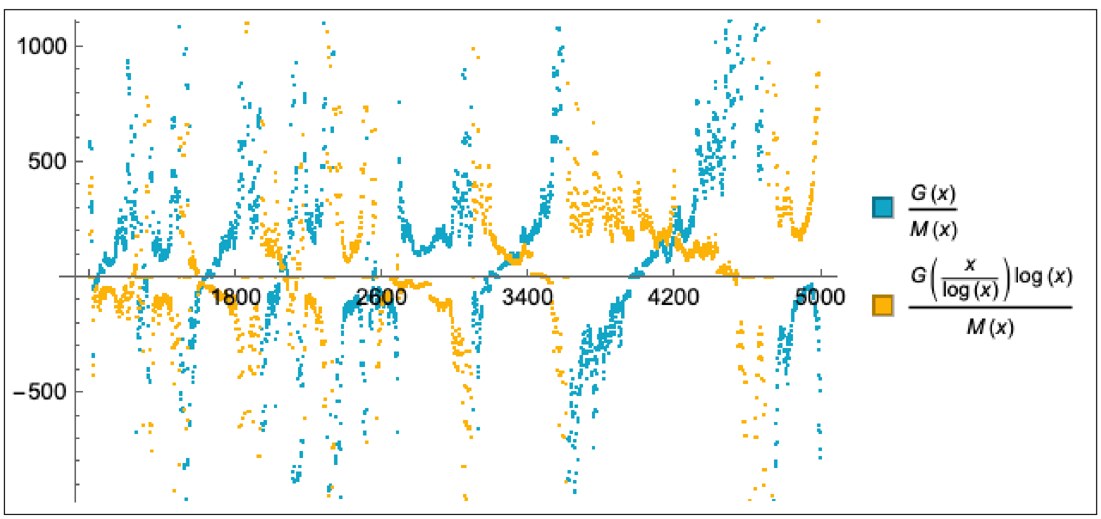
In this section, we prove the formulas for $M(x)$ involving the partial sums of the function $g(n)$ stated in Theorem 1.4. The new formulas exactly identify the Mertens function with partial sums of positive unsigned arithmetic functions whose summands are weighted by the sign of $\lambda(n)$. These formulas show that better understanding of the asymptotics of the summatory function of $g(n)$ provides insight into the behavior of $M(x)$.

Definition 5.1. The summatory functions of $g(n)$ and $|g(n)|$, respectively, are defined for all $x \geq 1$ by the partial sums

$$G(x) := \sum_{n \leq x} g(n) = \sum_{n \leq x} \lambda(n)|g(n)|, \text{ and } |G|(x) := \sum_{n \leq x} |g(n)|.$$



(a)



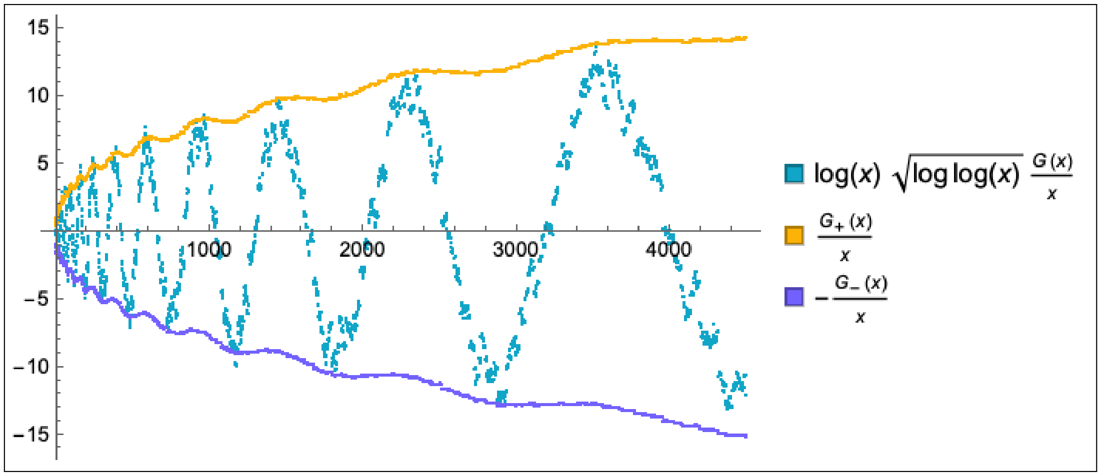
(b)

Figure 5.1

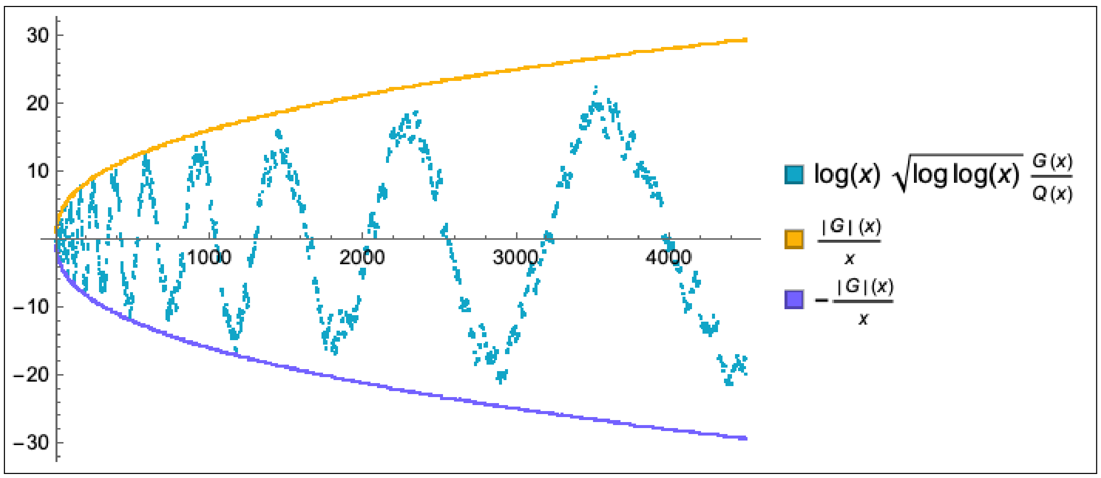
5.1 Proofs of the new formulas

Proof of (1.5a) *and* (1.5b) *of Theorem 1.4.* By applying Theorem 1.3 to equation (3.1) we have that

$$\begin{aligned}
M(x) &= \sum_{k=1}^{x} \left(\pi\left(\left\lfloor \frac{x}{k} \right\rfloor \right) + 1 \right) g(k) \\
&= G(x) + \sum_{k=1}^{\frac{x}{2}} \pi\left(\left\lfloor \frac{x}{k} \right\rfloor \right) g(k) \\
&= G(x) + G\left(\left\lfloor \frac{x}{2} \right\rfloor \right) + \sum_{k=1}^{\frac{x}{2}-1} \left(\pi\left(\left\lfloor \frac{x}{k} \right\rfloor \right) - \pi\left(\left\lfloor \frac{x}{k+1} \right\rfloor \right) \right) G(k).
\end{aligned}$$



(a)



(b)

Figure 5.2

The upper bound on the sum is truncated to $k \in \left[1, \frac{x}{2}\right]$ in the second equation above because $\pi(1) = 0$. The third formula above follows directly by summation by parts. $\square$

Proof of (1.5c) *of Theorem 1.4.* Lemma 3.4 shows that

$$G(x) = \sum_{d \leq x} \lambda(d) C_{\Omega}(d) M\left(\left\lfloor \frac{x}{d} \right\rfloor\right).$$

The identity in (3.1) implies

$$\lambda(d) C_{\Omega}(d) = (g * 1)(d) = (\chi_{\mathbb{P}} + \varepsilon)^{-1}(d).$$

We recover the stated result from the classical inversion of summatory functions in equation (3.2). $\square$

5.2 Discrete plots and numerical experiments

The plots shown in the figures in this section compare the values of $M(x)$ and $G(x)$ with scaled forms of related auxiliary partial sums:

- In Figure 5.1, we plot comparisons of $M(x)$ to scaled forms of $G(x)$ for $x \leq 5000$. The absolute constant $C_2 := \frac{\pi^2}{6}$ where the partial sums defined by the function $Q(x) := \sum_{n \leq x} \mu^2(n)$ count the number of squarefree integers $1 \leq n \leq x$. In (a) the shift to the left on the x-axis of the former function is compared and seen to be similar in shape to the magnitude of $M(x)$ on this initial subinterval. It is unknown whether the similar shape and magnitude of these two functions persists for much larger x. In (b) we have observed unusual reflections and symmetry between the two ratios plotted in the figure. We have numerically modified the plot values to shift the denominators of $M(x)$ by one at each $x \leq 5000$ for which $M(x) = 0$.

- In Figure 5.2, we compare envelopes on the logarithmically scaled values of $G(x)x^{-1}$ to other variants of the partial sums of $g(n)$ for $x \leq 4500$. In (a) we define $G(x) := G_+(x) - G_-(x)$ where the functions $G_+(x) \geq 0$ and $G_-(x) \geq 0$ for all $x \geq 1$, i.e., the signed component functions $G_\pm(x)$ denote the unsigned contributions of only those summands $|g(n)|$ over $n \leq x$ where $\lambda(n) = \pm 1$, respectively. The summatory function $Q(x) \sim \frac{6x}{\pi^2}$ in (b) has the same definition as in Figure 5.1 above. The second plot suggests that for large x
$$|G(x)| \ll \frac{|G|(x)}{(\log x)\sqrt{\log\log x}} = \frac{1}{(\log x)\sqrt{\log\log x}} \times \sum_{n \leq x} |g(n)|.$$

5.3 Local cancellation in the formulas involving the partial sums of $g(n)$

Definition 5.2. Let p_n denote the n^{th} prime for $n \geq 1$ [23, A000040]. The set of primorial integers is defined by [23, A002110]
$$\{n\#\}_{n \geq 1} = \left\{\prod_{k=1}^{n} p_k\right\}_{n \geq 1}.$$

Proposition 5.3. *As $m \to \infty$, each of the following holds:*
$$-G((4m+1)\#) \asymp (4m+1)!, \tag{A}$$
$$G\left(\frac{(4m+1)\#}{p_k}\right) \asymp (4m)!, \text{ for any } 1 \leq k \leq 4m+1. \tag{B}$$

Proof. We have by (3.7b) that for all squarefree integers $n \geq 1$
$$\begin{aligned}|g(n)| &= \sum_{j=0}^{\omega(n)} \binom{\omega(n)}{j} \times j! = (\omega(n))! \times \sum_{j=0}^{\omega(n)} \frac{1}{j!} \\ &= (\omega(n))! \times \left(e + O\left(\frac{1}{(\omega(n)+1)!}\right)\right).\end{aligned}$$

Let $m \geq 1$ be large. We obtain main terms of the form
$$\begin{aligned}\sum_{\substack{1 \leq n \leq (4m+1)\# \\ \omega(n) = \Omega(n)}} \lambda(n)|g(n)| &= \sum_{0 \leq k \leq 4m+1} \binom{4m+1}{k} (-1)^k k! \times \left(e + O\left(\frac{1}{(k+1)!}\right)\right) \\ &= -(4m+1)! + O\left(\frac{1}{4m+1}\right).\end{aligned} \tag{5.2}$$

The formula for $C_\Omega(n)$ stated in (1.6) implies the result in (A). In particular, the contributions from the summands of the inner summation on the right-hand-side of (5.2) off of the squarefree integers are at most a bounded multiple of $(-1)^k k!$ when $\Omega(n) = k$.

We can similarly show that for any $1 \leq k \leq 4m+1$
$$G\left(\frac{(4m+1)\#}{p_k}\right) \asymp \sum_{0 \leq k \leq 4m} \binom{4m}{k} (-1)^k k! \times \left(e + O\left(\frac{1}{(k+1)!}\right)\right) = (4m)! + O\left(\frac{1}{4m+1}\right). \qquad \square$$

Remark 5.4. The Riemann hypothesis (RH) is equivalent to showing that
$$M(x) = O\left(x^{\frac{1}{2}+\varepsilon}\right), \text{ for all } 0 < \varepsilon < \frac{1}{2}. \tag{5.3}$$

We expect that there is usually (almost always) a large amount cancellation between the successive values of the summatory function in (1.5c). Proposition 5.3 demonstrates the phenomenon well along the infinite subsequence of the primorials $\{(4m+1)\#\}_{m\geq 1}$. If the RH is true, the sums of the leading constants with opposing signs on the asymptotic bounds for the functions from the last proposition are necessarily required to match. In particular, we have that [6, 7]

$$n\# \sim e^{\vartheta(p_n)} \asymp n^n (\log n)^n e^{-n(1+o(1))}, \text{ as } n \to \infty.$$

The observation on the necessary cancellation in (1.5c) follows from the fact that if we obtain a contrary result, then for some fixed $\delta_0 > 0$

$$\frac{M((4m+1)\#)}{\sqrt{(4m+1)\#}} \gg \left[(4m+1)\#\right]^{\delta_0}, \text{ as } m \to \infty.$$

If the last equation holds, then we would find a contradiction to equation (5.3). Assuming the RH, we can state a stronger bound for the Mertens function along this subsequence by considering the error terms given in the proof of Proposition 5.3.

6 Conclusions

6.1 Summary

We have identified a sequence, $\{g(n)\}_{n\geq 1}$, that is the Dirichlet inverse of the shifted strongly additive function $\omega(n)$. There is a natural structure of the repetition of distinct values of $|g(n)|$ that depends on the configuration of the exponents of the distinct primes in the factorization of any $n \geq 2$. The definition of this auxiliary sequence provides new relations between the summatory function $G(x)$ to $M(x)$ and $L(x)$. The sign of $g(n)$ is given by $\lambda(n)$ for all $n \geq 1$. The new results proved within this article are significant in providing another lense through which we can view $M(x)$ in terms of the unsigned sequences, $C_\Omega(n)$ and $|g(n)|$, their partial sums, and the local sign changes of the function $\lambda(n)$.

6.2 Discussion of the new results

6.2.1 Randomized models of the Möbius function

Probabilistic models of the Möbius function lead us to consider the behavior of $M(x)$ as a sum of independent and identically distributed (i.i.d.) random variables. Suppose that $\{X_n\}_{n\geq 1}$ is a sequence of i.i.d. $\{-1,0,1\}$-valued random variables such that for all $n \geq 1$

$$\mathbb{P}[X_n = -1] = \mathbb{P}[X_n = +1] = \frac{3}{\pi^2}, \text{ and } \mathbb{P}[X_n = 0] = 1 - \frac{6}{\pi^2},$$

i.e., so that the sequence provides a randomized model of the values of $\mu(n)$ on the average. We may then approximate the partial sums as $M(x) \cong S_x$ where $S_x := \sum_{n\leq x} X_n$ for all $x \geq 1$. This viewpoint models predictions of certain limiting asymptotic behavior of the Mertens function including [5, Thm. 9.4; §9]

$$\mathbb{E}[S_x] = 0, \mathrm{Var}[S_x] = \frac{6x}{\pi^2}, \text{ and } \limsup_{x\to\infty} \frac{|S_x|}{\sqrt{x \log\log x}} = \frac{2\sqrt{3}}{\pi} \text{ (almost surely).}$$

6.2.2 Comparison of known formulas for $M(x)$ involving $\lambda(n)$

The Mertens function is related to the partial sums in (1.2) via the relation [15, 16]

$$M(x) = \sum_{d\leq\sqrt{x}} \mu(d) L\left(\left\lfloor \frac{x}{d^2} \right\rfloor\right), \text{ for } x \geq 1. \tag{6.1}$$

The relation in (6.1) gives an exact expression for $M(x)$ with summands involving $L(x)$ that are oscillatory. In contrast, the exact expansions for the Mertens function given in Theorem 1.4 express $M(x)$ as finite sums over $\lambda(n)$ with weighted coefficients that are unsigned. The property of the symmetry of the distinct values of $|g(n)|$ with respect to the prime factorizations of $n \geq 2$ in (4.4) suggests that the unsigned weights on $\lambda(n)$ in the new formulas from the theorem yield new insights compared to equation (6.1).

6.2.3 Adaptive strategies that leverage the unpredictability of $\lambda(n)$ versus $\mu(n)$

Stating tight bounds on the distribution of $L(x)$ is a problem that is equally as difficult as understanding the growth of $M(x)$ along infinite subsequences (cf. [13, 10, 25]). Indeed, $\lambda(n) = \mu(n)$ for all squarefree $n \geq 1$ so that $\lambda(n)$ agrees with $\mu(n)$ at most large n. We infer that $\lambda(n)$ must inherit the pseudo-randomized quirks of $\mu(n)$ predicted by Sarnak's conjecture. On the other hand, the formulas in Theorem 1.4 are more desirable to explore than other classical formulae for $M(x)$ for the following reasons:

- Breakthrough work in recent years due to Matomäki, Radziwiłł and Soundararajan to bound multiplicative functions in short intervals has proven fruitful when applied to $\lambda(n)$ [24, 17]. The analogs of results of this type corresponding to the Möbius function are not clearly attained;
- The squarefree $n \geq 1$ on which $\lambda(n)$ and $\mu(n)$ must identically agree are in some senses easier integer cases to handle insomuch as we can prove very regular properties that govern the distributions of the distinct values of $\omega(n)$, $\Omega(n)$ and their difference over $n \leq x$ as $x \to \infty$ [18, cf. §2.4; §7.4];
- The function $\lambda(n)$ is completely multiplicative. Hence, the non-zero function $\lambda(n)$ may be a kinder cousin to the multiplicative $\mu(n)$ on the integers $n \geq 4$ for which $\mu(n) = 0$.

Acknowledgements

The proofs of the results in Appendix B are closely adapted from communication with Gergő Nemes of the Alfréd Rényi Institute of Mathematics. We thank him for the helpful hints on applying the results from his articles.

Reference

[1] T. M. Apostol. *Introduction to Analytic Number Theory.* Springer–Verlag, 1976.

[2] R. Arratia, A. D. Barbour, and Simon Tavaré. *Logarithmic Combinatorial Structures: A Probabilistic Approach.* European Mathematical Society Publishing House, 2003.

[3] P. T. Bateman and H. G. Diamond. *Analytic Number Theory.* World Scientific Publishing, 2004.

[4] P. Billingsley. On the central limit theorem for the prime divisor function. *Amer. Math. Monthly*, 76(2):132–139, 1969.

[5] P. Billingsley. *Probability and measure.* Wiley, third edition, 1994.

[6] P. Dusart. The k^{th} prime is greater than $k(\log k + \log\log k - 1)$ for $k \geq 2$. *Math. Comp.*, 68(225):411–415, 1999.

[7] P. Dusart. Estimates of some functions over primes without R.H., 2010.

[8] P. D. T. A. Elliott. *Probabilistic Number Theory I: Mean-Value Theorems.* Springer New York, 1979.

[9] P. Erdős and M. Kac. The Gaussian errors in the theory of additive arithmetic functions. *American Journal of Mathematics*, 62(1):738–742, 1940.

[10] N. Frantzikinakis and B. Host. The logarithmic Sarnak conjecture for ergodic weights. *Ann. of Math. (2)*, 187(3):869–931, 2018.

[11] C. E. Fröberg. On the prime zeta function. *BIT Numerical Mathematics*, 8:87–202, 1968.

[12] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, 1994.

[13] B. Green and T. Tao. The Möbius function is strongly orthogonal to nilsequences. *Ann. of Math. (2)*, 175(2):541–566, 2012.

[14] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, 2008 (Sixth Edition).

[15] P. Humphries. The distribution of weighted sums of the Liouville function and Pólya's conjecture. *J. Number Theory*, 133:545–582, 2013.

[16] R. S. Lehman. On Liouville's function. *Math. Comput.*, 14:311–320, 1960.

[17] K. Matomäki and M. Radziwiłł. Multiplicative functions in short intervals. *Ann. of Math.*, 183:1015–1056, 2016.

[18] H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory: I, Classical Theory*. Cambridge, 2006.

[19] G. Nemes. The resurgence properties of the incomplete gamma function II. *Stud. Appl. Math.*, 135(1):86–116, 2015.

[20] G. Nemes. The resurgence properties of the incomplete gamma function I. *Anal. Appl. (Singap.)*, 14(5):631–677, 2016.

[21] G. Nemes and A. B. Olde Daalhuis. Asymptotic expansions for the incomplete gamma function in the transition regions. *Math. Comp.*, 88(318):1805–1827, 2019.

[22] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.

[23] N. J. A. Sloane. The Online Encyclopedia of Integer Sequences, 2021. http://oeis.org.

[24] K. Soundararajan. The Liouville function in short intervals (after Matomäki and Radziwiłł). *S einaire Boubake*, 2015-2016, no. 1119(390):453–472, 2017.

[25] T. Tao. The logarithmically averaged Chowla and Elliott conjectures for two-point correlations. *Forum of Mathematics*, 4:e8, 2016.

[26] G. Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*. American Mathematical Society, 2015.

A The distributions of $\omega(n)$ and $\Omega(n)$

As $n \to \infty$, we have that

$$\frac{1}{n} \times \sum_{k \leq n} \omega(k) = \log\log n + B_1 + o(1),$$

and

$$\frac{1}{n} \times \sum_{k \leq n} \Omega(k) = \log\log n + B_2 + o(1),$$

where $B_1 \approx 0.261497$ and $B_2 \approx 1.03465$ are absolute constants [14, §22.10]. The next theorems reproduced from [18, §7.4] bound the frequency of the number of times $\Omega(n)$ $n \leq x$ diverges substantially from its average order at integers $n \leq x$ when x is large (cf. [9, 4]).

Theorem A.1. *For $x \geq 2$ and $r > 0$, let*

$$A(x,r) := \#\left\{n \leq x : \Omega(n) \leq r \log\log x\right\},$$
$$B(x,r) := \#\left\{n \leq x : \Omega(n) \geq r \log\log x\right\}.$$

If $0 < r \leq 1$, then

$$A(x,r) \ll \quad x(\log x)^{r-1-r\log r}, \text{ as } x \to \infty.$$

If $1 \leq r < 2$, then

$$B(x,r) \ll x(\log x)^{r-1-r\log r}, \text{ as } x \to \infty.$$

Proof. The proof of this theorem is given in [18, Thm. 7.20; §7.4]. It uses an adaptation of Rankin's method in combination with the result in proved in [18, Thm. 7.18; §7.4] to obtain the two upper bounds. □

Theorem A.2. *For integers $k \geq 1$ and $x \geq 2$*

$$\widehat{\pi}_k(x) := \#\{1 \leq n \leq x : \Omega(n) = k\}.$$

For $0 \leq |z| < R$, we define the function

$$\mathcal{G}(z) := \frac{1}{\Gamma(1+z)} \times \prod_p \left(1 - \frac{z}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^z.$$

For $0 < R < 2$, uniformly for $1 \leq k \leq R \log\log x$

$$\widehat{\pi}_k(x) = \frac{x}{\log x} \times \mathcal{G}\left(\frac{k-1}{\log\log x}\right) \frac{(\log\log x)^{k-1}}{(k-1)!} \times \left(1 + O_R\left(\frac{k}{(\log\log x)^2}\right)\right), \text{ as } x \to \infty. \tag{A.1}$$

Proof. The proof of this theorem is given in [18, Thm. 7.19; §7.4]. The notation $\widehat{\pi}_k(x)$ is distinct from that used in other references [18, Eqn. (7.61)] [26, cf. §II.6]. □

Theorem A.3. *For integers $k \geq 1$ and $x \geq 2$, we define*

$$\pi_k(x) := \#\{2 \leq n \leq x : \omega(n) = k\}.$$

We define the function

$$\widetilde{\mathcal{G}}(z) := \frac{1}{\Gamma(1+z)} \times \prod_p \left(1 + \frac{z}{p-1}\right)\left(1 - \frac{1}{p}\right)^z, \text{ for } |z| \leq R < 2.$$

For fixed $0 < R < 2$, as $x \to \infty$ we have uniformly for $1 \leq k \leq R \log\log x$ that

$$\pi_k(x) = \frac{x}{\log x} \times \widetilde{\mathcal{G}}\left(\frac{k-1}{\log\log x}\right) \frac{(\log\log x)^{k-1}}{(k-1)!} \times \left(1 + O_R\left(\frac{k}{(\log\log x)^2}\right)\right). \tag{A.2}$$

Proof. We can extend the proofs in [18, §7] to obtain analogous results on the distribution of $\omega(n)$. This result is cited as an exercise in [18]. □

B The upper incomplete gamma function

Definition B.1. The (upper) incomplete gamma function is defined by [22, §8.4]

$$\Gamma(a,z) = \int_z^\infty t^{a-1}e^{-t}dt, \text{ for } a,z \in \mathbb{R}^+.$$

The function $\Gamma(a,z)$ can be continued to an analytic function of z on the universal covering of $\mathbb{C}\setminus\{0\}$. We similarly define the regularized incomplete gamma function for real $a,z>0$ by $Q(a,z) := \Gamma(a,z)\Gamma(a)^{-1}$.

The following properties are known [22, §8.4; §8.11(i)]:

$$Q(a,z) = e^{-z} \times \sum_{k=0}^{a-1} \frac{z^k}{k!}, \text{ for } a \in \mathbb{Z}^+ \text{ and } z \in \mathbb{R}^+, \tag{B.1a}$$

$$\Gamma(a,z) \sim z^{a-1}e^{-z}, \text{ for fixed } a>0 \text{ and } z>0 \text{ as } z\to\infty. \tag{B.1b}$$

For $z>0$, as $z\to\infty$ we have that [19]

$$\Gamma(z,z) = \sqrt{\frac{\pi}{2}} z^{z-\frac{1}{2}} e^{-z} \times \left(1 + O\left(\frac{1}{\sqrt{z}}\right)\right). \tag{B.1c}$$

For fixed, finite real $\rho>0$, we define the sequence $\{b_n(\rho)\}_{n\geq 0}$ by the following recurrence relation:

$$b_n(\rho) = \rho\cdot(1-\rho)b'_{n-1}(\rho) + \rho\cdot(2n-1)b_{n-1}(\rho) + \delta_{n,0}.$$

For fixed $\rho>0$, the sequence $\{b_n(\rho)\}_{n\geq 0}$ satisfies a Rodrigues type formula of the form [20, Thm. 1.1]

$$b_n(\rho) = (1-\rho)^n \times \frac{d^n}{dt^n}\left(\frac{(\rho-1)t}{\rho e^t - t - \rho}\right)^{n+1}\Bigg|_{t=0}.$$

If $z,a\to\infty$ with $z=\rho a$ for some $\rho>1$ such that $(\rho-1)^{-1} = o\left(\sqrt{|a|}\right)$, then [19]

$$\Gamma(a,z) \sim z^a e^{-z} \times \sum_{n\geq 0} \frac{(-a)^n b_n(\rho)}{(z-a)^{2n+1}}. \tag{B.1d}$$

Proposition B.2. *Let $a,z>0$ be taken such that as $a,z\to\infty$ (independently), we obtain a finite limit for the parameter $\rho = \frac{z}{a} > 0$. The following results hold:*

- *If $\rho\in(0,1)$, then as $z\to\infty$*

$$\Gamma(a,z) = \Gamma(a) + O_\rho\left(z^{a-1}e^{-z}\right). \tag{B.2a}$$

- *If $\rho>1$, then as $z\to\infty$*

$$\Gamma(a,z) = \frac{z^{a-1}e^{-z}}{1-\rho^{-1}} + O_\rho\left(z^{a-2}e^{-z}\right). \tag{B.2b}$$

- *If $\rho > W(1) > 0.56714$, then as $z\to\infty$*

$$\Gamma(a,ze^{\pm\pi\imath}) = -e^{\pm\pi\imath a}\frac{z^{a-1}e^{z}}{1+\rho^{-1}} + O_\rho\left(z^{a-2}e^{z}\right). \tag{B.2c}$$

Remark. The first two estimates in Proposition B.2 are only useful when ρ is bounded away from the transition point at one. We cannot write the last expansion above as $\Gamma(a,-z)$ directly unless $a\in\mathbb{Z}^+$ as the incomplete gamma function has a branch point at the origin with respect to its second variable. This function becomes a single-valued analytic function of its second input by continuation on the universal covering of $\mathbb{C}\setminus\{0\}$.

Proof of Proposition B.2. The first asymptotic estimate follows directly from the following asymptotic series expansion that holds as $z \to \infty$ [21, Eq. (2.1)]:

$$\Gamma(a,z) \sim \Gamma(a) + z^a e^{-z} \times \sum_{k\geq 0} \frac{(-a)^k b_k(\rho)}{(z-a)^{2k+1}}.$$

Suppose that $\rho > 0$. The notation from (B.1d) and [20, Thm. 1.1] shows that

$$\Gamma(a,z) = \frac{z^{a-1}e^{-z}}{1-\rho^{-1}} + z^a e^{-z} R_1(a,\rho).$$

From the bounds in [20, §3.1], we have

$$|z^a e^{-z} R_1(a,\rho)| \leq z^a e^{-z} \times \frac{a \cdot b_1(\rho)}{(z-a)^3} = \frac{z^{a-2}e^{-z}}{(1-\rho^{-1})^3}$$

The main and error terms in the previous equation can also be seen by applying the asymptotic series in (B.1d) directly.

The proof of the third equation above follows from the asymptotics [19, Eq. (1.1)]

$$\Gamma(-a,z) \sim z^{-a}e^{-z} \times \sum_{n\geq 0} \frac{a^n b_n(-\rho)}{(z+a)^{2n+1}},$$

by setting $(a,z) \mapsto (ae^{\pm\pi\imath}, ze^{\pm\pi\imath})$ so that $\rho = \frac{z}{a} > W(1)$. The restriction on the range of ρ over which the third formula holds is made to ensure that the formula from the reference is valid at negative real a. □

Lemma B.3. *As $x \to \infty$*

$$\frac{x}{\log x} \times \left| \sum_{1\leq k\leq \log\log x} \frac{(-1)^k (\log\log x)^{k-1}}{(k-1)!} \right| = \frac{x}{2\sqrt{2\pi \log\log x}} \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right). \tag{B.3a}$$

For any $a \in \left(1, W(1)^{-1}\right) \subset (1, 1.76321)$, as $x \to \infty$

$$\left| \sum_{k=1}^{a\log\log x} \frac{(-1)^k (\log\log x)^{k-1}}{(k-1)!} \right| = \frac{a^{\frac{1}{2}-\{a\log\log x\}}}{(1+a)} \times \frac{(\log x)^{a-a\log a}}{\sqrt{2\pi\log\log x}} \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right). \tag{B.3b}$$

The function $\{x\} = x - \lfloor x \rfloor \in [0,1)$ denotes the fractional part of any $x \in \mathbb{R}$.

Proof of Equation (B.3a). We have for $n \geq 1$ and any $t > 0$ by (B.1a) that

$$\sum_{1\leq k\leq n} \frac{(-1)^k t^{k-1}}{(k-1)!} = -e^{-t} \times \frac{\Gamma(n,-t)}{(n-1)!}.$$

Suppose that $t = n + \xi$ with $\xi = O(1)$. By the third formula in Proposition B.2 with the parameters $(a,z,\rho) \mapsto \left(n, t, 1 + \frac{\xi}{n}\right)$, we deduce that as $n, t \to \infty$.

$$\Gamma(n,-t) = (-1)^{n+1} \times \frac{t^n e^t}{t+n} + O\left(\frac{n t^n e^t}{(t+n)^3}\right) = \frac{(-1)^{n+1} t^n e^t}{2n} + O\left(\frac{t^{n-1}e^t}{n}\right).$$

Accordingly, we see that

$$\sum_{1\leq k\leq n} \frac{(-1)^k t^{k-1}}{(k-1)!} = \frac{(-1)^n t^n}{2n!} + O\left(\frac{t^{n-1}}{n!}\right).$$

The form of Stirling's formula in [22, cf. Eq. (5.11.8)] shows that

$$n! = \Gamma(1+t-\xi) = \sqrt{2\pi} t^{t-\xi+\frac{1}{2}} e^{-t} \times \left(1 + O\left(t^{-1}\right)\right) = \sqrt{2\pi} t^{n+\frac{1}{2}} e^{-t} \times \left(1 + O\left(t^{-1}\right)\right).$$

Hence, as $n \to \infty$ with $t := n + \xi$ and $\xi = O(1)$, we obtain that

$$\sum_{k=1}^{n} \frac{(-1)^k t^{k-1}}{(k-1)!} = \frac{(-1)^n e^t}{2\sqrt{2\pi t}} + O\left(e^t t^{-\frac{3}{2}}\right).$$

The conclusion follows by taking $n := \lfloor \log\log x \rfloor$ and $t := \log\log x$. □

Proof of Equation (B.3b). The argument is nearly identical to the proof of the first equation. The key modifications are to set $t := an + \xi$ where $\xi = O(1)$, take the parameters $(a, z, \rho) \mapsto \left(an, t, 1 + \frac{\xi}{an}\right)$, and use the identity that $a^{an} = e^{an \log a}$ to simplify the main term obtained from Stirling's formula. □

C Inversion of partial sums of Dirichlet convolutions

Proof of Theorem 1.3. Suppose that h, r are arithmetic functions such that $r(1) \neq 0$, i.e., so that the function r is invertible with respect to the operation of Dirichlet convolution. The following formulas hold for all $x \geq 1$:

$$\begin{aligned} S_{r*h}(x) &:= \sum_{n=1}^{x} \sum_{d|n} r(n) h\left(\frac{n}{d}\right) = \sum_{d=1}^{x} r(d) \times H\left(\left\lfloor \frac{x}{d} \right\rfloor\right) \\ &= \sum_{i=1}^{x} \left(R\left(\left\lfloor \frac{x}{i} \right\rfloor\right) - R\left(\left\lfloor \frac{x}{i+1} \right\rfloor\right)\right) \times H(i). \end{aligned} \tag{C.1}$$

The first formula on the right-hand-side above is well known from the references. The second formula is justified directly using summation by parts as [22, §2.10(ii)]

$$\begin{aligned} S_{r*h}(x) &= \sum_{d=1}^{x} h(d) \times R\left(\left\lfloor \frac{x}{d} \right\rfloor\right) \\ &= \sum_{i \leq x} \left(\sum_{j \leq i} h(j)\right) \times \left(R\left(\left\lfloor \frac{x}{i} \right\rfloor\right) - R\left(\left\lfloor \frac{x}{i+1} \right\rfloor\right)\right). \end{aligned}$$

For Boolean-valued conditions `cond`, we adopt Iverson's convention that $[\texttt{cond}]_\delta$ evaluates to one precisely when `cond` is true and to zero otherwise. We form the invertible matrix of coefficients (denoted by $\hat{R}$ below) associated with the linear system that defines $H(j)$ for $1 \leq j \leq x$ in (C.1) by defining

$$R_{x,j} := R\left(\left\lfloor \frac{x}{j} \right\rfloor\right) [j \leq x]_\delta,$$

and

$$r_{x,j} := R_{x,j} - R_{x,j+1}, \text{ for } 1 \leq j \leq x.$$

Since $r_{x,x} = R(1) = r(1) \neq 0$ for all $x \geq 1$ and $r_{x,j} = 0$ for all $j > x$, the matrix we have defined in this problem is lower triangular with a non-zero constant on its diagonals, and so is invertible. If we let $\hat{R} := (R_{x,j})$, then the next matrix is expressed by applying an invertible shift operation as

$$(r_{x,j}) = \hat{R}\left(I - U^T\right).$$

For any $N \geq 1$, the $N \times N$ square matrix U has $(i,j)^{th}$ entries for all $1 \leq i, j \leq N$ when $N \geq x$ that are defined by $(U)_{i,j} = \delta_{i+1,j}$ so that

$$\left[\left(I - U^T\right)^{-1}\right]_{i,j} = [j \leq i]_\delta.$$

We observe that

$$\left\lfloor \frac{x}{j} \right\rfloor - \left\lfloor \frac{x-1}{j} \right\rfloor = \begin{cases} 1, & \text{if } j|x; \\ 0, & \text{otherwise.} \end{cases}$$

The previous equation implies that

$$R\left(\left\lfloor \frac{x}{j} \right\rfloor\right) - R\left(\left\lfloor \frac{x-1}{j} \right\rfloor\right) = \begin{cases} r\left(\frac{x}{j}\right), & \text{if } j|x; \\ 0, & \text{otherwise.} \end{cases} \tag{C.2}$$

We use the property in (C.2) to shift the matrix $\hat{R}$, and then invert the result to obtain a matrix involving the Dirichlet inverse of r as

$$\left(\left(I - U^T\right)\hat{R}\right)^{-1} = \left(r\left(\frac{x}{j}\right)[j|x]_\delta\right)^{-1} = \left(r^{-1}\left(\frac{x}{j}\right)[j|x]_\delta\right).$$

Our target matrix is expressed by

$$(r_{x,j}) = \left(I - U^T\right)\left(r\left(\frac{x}{j}\right)[j|x]_\delta\right)\left(I - U^T\right)^{-1}.$$

We can evaluate its inverse by a similarity transformation conjugated by shift operators by

$$\begin{aligned} (r_{x,j})^{-1} &= \left(I - U^T\right)^{-1}\left(r^{-1}\left(\frac{x}{j}\right)[j|x]_\delta\right)\left(I - U^T\right) \\ &= \left(\sum_{k=1}^{\left\lfloor \frac{x}{j} \right\rfloor} r^{-1}(k)\right)\left(I - U^T\right) \\ &= \left(\sum_{k=1}^{\left\lfloor \frac{x}{j} \right\rfloor} r^{-1}(k) - \sum_{k=1}^{\left\lfloor \frac{x}{j+1} \right\rfloor} r^{-1}(k)\right). \end{aligned}$$

The summatory function $H(x)$ is given exactly by a vector product with the inverse matrix from the previous equation as

$$H(x) = \sum_{k=1}^{x}\left(\sum_{j=\left\lfloor \frac{x}{k+1} \right\rfloor+1}^{\left\lfloor \frac{x}{k} \right\rfloor} r^{-1}(j)\right) \times S_{r*h}(k), \text{ for } x \geq 1.$$

We can prove a second inversion formula by adapting our argument used to prove (C.1) above. This leads to the alternate expression for $H(x)$ given by

$$H(x) = \sum_{k=1}^{x} r^{-1}(k) \times S_{r*h}\left(\left\lfloor \frac{x}{k} \right\rfloor\right), \text{ for } x \geq 1. \qquad \square$$

D The proof of Theorem 1.6

Lemma D.1. *As $x \to \infty$*

$$\sum_{n \leq x} \log C_\Omega(n) = \sum_{k \geq 1} \#\{n \leq x : \Omega(n) = k\} \times \log(k!) \times \left(1 + O\left(\frac{1}{(\log\log x)^{\frac{1}{3}}}\right)\right). \tag{D.1}$$

Proof. Recall that an integer $n \geq 1$ is squarefree if and only if $\mu^2(n) = 1$. Equation (1.6) shows that

$$\sum_{\substack{n \leq x \\ \mu^2(n)=1}} \log C_\Omega(n) = \sum_{k \geq 1} \#\{n \leq x : \Omega(n) = k\} \times \log(k!),$$

where the sum in the last equation is finite since $\Omega(n) \leq \log_2(x)$ for all $x \geq 2$. The key to the rest of the proof is to understand that the main term of the sum on the left-hand-side of the equation is obtained by summing over only the squarefree $n \leq x$, i.e., the $n \leq x$ such that $\mu^2(n) = 1$. We claim that

$$\sum_{k\geq 1}\sum_{\substack{n\leq x\\ \Omega(n)=k}} \log C_\Omega(n) \sim \sum_{k\geq 1}\sum_{\substack{n\leq x\\ \mu^2(n)=1\\ \Omega(n)=k}} \log C_\Omega(n).$$

The function $\mathrm{rad}(n)$ is the radix (or squarefree part) of n which evaluates to the largest squarefree factor of n, or equivalently to the product of all primes $p|n$ [23, A007913]. For integers $x \geq 1$ and $1 \leq k \leq \log_2(x)$, let the sets

$$\mathcal{S}_k\left(\{\varpi_j\}_{j=1}^k; x\right) := \left\{2 \leq n \leq x : \mu(n) = 0, \omega(n) = k, \frac{n}{\mathrm{rad}(n)} = p_1^{\varpi_1} \times \cdots \times p_k^{\varpi_k},\ p_i \neq p_j \text{ prime for } 1 \leq i < j \leq k\right\}.$$

The crucial takeaway from the definition in the last equation is as follows: For every non-squarefree integer $n \in [2, x]$, there is some $1 \leq k \leq \log_2(x)$ and a sequence of positive integers $\{\varpi_j\}_{1\leq j\leq k}$ such that $n \in \mathcal{S}_k\left(\{\varpi_j\}_{j=1}^k; x\right)$. Let the function

$$\mathcal{N}_k(\varpi_1, \ldots, \varpi_k; x) := \frac{\left|\mathcal{S}_k\left(\{\varpi_j\}_{j=1}^k; x\right)\right|}{x}.$$

The special case where $\{\varpi_j^*\}_{1\leq j\leq k} \equiv \{0, 1\}$ (with value one of multiplicity exactly one) is denoted by

$$\widehat{T}_k(x) := \mathcal{N}_k\left(\varpi_1^*, \ldots, \varpi_k^*; x\right).$$

If $2 \leq n \leq x$ is not squarefree and $n \in \mathcal{S}_k\left(\{\varpi_j\}_{j=1}^k; x\right)$, then we must have that $\varpi_j \geq 1$ for at least one $1 \leq j \leq k$. Clearly for any $k \geq 1$

$$\mathcal{N}_k(\varpi_1, \ldots, \varpi_k; x) \ll \widehat{T}_k(x).$$

We claim that (see proof below)

$$\widehat{T}_k(x) \ll \frac{1}{(\log\log x)^{\frac{2}{3}}} \times \#\{n \leq x : \omega(n) = k\} \text{ for all } k \geq 1, \text{ as } x \to \infty. \tag{D.2}$$

The upper bounds on the functions $\widehat{T}_k(x)$ in equation (D.2) show that the sum of denominator differences from (1.6) we subtract from the main term contributions from the squarefree $n \leq x$ is asymptotically insubstantial. That is, we have proved that as $x \to \infty$

$$\sum_{\substack{n\leq x\\ \mu(n)=0}} \log C_\Omega(n) = o\left(\sum_{\substack{n\leq x\\ \mu^2(n)=1}} \log C_\Omega(n)\right). \qquad \square$$

Proof of Equation (D.2). We recall the next two famous asymptotic formulae. We know that $\pi(x) = \frac{x}{\log x} \times \left(1 + O\left(\frac{1}{\log x}\right)\right)$ as $x \to \infty$ [14, §22.4]. A theorem of Mertens is stated as follows: $\sum_{p\leq x} p^{-1} \sim \log\log x$ as $x \to \infty$ [14, §22.7–22.8].

The bound can be proved by induction on $k \geq 1$ using the inductive hypothesis (IH)

$$\widehat{T}_m(x) \ll \frac{x^{1-2^{-m}}(\log\log x)^m}{(\log x)^{1+2^{-m}}}, \text{ for all } 1 \leq m \leq k, \text{ as } x \to \infty. \tag{IH}$$

The case where $k := 1$ is evaluated by computation as follows:

$$\widehat{T}_1(x) = \sum_{p\leq\sqrt{x}} 1 \ll \frac{\sqrt{x}}{\log x}.$$

Suppose that $k \geq 1$ and that the IH holds at k. Theorem A.3 combined with the bounds in the next equations shows that equation (D.2) holds for all finite $k \geq 1$. In particular, we have by the IH and Hölder's inequality with $\left(p^{-1}, q^{-1}\right) = \left(1 - 2^{-k}, 2^{-k}\right)$ that

$$\begin{aligned}\widehat{T}_{k+1}(x) &\ll \sum_{p \leq \sqrt{x}} \widehat{T}_k\left(\frac{x}{p}\right) \ll \frac{x^{1-2^{-k}} (\log\log x)^k}{(\log x)^{1+2^{-k}}} \times \left(\sum_{p \leq x} p^{-1}\right)^{1-2^{-k}} \times \pi\left(\sqrt{x}\right)^{2^{-k}} \\ &\ll \frac{2^{2^{-k}} x^{1-2^{-(k+1)}} (\log\log x)^{k+1-2^{-k}}}{(\log x)^{1+2^{-(k-1)}}}, \text{ as } x \to \infty.\end{aligned}$$

Hence, if the following equation holds, then the proof by induction is complete:

$$\widehat{U}_{k+1}(x) := \frac{2^{2^{-k}}}{(\log\log x)^{2^{-k}} (\log x)^{3 \cdot 2^{-(k+1)}}} \ll 1, \text{ as } x \to \infty. \tag{D.3}$$

For any fixed finite large x, $1 \leq k \leq \log_2(x)$. In particular, we have that $2^{-k} \geq x^{-1}$. We can then establish the following upper bounds to prove equation (D.3):

$$\widehat{U}_{k+1}(x) \ll \left(\frac{2}{(\log\log x)(\log x)^{\frac{3}{2}}}\right)^{\frac{1}{x}} \ll \exp\left(-\frac{\log\log\log x}{x}\right) = 1 + O\left(\frac{\log\log\log x}{x}\right). \qquad \square$$

Proof of Theorem 1.6. We will split the full sum on the left-hand-side of (D.1) into two sums, each over disjoint indices, that form the main and error terms, $L_\Omega(x)$ and $\widehat{L}_\Omega(x)$, respectively. For $x \geq 3$, consider the following partial sums:

$$\begin{aligned}L_{M,\Omega}(x) &:= \sum_{\substack{n \leq x \\ \Omega(n) \leq \log\log x}} \log C_\Omega(n), \\ L_{E,\Omega}(x) &:= \sum_{\substack{n \leq x \\ \Omega(n) > \log\log x}} \log C_\Omega(n).\end{aligned}$$

We claim that the main term is given by

$$L_{M,\Omega}(x) = x(\log\log x)(\log\log\log x) \times \left(1 + O\left(\frac{1}{(\log\log x)^{\frac{1}{3}}}\right)\right). \tag{D.4}$$

To bound the error term, we claim that

$$L_{E,\Omega}(x) = o\left(x(\log\log x)^{\frac{2}{3}}(\log\log\log x)\right), \text{ as } x \to \infty. \tag{D.5}$$

Equations (D.4) and (D.5) yield the conclusion of Theorem 1.6 since

$$\sum_{n \leq x} \log C_\Omega(n) = L_{M,\Omega}(x) + L_{E,\Omega}(x), \text{ for all } x > e^e.$$

The proofs of equations (D.4) and (D.10) are completed below. $\square$

Proof of Equation (D.4). Lemma D.1 and Theorem A.2 from the appendix show that

$$L_{M,\Omega}(x) = \frac{x}{\log x} \times \sum_{1 \leq k \leq \log\log x} \mathcal{G}\left(\frac{k-1}{\log\log x}\right) \frac{(\log\log x)^{k-1}}{(k-1)!} \log(k!) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right),$$

where

$$\mathcal{G}(z) := \frac{1}{\Gamma(1+z)} \times \prod_p \left(1 - \frac{z}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^z, \text{ for } |z| < 2.$$

Let the function

$$g(x,k) := \frac{(\log\log x)^{k-1}}{(k-1)!} \times \left(\frac{(2k+1)}{2}\log(1+k) - k + O(1)\right).$$

For any $j \geq 0$, Binet's formula for the log-gamma function is stated as follows [22, §5.9(i)]:

$$\log j! = \left(j + \frac{1}{2}\right)\log(1+j) - j + O(1),$$

where $z! := \Gamma(1+z)$ for $z \geq 0$. Binet's formula shows that

$$L_{M,\Omega}(x) = \frac{x}{\log x} \times \sum_{1\leq k\leq \log\log x} \mathcal{G}\left(\frac{k-1}{\log\log x}\right) g(x,k) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right).$$

The Euler-Maclaurin summation (EM) formula [12, §9.5] shows that for each fixed integer $p \geq 1$ and function $f(t)$ that is p times continuously differentiable on $(0,\infty)$ [23, A000367; A002445]

$$\begin{aligned} L_{M,\Omega}(x) = \frac{x}{\log x} \times \Bigg(\int_1^{\log\log x} \mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)dt + \frac{1}{2}\mathcal{G}\left(1 - \frac{1}{\log\log x}\right) g\left(x, \log\log x\right) - \frac{g(x,1)}{2} \\ + O\left(\sum_{k=1}^{p} \frac{B_k}{k!} \times \frac{\partial^{(k-1)}}{\partial t^{(k-1)}}\left[\mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)\right]_{t=1}^{t=\log\log x} + \widehat{R}_p[\mathcal{G}\cdot g]\right)\Bigg) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right). \end{aligned}$$

The degree-p remainder term is bounded by

$$\left|\widehat{R}_p[\mathcal{G}\cdot g]\right| = O\left(\frac{1}{p!} \times \int_1^{\log\log x} B_p(\{t\}) f^{(p)}(t)dt\right).$$

It suffices to choose $p := 1$ in the expansion above. When $f(t) \equiv f_x(t) := \mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)$, we denote the degree-$p$ EM formula error term by

$$E_p(x) := \sum_{k=1}^{p} \frac{B_k}{k!} \times \frac{\partial^{(k-1)}}{\partial t^{(k-1)}}\left[\mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)\right]_{t=1}^{t=\log\log x} + \frac{1}{m!} \times \int_1^{\log\log x} |B_p(\{t\})| f_x^{(m)}(t)dt$$

Specializing to the case where $p := 1$ yields the upper bound

$$\begin{aligned} |E_1(x)| &\ll \frac{(\log x)(\log\log\log x)}{\sqrt{\log\log x}} + \underbrace{\int_1^{\log\log x} t^2 \log(1+t) \times \frac{(\log\log x)^t}{(t+1)!} dt}_{:=I_1(x)} \qquad \text{(D.6)} \\ &\ll (\log x)\sqrt{\log\log x}(\log\log\log x), \text{ as } x \to \infty. \end{aligned}$$

The integral term, $I_1(x)$, in equation (D.6) is bounded using Hölder's inequality with $p, q := 1, \infty$ as follows:

$$\begin{aligned} |I_1(x)| &\ll (\log\log x) \times \max_{1\leq t\leq \log\log x} \frac{(\log\log\log x)(\log\log x)^{t+2}}{(1+t)\Gamma(1+t)} \\ &\ll (\log\log x)^3 (\log\log\log x) \times \max_{1\leq t\leq \log\log x} \frac{(\log\log x)^t e^t}{t^{t+\frac{3}{2}}} \\ &\ll (\log x)\sqrt{\log\log x}(\log\log\log x), \text{ as } x \to \infty. \qquad \text{(D.7)} \end{aligned}$$

The maximum in the previous equations is attained when $t = \log\log x$.

The mean value theorem states that for all large x there is some $c \in [1, \log\log x]$ such that

$$\int_1^{\log\log x} \mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)dt = \mathcal{G}\left(\frac{c-1}{\log\log x}\right) \times \int_1^{\log\log x} g(x,t)dt. \qquad \text{(D.8a)}$$

For any real $y > e^e$, let

$$c(y) := \inf\left\{c \in [1, \log\log y] : \text{ equation (D.8a) holds}\right\}. \tag{D.8b}$$

Let $B_0^*(x) := \mathcal{G}\left(\frac{c(x)-1}{\log\log x}\right)$ denote the multiplier function (depending on x) given by equations (D.8). We can apply the EM formula again to see that

$$\begin{aligned}L_{M,\Omega}(x) &= \frac{x}{\log x} \times \left(\sum_{k=1}^{\log\log x} B_0^*(x) g(x,k) + \frac{1}{2}\left(1 - B_0^*(x)\right)\mathcal{G}\left(1 - \frac{1}{\log\log x}\right) g\left(x, \log\log x\right)\right. \\ &\left. + O\left(1 + \sum_{k=1}^{p} \frac{B_k}{k!} \times \frac{\partial^{(k-1)}}{\partial t^{(k-1)}}\left[\left(1 + \mathcal{G}\left(\frac{t-1}{\log\log x}\right)\right) g(x,t)\right]_{t=1}^{t=\log\log x} + \widehat{R}_p[(1+\mathcal{G})\cdot g]\right)\right) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right).\end{aligned} \tag{D.9}$$

For $p := 1$, the EM formula error term from equation (D.4) satisfies $\left|\widehat{R}_p[(1+\mathcal{G})\cdot g]\right| \ll |I_1(x)|$ for all sufficiently large large x.

We have two remaining steps to establish equation (D.4):

(i) To show that the sums on the right-hand-side of equation (D.9) give the main term of this expression for $L_{M,\Omega}(x)$ (up to a factor of the bounded function $B_0^*(x)$); and

(ii) To show that there is actually a limiting constant insomuch as $B_0^*(x) \xrightarrow{x\to\infty} 1$.

The sums in the previous equation are approximated using Abel summation applied to the following functions for $1 \le u \le \log\log x$:

$$A_x(u) := \sum_{1\le k\le u} \frac{x(\log\log x)^{k-1}}{(\log x)(k-1)!} = \frac{x\Gamma\left(u, \log\log x\right)}{\Gamma\left(u\right)}; \text{ and } f(u) := \frac{(2u+1)}{2}\log\left(1+u\right) - u + O(1).$$

That is, we have by Proposition B.2 that

$$\begin{aligned}&\frac{x}{\log x} \times \sum_{1\le k\le \log\log x} g(x,k) \\ &= A_x\left(\log\log x\right) f\left(\log\log x\right) - \frac{1}{\log\log x} \times \int_0^1 A_x(\alpha\log\log x) f'(\alpha\log\log x) d\alpha \\ &= x(\log\log x)(\log\log\log x) \times \left(1 + O\left(\frac{1}{\sqrt{\log\log x}}\right)\right) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right) \\ &\quad - \frac{x}{\log\log x} \times \int_0^1 f'(\alpha\log\log x) \times \left(1 + O\left(\frac{\sqrt{\alpha}(\log x)^{\alpha-1-\alpha\log\alpha}}{\sqrt{\log\log x}}\right)\right) \\ &= x(\log\log x)(\log\log\log x) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right).\end{aligned} \tag{D.10}$$

The following observations are not difficult to see:

$$\mathcal{G}\left(1 - \frac{1}{\log\log x}\right) = \mathcal{G}\left(1\right) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right), \text{ as } x \to \infty, \tag{D.11a}$$

$$g\left(x, \log\log x\right) = (\log\log\log x) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right), \text{ as } x \to \infty. \tag{D.11b}$$

Equations (D.9) and (D.10) and (D.11) show that

$$L_{M,\Omega}(x) = B_0^*(x) x (\log\log x)(\log\log\log x)\left(1 + O\left(\frac{1}{(\log\log x)^{\frac{1}{3}}}\right)\right), \text{ as } x \to \infty. \tag{D.12}$$

For technical reasons we have chosen a sub-optimal error term to express equation (D.12). This observation accomplishes step **(i)**.

Let $\mathcal{C}(t) := \mathcal{G}\left(\frac{c(t)-1}{\log\log[t]}\right)$ where $c(t)$ is defined as in equation (D.8b) for all real $t \geq 19$. It is not difficult to see that $\mathcal{C}(t)$ is continuous and differentiable at all $t \in (e^e, \infty)$. We can see by computation from equation (D.8a) that for all sufficiently large $t > e^e$, the derivative of this function satisfies $\mathcal{C}'(t) < 0$. Moreover, we have that $1 \leq \mathcal{C}(t) < 2$ for all $t \geq \lceil e^e \rceil$ where $0 \leq \frac{c(x)-1}{\log\log x} \leq 1$ for all integers $x \geq 16$. This means that

$$\lim_{x\to\infty} \mathcal{C}(x) = \mathcal{G}(0) = 1,$$

and so $B_0^*(x) \xrightarrow{x\to\infty} 1$. This completes step **(ii)**. Step **(ii)** combined with equation (D.12) (e.g., with step **(i)**) shows that equation (D.4) holds. □

Proof of Equation (D.5). The following equation holds:

$$\log C_\Omega(n) \ll \Omega(n)\log\Omega(n), \text{ for } n \leq x, \text{ as } x \to \infty. \tag{D.13}$$

The bound for $\log C_\Omega(n)$ in (D.13) stated in terms of the variable $n \leq x$ holds as the upper bound on the interval $x \to \infty$. The right-hand-side terms involving $\Omega(n) \in [1, \log_2(x)]$ oscillate in magnitude over the integers $1 \leq n \leq x$. The statement in the last equation follows by maximizing (and minimizing) the ratio of the right-hand-side of (D.13) to Binet's log-gamma formula. We use the notation $\mathcal{R}(n) \equiv \mathcal{R}(n,z)$ when $z \equiv \Omega(n)$ to express this ratio at all $n \leq x$. Numerical methods show that this ratio is absolutely bounded by $0 \leq \mathcal{R}(n) < +\infty$ for all $16 \leq n \leq x$ as $x \to \infty$. The global extrema of this function on the positive integers are both attained at finite integers $2 \leq n_\ell, n_u < +\infty$ and $z_\ell, z_u = \Omega(n_\ell), \Omega(n_u) \in [1,12]$.

We have that as $x \to \infty$

$$\begin{aligned}
L_{E,\Omega}(x) &\ll \sum_{\substack{n\leq x\\ \Omega(n)\geq\log\log x}} \Omega(n)\log\Omega(n) \\
&\ll \left(\sum_{\substack{n\leq x\\ \Omega(n)\geq\log\log x}} \Omega(n)^q\right)^{\frac{1}{q}} \times \left(\sum_{\substack{n\leq x\\ \Omega(n)\geq\log\log x}} (\log\Omega(n))^{\frac{q}{q-1}}\right)^{\frac{q-1}{q}} \\
&\ll \sqrt{\log\log x} \times \left(\frac{x}{\sqrt{\log\log x}} \times \int_{\log\log x}^{\log_2(x)} t^q e^{-\frac{(t-\log\log x)^2}{\log\log x}} dt\right)^{\frac{1}{q}} \times \#\{n \leq x : \Omega(n) \geq \log\log x\}^{\frac{q-1}{q}} \\
&\qquad \left(\text{Set } u := \tfrac{t-\log\log x}{\sqrt{\log\log x}}\right) \\
&\ll x(\log\log x)^{\frac{1}{2}\left(1+q-\frac{1}{q}\right)}(\log\log\log x), \text{ for any finite } q \in (1,\infty).
\end{aligned}$$

We applied Theorem A.1 to reach the last equation. The exponent of $\log\log x$ on the right-hand-side of the last equation satisfies the following bound that suffices to yield equation (D.5):

$$\epsilon(q) := \frac{1}{2}\left(1 + q - \frac{1}{q}\right) < \frac{2}{3}, \text{ for all } q \in \left(1, \frac{1+\sqrt{37}}{6}\right).$$

We can select $q := \frac{46}{39}$ to obtain the desired bound[1]. This shows that equation (D.5) holds. □

[1] This choice of q is made by rationalizing the upper bound $\frac{1+\sqrt{37}}{6} \approx 1.18046$ to a nearby rational number with a small denominator that is within 1×10^{-3} of the exact value of 1.18. This yields the exponent $\epsilon\left(\frac{46}{39}\right) = \frac{2389}{3588} \approx 0.66583$.