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 $\mathrm{Re}(s) > 1$ 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 $\mathrm{Re}(s) > 1$. We establish formulas for the average order and variance of $\log C_\Omega(n)$ and prove a central limit theorem for the distribution of its values on the integers $n \leq x$ as $x \to \infty$. Discrete convolutions of the partial sums of $g(n)$ with the prime counting function provide new exact formulas for $M(x)$.

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

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

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

That is, if $n = p_1^{\alpha_1} \times \cdots p_r^{\alpha_r}$ is the factorization of n into powers of distinct primes, then $\omega(n) = r$ and $\Omega(n) = \alpha_1 + \cdots + \alpha_r$. We use the convention that the functions $\omega(1) = \Omega(1) = 0$. The Möbius function is the multiplicative function defined by [22, 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 [22, 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 $\lambda(n) := (-1)^{\Omega(n)}$ [22, A008836]. The partial sums of this function are defined by [22, 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 [22, 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 [22, 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

1.2.1 Partial summation identities for the Mertens function

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

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

Theorem 1.3 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 $R(x) := \sum_{n \le x} r(n)$, $H(x) := \sum_{n \le x} h(n)$, and that $R^{-1}(x) := \sum_{n \le x} r^{-1}(n)$ for $x \ge 1$. The following holds for all $x \ge 1$:*

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

Moreover, for all $x \ge 1$

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

For integers $x \ge 1$, the function $\pi(x) := \sum_{p \le x} 1$ denotes the classical prime counting function [22, A000720]. We find exact formulas for $M(x)$ by applying Theorem 1.3 to the expansion of the partial sums (see Section 3.1)

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

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

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

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

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

The new results proved in this article provide a new lense through which we can view $M(x)$ in terms of sums of auxiliary unsigned functions sign weighted by $\lambda(n)$.

1.2.2 Distributions of auxiliary unsigned functions

The unsigned function $C_\Omega(n)$ is studied by Fröberg in [10]. This function has the exact formula

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

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 $n \to \infty$*
$$\frac{1}{n} \times \sum_{k\leq n} \log C_\Omega(k) = (\log\log n)(\log\log\log n)\left(1 + O\left(\frac{1}{\sqrt{\log\log n}}\right)\right).$$

A proof of Theorem 1.6 is given in Appendix D (cf. Proposition 2.4).

Definition 1.7. The cumulative density function of the standard normal distribution at z is
$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \times \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt, \text{ for any } z \in (-\infty, \infty).$$

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 $|g(n)|$ through 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 variant, we suppress the duplicate index n by writing $C_\Omega(n) := C_{\Omega(n)}(n)$ [22, 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 when $1 \leq k \leq \Omega(n)$ 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.

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

The variance of this function is given by the centralized second moments
$$\operatorname{Var}[\log C_\Omega(x)] := \frac{1}{x} \times \sum_{n\leq x} (\log C_\Omega(n) - \mathbb{E}[\log C_\Omega(x)])^2.$$

Proposition 2.4. *For all sufficiently large* $n > e^e$

$$\sqrt{\mathrm{Var}\left[\log C_{\Omega}(n)\right]} = (\log\log n)(\log\log\log n)\left(1 + O\left(\frac{1}{(\log\log n)^{\frac{1}{3}}}\right)\right), \text{ as } n \to \infty.$$

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

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

We define the sums

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

We have that

$$S_{2,\Omega}(n) = n^2 E_{\Omega}^2(n) - nV_{\Omega}^2(n). \tag{2.3a}$$

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

$$\begin{aligned} S_{2,\Omega}(n) &= \sum_{1\leq j<n} 2\log C_{\Omega}(j)\left(nE_{\Omega}(n) - jE_{\Omega}(j)\right), \qquad (2.3b) \\ &= 2n^2 E_{\Omega}^2(n) - 2 \times \int_{e^e}^{n-1} tE_{\Omega}(t) \times \frac{d}{dt}\left[tE_{\Omega}(t)\right] dt, \\ &= n^2 E_{\Omega}^2(n) - (2n-1)E_{\Omega}^2(n)\left(1 + O\left(\frac{1}{n}\right)\right). \end{aligned}$$

Equations (2.3a) and (2.3b)show that

$$V_{\Omega}^2(n) = 2E_{\Omega}^2(n)\left(1 + O\left(\frac{1}{n}\right)\right), \text{ as } n \to \infty. \tag{2.4}$$

The variance of the function $\log C_{\Omega}(n)$ is given by $\sigma_n^2 = V_{\Omega}^2(n) - E_{\Omega}^2(n)$. Theorem 1.6 applied to the right-hand-side of equation (2.4) completes the proof. □

2.3 Remarks

Asymptotic formulae for the moments of the function $C_{\Omega}(n)$ on the positive integers $n \leq x$ as $x \to \infty$ are required to evaluate the average order of $|g(n)|$. Proofs to evaluate the centralized moments of the former function are not nearly as straightforward as the methods we 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 $0 < R, M < +\infty$ be fixed. An approach to the average order of $C_{\Omega}(n)$ invokes the Selberg-Delange method [25, §II.6.1] [17, §7.4] in evaluating the partial sums of the form

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

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

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

A proof of the average order of the function $(-1)^{\omega(n)} C_{\Omega}(n)$ requires technical arguments that are non-trivial extensions of the proofs in [17, 25]. 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 inverse function with respect to the operation of Dirichlet convolution by

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

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

Remark 3.2 (Motivation)**.** Let $\chi_{\mathbb{P}}(n)$ denote the characteristic function of the primes, suppose that $\varepsilon(n) = \delta_{n,1}$ is the multiplicative identity with respect to Dirichlet convolution, and define the function $\mathbb{1}(n)$ to be identically equal to one for all $n \geq 1$. We find 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

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

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, we may consider the inversion of the following partial sums to study the Mertens function:

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

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 $\mathrm{Re}(s) > \sigma_f$ 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 $\mathrm{Re}(s) > 1$. Whenever $f(1) \neq 0$ the DGF of $f^{-1}(n)$ is $D_f(s)^{-1}$. By equation (3.1) we have

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

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

We recover exactly that [10, 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 $\mathrm{Re}(s) > 1.39944$, we count that

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

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

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

Remark 3.4. 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 $\mathrm{Re}(s) > 1$. We can easily extend the last proof to see that $C_\Omega(n)$ has DGF $(1-P(s))^{-1}$ for all $\mathrm{Re}(s) > 1$ (see Remark 3.6).

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

Lemma 3.5. *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

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

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 $(g*1)(n) = G_m(n)$ where $G_m(n) := (-1)^m(C_m(-) * g)(n)$ is expanded as

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

When $n \geq 2$ and $m := \Omega(n)$, i.e., with these expansions carried out to a maximal depth, we obtain

$$(g*1)(n) = \lambda(n)C_\Omega(n). \qquad (3.6)$$

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

Proof of Proposition 1.5. The result follows from Lemma 3.5, 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

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

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

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. $\square$

Remark 3.6. The following are consequences of Proposition 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 $n \geq 1$, 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 $\sigma > 1$.
- The average order of $|g(n)|$ is given by [13, §18.6]
$$\frac{1}{x} \times \sum_{n \leq x} |g(n)| = \frac{1}{x} \times \sum_{n \leq x} C_{\Omega}(n) \left(\sum_{k \leq \left\lfloor \frac{x}{n} \right\rfloor} \mu^2(k) \right) = \frac{6}{\pi^2} \times \sum_{n \leq x} \frac{C_{\Omega}(n)}{n} \left(1 + O\left(\sqrt{\frac{n}{x}}\right)\right). \tag{3.7c}$$

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

In this section, we motivate and prove a central limit theorem for the distribution of the function $\log C_{\Omega}(n)$.

4.1 Proof of Theorem 1.8

Proof. 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 primes p as $x \to \infty$, we have the 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} [2, §1.2]. In other words, for any prime p and $k \geq 1$ we have that
$$\lim_{x \to \infty} \mathbb{P}\left[C_p^{(x)} = k\right] = \left(1 - \frac{1}{p}\right) \left(\frac{1}{p}\right)^k.$$
- For $n \geq 1$, we use equation (1.6) and Binet's log-gamma formula [21, §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., at prime n), it suffices to restrict our considerations to the cases where $\Omega(n) \geq 2$.

- For $x \geq 2$, let

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

We can write the expansion from equation (4.1) as the difference

$$\log C_\Omega(N(x)) := \Theta_{N(x)} - A_{N(x)} + O(1), \text{ as } x \to \infty.$$

Moreover, 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 $x \geq 1$, we define

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

We calculate that

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

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

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

We say that the Lindeberg condition is satisfied when the following is true for every fixed $\varepsilon > 0$:

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

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

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

- 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 $xe^x = y$. For any $M > 0$, the condition that $k \log k > M\sigma_x$ is true when $k > \frac{M\sigma_x}{W(M\sigma_x)} \sim M \log\log x$ as $x \to \infty$ [6]. The inequality $t \log(1+t) \geq \left(t + \frac{1}{2}\right)\log(1+t) - t$ is satisfied for all real $t > 0$.
- Suppose that $\varepsilon \in (0,1)$ is fixed and let $\mathcal{I}(\varepsilon, x) := [(1-\varepsilon)\log\log x, (1+\varepsilon)\log\log x] \bigcap \mathbb{Z}^+$. We have that

$$\begin{aligned}
\widetilde{E}_\Omega(\varepsilon, x) &\ll \frac{1}{\sigma_x^2} \times \sum_{n \leq x} \mathbb{E}\left[\log^2 C_\Omega(n)\left(\chi_{x,n} - \frac{1}{x}\right)^2 \mathbb{1}_{\{|S_x - \mu_x| > \varepsilon\sigma_x\}}\right] \\
&\ll \frac{1}{\sigma_x^2} \times \sum_{k \in \mathcal{I}(\varepsilon, x)} \log^2(k!) \times \mathbb{P}[\Omega(N(x)) = k] \times \sum_{n \leq x} \mathbb{E}\left(\chi_{x,n} - \frac{1}{x}\right)^2 \\
&\ll \frac{1}{x\sigma_x^2} \times \sum_{k \in \mathcal{I}(\varepsilon, x)} \log^2(k!) \times \mathbb{P}[\Omega(N(x)) = k] \times \left(1 - \frac{1}{x}\right).
\end{aligned} \tag{4.4}$$

Proposition 2.4, Theorem A.1 and equation (4.4) show that

$$\widetilde{E}_{\Omega}(\varepsilon, x) \ll \left|(\log x)^{-\varepsilon-(1-\varepsilon)\log(1-\varepsilon)} - (\log x)^{\varepsilon-(1+\varepsilon)\log(1+\varepsilon)}\right|.$$

The function $f_{\pm}(t) := \pm t - (1 \pm t)\log(1 \pm t)$ is strictly negative and monotone decreasing for $t \in (0,1)$. It follows that equation (4.2) is satisfied for all $\varepsilon \in (0,1)$. We can modify the argument given above using equation (D.5b) from the appendix to show that (4.2) is also satisfied whenever $\varepsilon \in [1, \infty)$. We conclude from the Lindeberg CLT that (4.3) holds. □

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 $n_1, n_2 \geq 2$

$$\mathcal{E}[n_1] = \mathcal{E}[n_2] \implies C_{\Omega}(n_1) = C_{\Omega}(n_2) \text{ and } g(n_1) = g(n_2). \tag{4.5}$$

This property shows that there is a deep structure to these functions connected to the prime divisors of the positive integers $n \geq 2$. On the other hand, since the multiplicative function $\mu^2(n)$ and the strongly additive functions $\omega(n)$ and $\Omega(n)$ also satisfy their analog to equation (4.5). Thus, this property alone is insufficient to predict the CLT type result for functions of this type we see in Theorem 1.8.

More intuition about why the distribution of $\log C_{\Omega}(n)$ obeys 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)

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

In the sense that multiple (Dirichlet) convolutions reflect a qualitative smoothing operation on average, the CLT statement for $\omega(n)$ above should predict a smooth limiting distribution. Incidentally, equation (1.6) shows that the normalized function $\frac{C_{\Omega}(n)}{(\Omega(n))!}$ is multiplicative (cf. [7]).

5 Applications to the Mertens function

In this section, we prove Theorem 1.4. The new formulas precisely connect the Mertens function with partial sums of positive unsigned arithmetic functions whose summands are weighted by the sign of $\lambda(n)$.

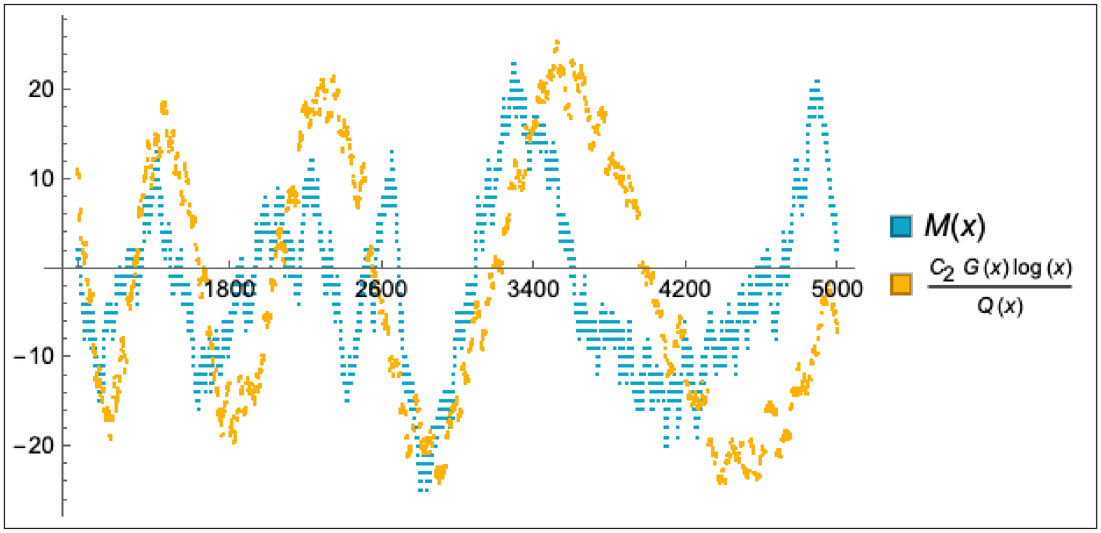
Definition 5.1. The summatory functions of $g(n)$ and $|g(n)|$, respectively, are defined for all $x \geq 1$ by the partial sums

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

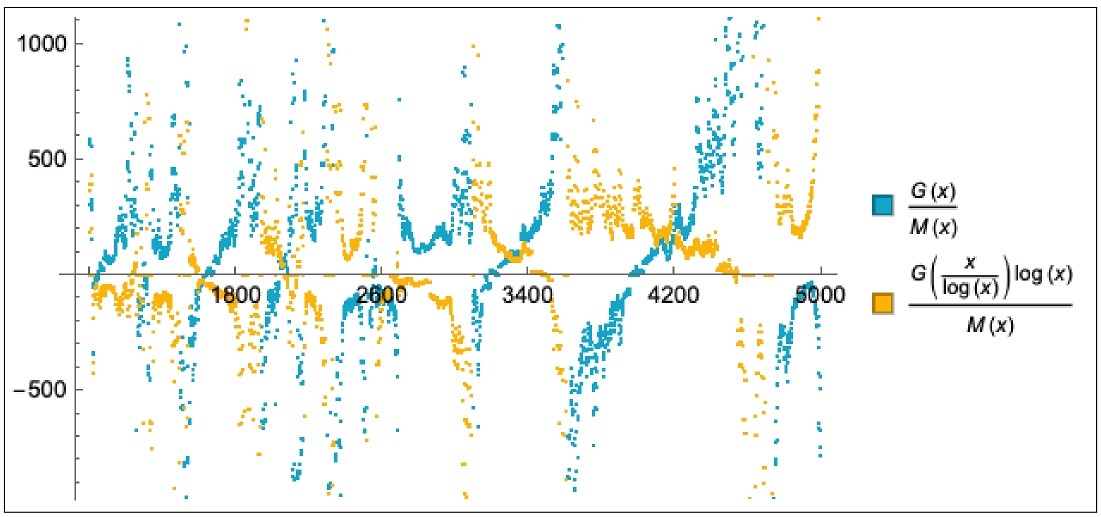
5.1 Proofs of the new formulas

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

$$M(x) = \sum_{k=1}^{x} \left(1 + \pi\left(\left\lfloor \frac{x}{k} \right\rfloor\right)\right) g(k)$$



(a)



(b)

Figure 5.1

$$= G(x) + \sum_{k=1}^{\frac{x}{2}} \pi\left(\left\lfloor \frac{x}{k} \right\rfloor\right) g(k)$$

$$= G(x) + G\left(\left\lfloor \frac{x}{2} \right\rfloor\right) + \sum_{k=1}^{\frac{x}{2}-1} \left(\pi\left(\left\lfloor \frac{x}{k} \right\rfloor\right) - \pi\left(\left\lfloor \frac{x}{k+1} \right\rfloor\right)\right) \times G(k).$$

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. $\square$

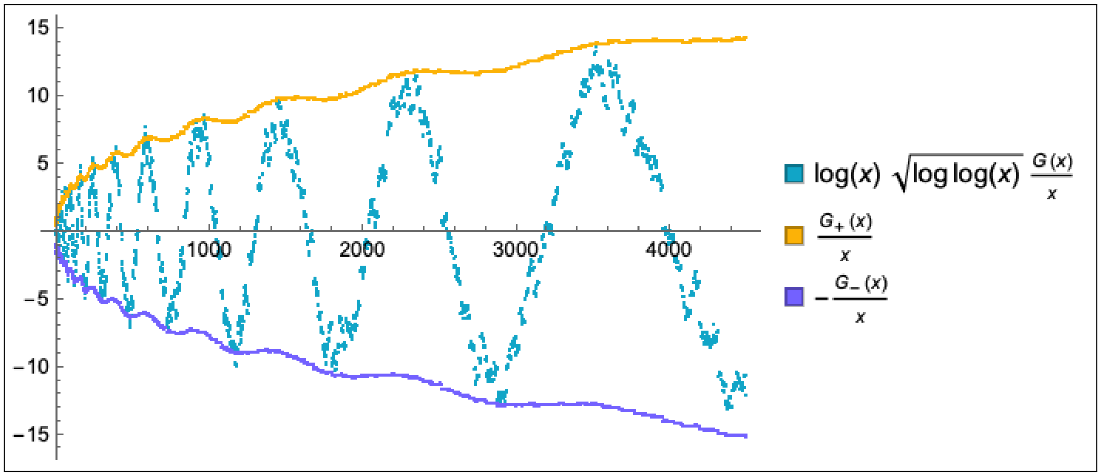
Proof of (1.5c) *of Theorem 1.4.* Lemma 3.5 shows that

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

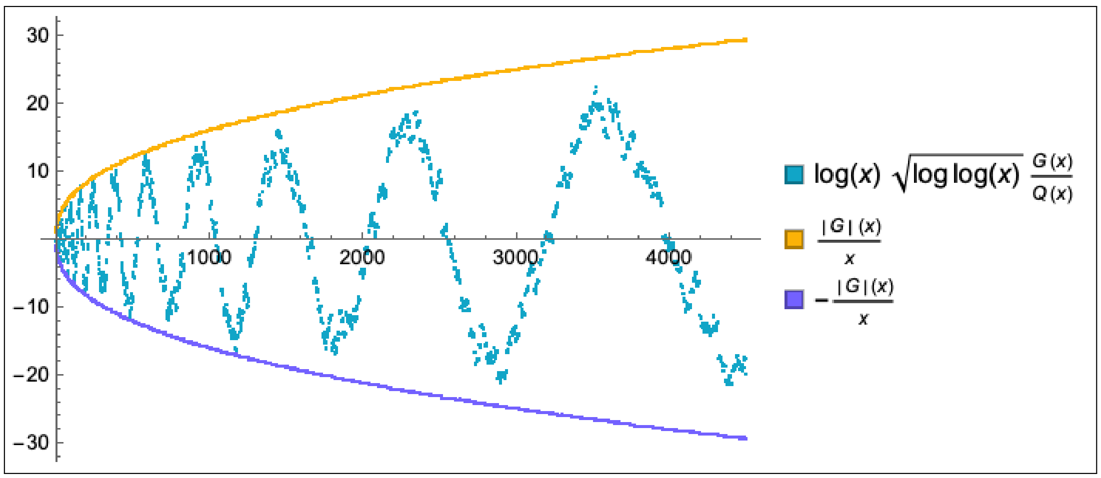
The identity in (3.1) implies

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

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



(a)



(b)

Figure 5.2

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 related auxiliary partial sums. These plots showcase interesting phenomena observed for small x:

- In Figure 5.1, we plot comparisons of $M(x)$ to scaled forms of $G(x)$ for $x \leq 5000$. The absolute constant is $C_2 := \frac{\pi^2}{6}$. 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 $Q(x)$ in **(b)** has the same definition as in Figure 5.1 above. This plot suggests that for large x

$$|G(x)| \ll \frac{|G|(x)}{(\log x)\sqrt{\log\log x}} = \frac{1}{(\log x)\sqrt{\log\log x}} \times \sum_{n \leq x} |g(n)|, \text{ as } x \to \infty.$$

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 distributions of the unsigned functions $C_\Omega(n)$ and $|g(n)|$ provide new information about $M(x)$ via the formulas proved in Theorem 1.4. These formulas are expressed in terms of weighted partial sums with terms signed by $\lambda(n)$

6.2 Discussion of the new results

6.2.1 Randomized models of the Möbius function

Some natural 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 $n \geq 1$

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

That is, the sequence $\{X_n\}_{n\geq 1}$ provides a randomized model of the values of $\mu(n)$ on average. We may approximate limiting properties of the partial sums as $M(x) \cong \overline{S}_x$ where $\overline{S}_x := \sum_{n\leq x} X_n$ for all $x \geq 1$. This viewpoint models predictions of certain limiting asymptotic behavior of the Mertens function such as [5, Thm. 9.4; §9]

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

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 [14, 15]

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

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 equation (4.5) suggests that the unsigned weights on $\lambda(n)$ in the new formulas from the theorem may yield new insights compared to formulas like equation (6.1).

6.2.3 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. [12, 9, 24]). Indeed, $\lambda(n) = \mu(n)$ for all squarefree $n \geq 1$ so that $\lambda(n)$ agrees with $\mu(n)$ at most large n. It can be inferred 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 $M(x)$ according to the rationale in the following points:

- Breakthrough work in recent years due to Matomäki, Radziwiłł and Soundararajan to bound multiplicative functions in short intervals has proven fruitful when applied to $\lambda(n)$ [23, 16]. 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 limiting distributions for the distinct values of $\omega(n)$, $\Omega(n)$, and their difference, over $n \leq x$ as $x \to \infty$ [17, cf. §2.4; §7.4];
- The function $\lambda(n)$ is completely multiplicative. Hence, the values of the non-zero function $\lambda(n)$ may be more regular in certain ways on the integers $n \geq 4$ for which $\mu(n) = 0$.

Acknowledgements

We credit Appendix B to correspondence with Gergő Nemes from the Alfréd Rényi Institute of Mathematics.

References

[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] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert w-function. *Advances in Computational Mathematics*, 5:329–359, 1996. https://cs.uwaterloo.ca/research/tr/1993/03/W.pdf.

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

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

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

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

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

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

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

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

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

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

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

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

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

[20] 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.

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

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

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

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

[25] 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

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

and

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

where $B_1 \approx 0.261497$ and $B_2 \approx 1.03465$ are absolute constants [13, §22.10]. The next theorems from [17, §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. [8, 4]).

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

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

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 [17, Thm. 7.20; §7.4]. It uses an adaptation of Rankin's method in combination with [17, 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 $1 \leq k \leq R \log\log x$

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

Proof. The proof of this theorem is given in [17, Thm. 7.19; §7.4]. The notation $\widehat{\pi}_k(x)$ is distinct from that other references that use $N_k(x)$ and $\sigma_k(x)$, respectively [17, Eqn. (7.61)] [25, 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

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

Proof. We can extend the proofs in [17, §7] to obtain analogous results on the distribution of $\omega(n)$. This result is cited as an exercise in [17]. □

B The upper incomplete gamma function

Definition B.1. The (upper) incomplete gamma function is defined by [21, §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 upper incomplete gamma function for real $a,z>0$ by $Q(a,z) := \Gamma(a,z)\Gamma(a)^{-1}$.

The following properties are known [21, §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 [18]

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

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

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

For fixed $\rho>0$, the sequence $\{b_n(\rho)\}_{n\geq 0}$ satisfies a Rodrigues type formula of the form [19, Thm. 1.1]

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

If $z,a\to\infty$ with $z=\rho a$ for some $\rho>1$ such that $(\rho-1)^{-1} = o\left(\sqrt{a}\right)$, then [18]

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

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

- *If $\rho \in (0,1)$, then*
$$\Gamma(a,z) = \Gamma(a) + O_\rho\left(z^{a-1}e^{-z}\right). \tag{B.2a}$$
- *If $\rho>1$, then*
$$\Gamma(a,z) = \frac{z^{a-1}e^{-z}}{1-\rho^{-1}} + O_\rho\left(z^{a-2}e^{-z}\right). \tag{B.2b}$$
- *If $\rho > W(1) > 0.56714$, then*
$$\Gamma(a, ze^{\pm\pi\imath}) = -e^{\pm\pi\imath a}\frac{z^{a-1}e^{z}}{1+\rho^{-1}} + O_\rho\left(z^{a-2}e^{z}\right). \tag{B.2c}$$

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 $z \to \infty$ [20, 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 [19, 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 [19, §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 [18, 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. $\square$

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

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

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

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

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

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 [21, cf. Eq. (5.11.8)] shows that

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

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 expansion $a^{an} = e^{an \log a}$, 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 $r(1) \neq 0$. The following holds 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 for $S_{r*h}(x)$ is well known in the references. The second formula is justified directly using summation by parts as follows [21, §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 $[\texttt{cond}]_\delta \in \{0, 1\}$ evaluates to one precisely when `cond` is true and to zero otherwise.

We form the invertible matrix of coefficients (denoted by $\hat{R}$ below) by defining

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

and

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

If we let $\hat{R} := (R_{x,j})$, then the next matrix is expressed by applying an invertible shift operation as

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

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.

For any fixed $N \geq 1$, 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 follows:

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

The target matrix is expressed by

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

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

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

The summatory function $H(x)$ is expressed 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 another inversion formula by adapting our argument used to prove (C.1) above. This leads to an 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* $x \to \infty$

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

Proof. Equation (1.6) shows that

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

The sum on the right-hand-side o 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 equation (D.1) is obtained by summing over only the squarefree $n\leq x$, i.e., the $n\leq x$ such that $\mu^2(n)=1$. That is, we claim that

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

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 $p|n$ [22, A007913]. For integers $x\geq 1$ and $1\leq k\leq \log_2(x)$, define the sets

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

Recall that our goal is to show that the sums of $\log C_\Omega(n)$ at the non-squarefree $n\leq x$ corresponds to an error term on the right-hand-side of equation (D.2). Then the idea behind the definition in the previous 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 non-empty sequence of k *positive* integers $\{\varpi_j\}_{1\leq j\leq k}$ such that $n\in\mathcal{S}_k\left(\{\varpi_j\}_{j=1}^k; x\right)$. Hence, we need to bound the growth of this function at non-empty sets $\{\varpi_j\}_{j=1}^k\subseteq\mathbb{Z}^+$ for $k\geq 1$.

Let the function

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

The special case where $\mathcal{W}_k := \{\varpi_j^*\}_{1\leq j\leq k}\equiv\{0,1\}$ is the set with unit value of multiplicity exactly one is denoted by

$$\widehat{T}_k(x) := \mathcal{N}_k\left(\mathcal{W}_k; x\right).$$

If $n\in[2,x]$ 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 $1\leq j\leq k$. For any $k\geq 1$ and non-empty set $\{\varpi_j\}_{1\leq j\leq k}\subset\mathbb{Z}^+$ of k elements we have

$$\mathcal{N}_k\left(\{\varpi_j\}_{j=1}^k; x\right) \ll \widehat{T}_k(x).$$

We claim that (proof below)

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

Suppose that equation (D.3a) is true. Theorem A.2, Theorem A.3 and equation (D.3a) imply that

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

The upper bound on $\widehat{T}_k(x)$ in equation (D.3b) shows that the sum of denominator differences from (1.6) we subtract from the main term over the squarefree $n\leq x$ is asymptotically insubstantial. That is, we have proved that

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

Recall the next two famous asymptotic formulae:

1. As $x \to \infty$ [13, §22.4]

$$\pi(x) = \frac{x}{\log x} \times \left(1 + O\left(\frac{1}{\log x}\right)\right).$$

2. A theorem of Mertens states that as $x \to \infty$ [13, §22.7–22.8]

$$\sum_{p \leq x} p^{-1} \sim \log \log x.$$

Proof of Equation (D.3a). The bound can be proved by induction on $k \geq 1$ using the inductive hypothesis

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

The case where $k := 1$ is evaluated 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 and the bounds in the next equations show that equation (D.3a) 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

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

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

$$\widehat{U}_{k+1}(x) := \frac{2^{2^{-k}}}{(\log \log x)^{2^{-k}} (\log x)^{3 \cdot 2^{-(k+1)}}} \ll 1, \text{ as } x \to \infty. \tag{D.4}$$

For any fixed finite large x, we have that $1 \leq k \leq \log_2(x)$. In particular, we have that $\frac{1}{x} \leq 2^{-k} \leq \frac{1}{2}$ for all k and large x. We can then establish the following to prove that equation (D.4) holds:

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

The proof of equation (D.3a) is completed by applying Theorem A.3. □

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 $x \geq 19$, consider the following partial sums[1]:

$$\begin{aligned}
L_{M,\Omega}(x) &:= \sum_{\substack{n \leq x \\ \Omega(n) \leq \frac{3}{2} \log \log x}} \log C_\Omega(n), \\
L_{E,\Omega}(x) &:= \sum_{\substack{n \leq x \\ \Omega(n) > \frac{3}{2} \log \log x}} \log C_\Omega(n).
\end{aligned}$$

[1] Note that the choice of the constant $\frac{3}{2}$ is arbitrary. We can obtain the same result by splitting the upper (lower) bound on $\Omega(n)$ at $r \log \log x$ for any fixed $r > 1$.

We claim that the main term is given by

$$L_{M,\Omega}(x) = x(\log\log x)(\log\log\log x)\left(1 + O\left(\frac{1}{\sqrt{\log\log x}}\right)\right). \tag{D.5a}$$

To bound the error term, we claim that

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

The proofs of equations (D.5a) and (D.5b) are completed below. Equations (D.5a) and (D.5b) yield the conclusion of Theorem 1.6 since

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

Proof of Equation (D.5a). Lemma D.1 and Theorem A.2 from the appendix show that

$$L_{M,\Omega}(x) = \frac{x}{\log x} \times \sum_{1 \leq k \leq \frac{3}{2}\log\log x} \mathcal{G}\left(\frac{k-1}{\log\log x}\right) \frac{(\log\log x)^{k-1}}{(k-1)!} \log(k!) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right),$$

where

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

For any $j \geq 0$, Binet's formula for the log-gamma function is stated as [21, §5.9(i)]

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

where $z! := \Gamma(1+z)$ for any real-valued $z \geq 0$. Let the function

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

Binet's formula shows that

$$L_{M,\Omega}(x) = \frac{x}{\log x} \times \sum_{1 \leq k \leq \frac{3}{2}\log\log x} \mathcal{G}\left(\frac{k-1}{\log\log x}\right) g(x,k) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right).$$

The Euler-Maclaurin summation (EM) formula [11, §9.5] shows that for each fixed integer $m \geq 1$ and function $f(t)$ that is m times continuously differentiable on $(0,\infty)$ [22, A000367; A002445]

$$\begin{aligned} L_{M,\Omega}(x) = \frac{x}{\log x} \times \Bigg(& \int_1^{\frac{3}{2}\log\log x} \mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)dt + \frac{1}{2}\mathcal{G}\left(\frac{3}{2} - \frac{1}{\log\log x}\right) g\left(x, \frac{3}{2}\log\log x\right) - \frac{g(x,1)}{2} \\ & + O\left(\sum_{k=1}^m \frac{B_k}{k!} \times \frac{\partial^{(k-1)}}{\partial t^{(k-1)}}\left[\mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)\right]_{t=1}^{t=\frac{3}{2}\log\log x} + \widehat{R}_m[\mathcal{G}\cdot g]\right)\Bigg) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right). \end{aligned}$$

The degree-m remainder term in the previous equation is bounded by

$$\left|\widehat{R}_m[\mathcal{G}\cdot g]\right| = O\left(\frac{1}{m!} \times \int_1^{\frac{3}{2}\log\log x} B_m(\{t\}) f^{(m)}(t)dt\right).$$

When $f(t) \equiv f_x(t) := \mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)$, we denote the degree-$m$ EM formula error term by

$$E_m(x) := \sum_{k=1}^m \frac{B_k}{k!} \times \frac{\partial^{(k-1)}}{\partial t^{(k-1)}}\left[\mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t)\right]_{t=1}^{t=\frac{3}{2}\log\log x} + \frac{1}{m!} \times \int_1^{\frac{3}{2}\log\log x} |B_m(\{t\})| f_x^{(m)}(t)dt$$

It suffices to choose $m := 1$ in the expansion above. Specializing to the case where $m := 1$ yields the upper bound

$$|E_1(x)| \ll \frac{(\log x)(\log\log\log x)}{\sqrt{\log\log x}} + \underbrace{\int_1^{\frac{3}{2}\log\log x} t^2 \log(1+t) \times \frac{(\log\log x)^t}{(t+1)!} dt}_{:=I_1(x)}, \text{ as } x \to \infty. \tag{D.6a}$$

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

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

The maximum in the previous equations is attained when $t = \frac{3}{2}\log\log x$. This shows that

$$\begin{aligned}\frac{x}{\log x} \times |E_1(x)| \ll \frac{x}{\log x} \times |I_1(x)| &\ll x(\log x)^{\frac{1}{2}(1-3\log(\frac{3}{2}))}(\log\log x)^{\frac{3}{2}}(\log\log\log x) \\ &= o\left(x\sqrt{\log\log x}(\log\log\log x)\right), \text{ as } x \to \infty.\end{aligned} \tag{D.6b}$$

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

$$\int_1^{\frac{3}{2}\log\log x} \mathcal{G}\left(\frac{t-1}{\log\log x}\right) g(x,t) dt = \mathcal{G}\left(\frac{c-1}{\log\log x}\right) \times \int_1^{\frac{3}{2}\log\log x} g(x,t) dt. \tag{D.7a}$$

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

$$c(y) := \inf\left\{c \in \left[1, \frac{3}{2}\log\log y\right] : \text{ equation (D.7a) holds}\right\}. \tag{D.7b}$$

Let $B_0^*(x) := \mathcal{G}\left(\frac{c(x)-1}{\log\log x}\right)$. We can apply the EM formula again to see that

$$\begin{aligned}L_{M,\Omega}(x) &= \frac{x}{\log x} \times \left(\sum_{1 \le k \le \frac{3}{2}\log\log x} B_0^*(x) g(x,k) + \frac{1}{2}\left(1 - B_0^*(x)\right)\mathcal{G}\left(\frac{3}{2} - \frac{1}{\log\log x}\right) g\left(x, \frac{3}{2}\log\log x\right)\right. \\ &\left. + O\left(1 + \sum_{k=1}^m \frac{B_k}{k!} \times \frac{\partial^{(k-1)}}{\partial t^{(k-1)}}\left[\left(1 + \mathcal{G}\left(\frac{t-1}{\log\log x}\right)\right) g(x,t)\right]_{t=1}^{t=\frac{3}{2}\log\log x} + \widehat{R}_m[(1+\mathcal{G}) \cdot g]\right)\right) \times \left(1 + O\left(\frac{1}{\log\log x}\right)\right).\end{aligned} \tag{D.8}$$

For $m := 1$, the EM formula error term satisfies $\left|\widehat{R}_m[(1+\mathcal{G}) \cdot g]\right| \ll |I_1(x)|$ for all sufficiently large large x. Equation (D.6b) provides bounds on $|E_1(x)|$ and $|I_1(x)|$.

We have two remaining steps to establish equation (D.5a):

- **(i)** To show that the sums on the right-hand-side of equation (D.8) 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 a limiting constant with $B_0^*(x) \xrightarrow{x\to\infty} 1$.

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

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

That is, we have by Proposition B.2 that

$$
\begin{aligned}
&\frac{x}{\log x} \times \sum_{1 \leq k \leq \log\log x} g(x,k) \\
&\quad = A_x\left(\frac{3}{2}\log\log x\right) f\left(\frac{3}{2}\log\log x\right) - (\log\log x) \times \int_{\frac{1}{\log\log x}}^{\frac{3}{2}} A_x(\alpha \log\log x) f'(\alpha\log\log x) d\alpha + O\left(\frac{x}{\log x}\right) \\
&\quad = \frac{3x}{2}(\log\log x)(\log\log\log x)\left(1 + O\left(\frac{1}{\log\log x}\right)\right) + \int_1^{\frac{3}{2}} f'(\alpha\log\log x)\left(1 + O\left(\frac{1}{\sqrt{\log\log x}}\right)\right) d\alpha \\
&\qquad\qquad + O\left(x\sqrt{\log\log x} \times \int_0^1 f'(\alpha\log\log x) \times (\log x)^{\alpha-1} d\alpha\right) \\
&\quad = x(\log\log x)(\log\log\log x)\left(1 + O\left(\frac{1}{\sqrt{\log\log x}}\right)\right). && \text{(D.9)}
\end{aligned}
$$

We have the following observations:

$$
\begin{aligned}
\mathcal{G}\left(\frac{3}{2} - \frac{1}{\log\log x}\right) &= \mathcal{G}\left(\frac{3}{2}\right)\left(1 + O\left(\frac{1}{\log\log x}\right)\right), \text{ as } x \to \infty, && \text{(D.10a)} \\
g\left(x, \frac{3}{2}\log\log x\right) &\ll (\log x)^{\frac{3}{2}(1-\log(\frac{3}{2}))}\sqrt{\log\log x}(\log\log\log x) && \text{(D.10b)} \\
&= o\left((\log x)\sqrt{\log\log x}(\log\log\log x)\right), \text{ as } x \to \infty.
\end{aligned}
$$

Equations (D.8) and (D.9) and (D.10) show that

$$
L_{M,\Omega}(x) = B_0^*(x) x (\log\log x)(\log\log\log x)\left(1 + O\left(\frac{1}{\sqrt{\log\log x}}\right)\right), \text{ as } x \to \infty. \tag{D.11}
$$

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.7b) 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.7a) 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 $x \geq 19$. 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.11) (e.g., with step **(i)**) shows that equation (D.5a) holds. □

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

$$
\log C_\Omega(n) \ll \Omega(n)\log\Omega(n), \text{ for } n \leq x, \text{ as } x \to \infty. \tag{D.12}
$$

The bound for $\log C_\Omega(n)$ in (D.12) 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.12) to Binet's log-gamma formula. We use the following notation to express this ratio as

$$
\mathcal{R}(n,j) := \frac{\Omega(n)\log\Omega(n)}{\left(j + \frac{1}{2}\right)\log(1+j) - j}.
$$

Numerical methods show that $\mathcal{R}(n, \Omega(n))$ is absolutely bounded for all $16 \leq n \leq x$ as $x \to \infty$. The global extrema of this function on the positive integers are each attained as $\mathcal{R}(n_\ell, j_\ell), \mathcal{R}(n_u, j_u)$ for finite integers $2 \leq n_\ell, n_u < +\infty$ and $j_\ell, j_u = \Omega(n_\ell), \Omega(n_u) \in [1, 12]$.

The analog of the Erdős-Kac theorem for the function $\Omega(n)$ is given by [17, Thm. 7.21; §7.4]

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

Then we have that as $x \to \infty$

$$\begin{aligned} L_{E,\Omega}(x) &\ll \sum_{\substack{n \leq x \\ \Omega(n) \geq \frac{3}{2}\log\log x}} \Omega(n) \log \Omega(n) \ll \left(\sum_{\substack{n \leq x \\ \Omega(n) \geq \frac{3}{2}\log\log x}} \Omega(n)^2\right)^{\frac{1}{2}} \times \left(\sum_{\substack{n \leq x \\ \Omega(n) \geq \frac{3}{2}\log\log x}} \log^2 \Omega(n)\right)^{\frac{1}{2}} \\ &\ll (\log\log x) \times \left(\frac{x}{\sqrt{\log\log x}} \times \int_{\log\log x}^{\log_2(x)} t^2 e^{-\frac{(t-\log\log x)^2}{2\log\log x}} dt\right)^{\frac{1}{2}} \times \#\left\{n \leq x : \Omega(n) \geq \frac{3}{2}\log\log x\right\}^{\frac{1}{2}}. \end{aligned}$$

The change of variable $u := \frac{t-\log\log x}{\sqrt{\log\log x}}$ and Theorem A.1 applied to the previous equation show that

$$L_{E,\Omega}(x) \ll x(\log\log x)^2 (\log x)^{\frac{1}{4}\left(1-3\log\left(\frac{3}{2}\right)\right)} \ll x(\log\log x)^2(\log x)^{-0.0540987}. \qquad \square$$