Weighted Erdős–Kac theorems via computing moments

by

Kai (Steve) Fan

Abstract. By adapting the moment method developed by Granville and Soundararajan (2007), Khan, Milinovich and Subedi (2022) obtained a weighted version of the Erdős–Kac theorem for $\omega(n)$ with multiplicative weight $d_k(n)$, where $\omega(n)$ denotes the number of distinct prime divisors of a positive integer n, and $d_k(n)$ is the k-fold divisor function with $k \in \mathbb{N}$. In the present paper, we generalize their method to study the distribution of additive functions $f(n)$ weighted by nonnegative multiplicative functions $\alpha(n)$ in a wide class. In particular, we establish uniform asymptotic formulas for the moments of $f(n)$ with suitable growth rates. We also prove a qualitative result on the moments which extends a theorem of Delange and Halberstam (1957). As a consequence, we obtain a weighted analogue of the Kubilius–Shapiro theorem.

Contents

1. Introduction. The celebrated Erdős–Kac theorem, first proved by Erdős and Kac [15] in 1940, states that if $\omega(n)$ denotes the number of distinct prime divisors of a positive integer n, then

$$(1.1) \qquad \lim_{x \to \infty} \frac{1}{x} \cdot \# \left\{ n \le x \colon \frac{\omega(n) - \log \log n}{\sqrt{\log \log n}} \le V \right\} = \Phi(V)$$

[1]

for any given $V \in \mathbb{R}$, where

$$\Phi(V) := \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{V} e^{-v^2/2}\,dv$$

is the cumulative distribution function of the standard Gaussian distribution. This statistical result is a direct upgrade of an earlier theorem of Hardy and Ramanujan on the *normal order* of ω (see [21] and [22, Theorem 431]), which asserts that given any $\epsilon > 0$, the inequality $|\omega(n) - \log\log n| < \epsilon \log\log n$ holds for all but $o(x)$ values of $n \le x$. In fact, Erdős and Kac proved in the same paper a more general result in which the function ω is replaced by any strongly additive function f that is bounded on primes and admits an unbounded "variance" $\sum_{p \le x} f(p)^2/p$. Recall that an arithmetic function $f \colon \mathbb{N} \to \mathbb{C}$ is *additive* if $f(mn) = f(m) + f(n)$ for all positive integers $m, n \in \mathbb{N}$ with $\gcd(m,n) = 1$, and it is *strongly additive* if it also satisfies $f(p^\nu) = f(p)$ for all prime powers p^ν. Thus, strongly additive functions are completely determined by their values at primes, which makes them a particularly nice subclass of additive functions. In fact, it can be shown that (1.1) also holds for $\Omega(n)$ in place of $\omega(n)$, where $\Omega(n)$ denotes the total number of prime factors of n, counting multiplicity, by exploiting the fact that ω and its cousin Ω do not differ very much on average:

$$(1.2) \qquad \sum_{n \le x} (\Omega(n) - \omega(n)) = O(x).$$

The original proof of the Erdős–Kac theorem (1.1) combines the central limit theorem with Brun's sieve and is quite complicated. Later, LeVeque [25, Theorem 1] introduced some modifications and obtained a quantitative version of (1.1) with a rate of convergence $O(\log\log\log x/\sqrt{\log\log x})$. Using a deep analytic approach, Rényi and Turán [28] improved upon LeVeque's result with a rate of convergence $O(1/\sqrt{\log\log x})$, which is best possible in the sense that it cannot be improved to $o(1/\sqrt{\log\log x})$ without loss of uniformity in V.

A third approach to (1.1), first suggested by Kac [23], examines the moments of ω. In probability theory, the moments of a random variable X often provide valuable information about its distribution. For example, an application of Markov's inequality yields $\mathbb{P}(|X| > c) \le \mathbb{E}|X|^k/c^k$. Given all the moments $\mathbb{E}|X|^k < \infty$, one may select k that minimizes this tail estimate. If X happens to obey a Gaussian law, then it is completely determined by its moments, a direct consequence of [3, Theorem 30.1] or [10, Theorem 3.3.26]. Consequently, by [3, Theorem 30.2], one reduces the proof of (1.1) to that of the asymptotic formula

$$(1.3) \qquad \frac{1}{x} \sum_{n \le x} (\omega(n) - \log\log x)^m = (\mu_m + o(1))(\log\log x)^{m/2}$$

for every $m \in \mathbb{N}$. Here μ_m is the mth moment of a standard Gaussian distribution given by

$$\mu_m = \begin{cases} m!/m!! & \text{if } 2 \mid m, \\ 0 & \text{otherwise.} \end{cases}$$

The case $m = 1$ follows from Mertens' second theorem [22, Theorem 427], and the case $m = 2$ was settled by Turán [34]. Early proofs of (1.3) for all k are due to Delange [6] in 1953 and Halberstam [18] in 1955, both of which are very complicated. Later, Delange [7] provided an elementary proof of (1.3) for strongly additive functions. Halberstam's proof was also simplified and rendered more transparent by Billingsley [2] in 1969. In 2007, Granville and Soundararajan [17] derived asymptotic formulas for the moments which hold uniformly in the range $m \le (\log \log x)^{1/3}$. Their method, which may be viewed to some extent as a clever and efficient repacking of Billingsley's, is so flexible that it can be modified to study the distribution of values of additive functions in a rather general sieve-theoretic framework.

More generally, one can study the distribution of values of $\omega(n)$ weighted by certain nonnegative multiplicative functions $\alpha(n)$. For instance, Elliott [13] showed, based on the Landau–Selberg–Delange method, that

$$\lim_{x \to \infty} \left(\sum_{n \le x} d(n)^c \right)^{-1} \sum_{\substack{n \le x \\ \omega(n) \le 2^c \log \log x + V \sqrt{2^c \log \log x}}} d(n)^c = \Phi(V)$$

for any given $c \in \mathbb{R}$ and $V \in \mathbb{R}$, where $d(n)$ denotes the number of positive divisors of n. Building on the method of Granville and Soundararajan, Khan, Milinovich and Subedi [24] recently proved an analogue for the weight $d_k(n)$ with mean $k \log \log x$ and variance $k \log \log x$, where $d_k(n) := \#\{(a_1, \ldots, a_k) \in \mathbb{N}^k : a_1 \cdots a_k = n\}$. There is now a vast literature on weighted versions of the Erdős–Kac theorem with general weights, including the early work of Alladi [1] and the more recent works of Elboim and Gorodetsky [11] and Tenenbaum [32, 33]. Alladi made use of Halberstam's approach to prove weighted Erdős–Kac type results for strongly additive functions with the weights being the characteristic functions of the subsets of $\mathbb{N}$ which satisfy certain sieve type conditions. On nonnegattive multiplicative weights α, Elboim and Gorodetsky [11, Theorem 1.1] generalized Billingsley's proof [2] to handle the distribution of $\Omega(n)$ weighted by those α having constant mean values and satisfying certain growth conditions, while Tenenbaum's result [32, Corollary 2.5] was proved by means of charactersitic functions and allows for general additive functions and a large class of multiplicative weights α with the property that $\alpha(p) = O(p^{\sigma_0 - 1})$ for some constant $\sigma_0 > 0$.

The main purpose of this paper is to establish weighted versions of the Erdős–Kac theorem by pushing the method of moments of Granville and Soundararajan to its limit. Our work is the first to apply this method to prove weighted Erdős–Kac theorems with general additive functions and multiplicative weights. We obtain uniform estimates for moments of strength comparable to that of the original estimate of Granville and Soundararajan for $\omega(n)$. With our emphasis on the strength of the method, we have refrained from pursuing the most general theorems at the risk of complicating our exposition. Despite this compromise, our results have some advantages over the results in [11, 32, 33]. Our approach is elementary and flexible, and it can be applied to handle certain arithmetic functions of special interest which were studied previously by different methods. Some examples are discussed in the comment below Corollary 2.5.

Definitions and notation. We introduce some terminology and notation that will be adopted throughout this paper without further clarification. Given any real- or complex-valued functions $f(x)$ and $g(x)$ with a common domain $\mathcal{D} \subseteq \mathbb{R}$, we shall use Landau's big-O notation $f(x) = O(g(x))$ and Vinogradov's notation $f(x) \ll g(x)$ interchangeably to mean that there exists an absolute constant $C > 0$ such that $|f(x)| \leq C|g(x)|$ for all $x \in \mathcal{D}$. Likewise, we shall use the notation $f(x) \gg g(x)$ interchangeably with $g(x) = O(f(x))$. If $f(x) = O(g(x))$ and $g(x) = O(f(x))$ hold simultaneously, then we adopt the shorthand notation $f(x) \asymp g(x)$. If $\mathcal{D}$ contains a neighborhood of ∞, then we write $f(x) = o(g(x))$ when $f(x)/g(x) \to 0$ as $x \to \infty$ and $f(x) \sim g(x)$ when $f(x)/g(x) \to 1$ as $x \to \infty$. We shall occasionally make use of the characteristic function $\epsilon_{a,b}$ of the condition $a \neq b$ for any $a, b \in \mathbb{R}$. Equivalently, $\epsilon_{a,b} = 1 - \delta_{a,b}$, where $\delta_{a,b}$ is the Kronecker delta function.

Throughout, the letter p denotes a prime, and we write $\pi(x)$ for the prime counting function, namely, $\pi(x) = \sum_{p \leq x} 1$. For any $x \in \mathbb{R}$, we write $\lfloor x \rfloor$ for the integer part of x, and $\lceil x \rceil$ for the least integer $\geq x$. For every $n \in \mathbb{N}$, denote by $P^-(n)$ and $P^+(n)$ the least and the greatest prime factor of n, respectively, with the convention that $P^-(1) = \infty$ and $P^+(1) = 1$. We say that $n \in \mathbb{Z} \setminus \{0\}$ is *squareful* if for any prime $p \mid n$ one has $p^2 \mid n$. Given any prime power p^ν, the relation $p^\nu \, \| \, n$ means that $p^\nu \mid n$ but $p^{\nu+1} \nmid n$. In addition, we denote by R_n the *radical* of n, i.e., $R_n := \mathrm{rad}(n) = \prod_{p \mid n} p$. Finally, we write $\binom{m}{m_1, \ldots, m_k} := \frac{m!}{m_1! \cdots m_k!}$ for the multinomial coefficient of shape $(m_1, \ldots, m_k)$ of size $m = m_1 + \cdots + m_k$.

2. Main results. The weights $\alpha \colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ that we shall consider in this paper form a nice subclass $\mathcal{M}^*$ of nonnegative multiplicative functions, nice in the sense that there exist constants $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \geq 0$, $\varrho_0 \in [0, 1)$ and $r \in (0, 1)$ such that

$$(2.1) \qquad \alpha(p^\nu) \ll p^{(\varrho_0 + \sigma_0 - 1)\nu},$$

$$(2.2) \qquad \sum_{p \leq x} \frac{\alpha(p) \log p}{p^{\sigma_0 - 1}} = \beta x + O\left(\frac{x}{(\log x)^{A_0}}\right),$$

$$(2.3) \qquad {\sum_{p}}' \left(\frac{\alpha(p)^2}{p^{2(r+\sigma_0-1)}} + \sum_{\nu \geq 2} \frac{\alpha(p^\nu)}{p^{(r+\sigma_0-1)\nu}}\right) < \infty,$$

$$(2.4) \qquad \sum_{\nu \geq 1} \frac{\nu \alpha(p^\nu)}{p^{\sigma_0 \nu}} \ll \frac{(\log\log(p+1))^{\vartheta_0}}{p},$$

where the sum $\sum_p'$ is over all but finitely many primes p. It is not hard to verify that $\mathcal{M}^*$ is closed under Dirichlet convolution. Despite some overlaps between our class $\mathcal{M}^*$ and the class of multiplicative functions studied by Elboim and Gorodetsky [11], neither of them strictly contains the other. On the one hand, the multiplicative function α defined by $\alpha(p) = 1$ for all primes p and $\alpha(p^\nu) = p^{\nu/3}$ for all prime powers p^ν with $\nu \geq 2$ falls into $\mathcal{M}^*$ but is not covered by the first part of [11, Theorem 1.1]. On the other hand, (2.4) implies the more restrictive growth condition $\alpha(p)/p^{\sigma_0-1} \ll (\log\log(p+1))^{\vartheta_0}$, which is not required in [11, Theorem 1.1].

The class $\mathcal{M}^*$ contains many familiar multiplicative functions, including the κ-fold divisor function $d_\kappa(n)$ for $\kappa > 0$, the sum-of-divisors function $\sigma_\lambda(n)$ for $\lambda > -1$, Euler's totient function $\varphi(n)$, the characteristic function $\mu(n)^2$ of squarefree numbers, and the function $r_2(n)/4$, where $\mu(n)$ is the Möbius function and $r_2(n) := \#\{(a,b) \in \mathbb{Z}^2 : n = a^2 + b^2\}$. Less obvious examples include $\rho_g(n)$, which denotes the number of zeros in $\mathbb{Z}/n\mathbb{Z}$ of a nonconstant irreducible polynomial $g \in \mathbb{Z}[x]$, and Ramanujan's τ-function $\tau(n)$, which may be defined as the nth Fourier coefficient of the modular discriminant. We leave the verification of these claims to the interested reader.

Let $\alpha \in \mathcal{M}^*$ with parameters $A_0, \beta, \sigma_0, \vartheta_0, \varrho_0, r$ and set

$$S(x) = S_\alpha(x) := \sum_{n \leq x} \alpha(n).$$

For any additive function $f \colon \mathbb{N} \to \mathbb{R}$, define

$$A(x) = A_{\alpha,f}(x) := \sum_{p \leq x} \alpha(p) \frac{f(p)}{p^{\sigma_0}}, \qquad B(x) = B_{\alpha,f}(x) := \sum_{p \leq x} \alpha(p) \frac{f(p)^2}{p^{\sigma_0}}.$$

One may think of n as a random variable defined on the sample space $\mathbb{N} \cap [1, x]$ with the natural probability measure induced by α, that is, $\mathrm{Prob}(n = k) = \alpha(k)/S(x)$ for every $k \in \mathbb{N} \cap [1, x]$. We shall show, by estimating the weighted mth moment defined by

$$M(x; m) = M_{\alpha,f}(x; m) := S(x)^{-1} \sum_{n \leq x} \alpha(n)(f(n) - A(x))^m$$

for every $m \in \mathbb{N}$, that for certain additive functions f, the limiting distribution of the normalization $(f(n) - A(x))/\sqrt{B(x)}$ is standard Gaussian. To state our results in a coherent manner, we set $\chi_m := (1 + (-1)^m)/2$, the characteristic function of even integers, and

$$C_m := \frac{m!}{2^{m/2}\Gamma(m/2 + 1)}$$

for $m \in \mathbb{N}$, where Γ is the Gamma function. One quickly notes that $C_m = \mu_m = (m-1)!!$ for m even. The numbers C_m play a nonnegligible role in the error terms of our uniform estimates for $M(x; m)$. Our first result is the following theorem.

THEOREM 2.1. *Let $f \colon \mathbb{N} \to \mathbb{R}$ be a strongly additive function with $|f(p)| \leq M$ for all p, where $M > 0$ is constant, and let $\alpha \in \mathcal{M}^*$ with parameters $A_0, \beta, \sigma_0, \vartheta_0, \varrho_0, r$. If $\beta = 1$ and $0 < h_0 < (3/2)^{2/3}$ is arbitrary, and if $B(x) \to \infty$ as $x \to \infty$, then*

$$M(x; m) = C_m B(x)^{m/2}\left(\chi_m + O\left(\frac{Mm^{3/2}}{\sqrt{B(x)}}\right)\right)$$

uniformly for all sufficiently large x and all $1 \leq m \leq h_0(B(x)/M^2)^{1/3}$. If $\beta \neq 1$ and if $B(x)/(\log\log\log x)^2 \to \infty$ as $x \to \infty$, then

$$M(x; m) = C_m B(x)^{m/2}\left(\chi_m + O\left(\frac{Mm^{3/2}\log\log\log x}{\sqrt{B(x)}}\right)\right)$$

uniformly for all sufficiently large x and all

$$1 \leq m \ll B(x)^{1/3}/(\log\log\log x)^{2/3}.$$

The implicit constants in the error terms of both asymptotic formulas above depend at most on the explicit and implicit constants in the hypotheses except for M.

REMARK 2.1. It may be worth pointing out that in the same way as in Theorem 2.1, the implicit constants in the estimates appearing in the rest of the paper depend at most on the explicit and implicit constants in the hypotheses unless stated otherwise.

In the case where $\alpha = 1$ and $f = \omega$, we recover [17, Theorem 1] with a slightly wider range $1 \leq m \leq h_0(\log\log x)^{1/3}$ compared to the original range $1 \leq m \leq (\log\log x)^{1/3}$. Though Theorem 2.1 is formulated for strongly additive functions, similar things can be said about additive functions whose values at prime powers do not grow too rapidly and are hence not expected to contribute very much. A simple example of such functions is $\Omega(n)$. Since $\Omega(p^\nu) = \nu$ for all p^ν, one can show, by establishing a weighted version of (1.2), that $\Omega(n)$ does not differ from its cousin $\omega(n)$ very much for most values of n, and so they are expected to have the same distribution. More

generally, we shall prove the following variant of Theorem 2.1 for additive functions. For simplicity, we shall focus on a subclass of the multiplicative functions in $\mathcal{M}^*$.

THEOREM 2.2. *Let $f\colon \mathbb{N} \to \mathbb{R}$ be an additive function such that $f(p^\nu) \leq M\nu^\kappa$ for all prime powers p^ν, where $M > 0$ and $\kappa \geq 0$ are constant. Let $\alpha\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a multiplicative function, and suppose that there exist constants $A_0, \beta, \sigma_0 > 0$, $\vartheta_0 \geq 0$, $\varrho_0 \in [0, 1/2)$ and $\lambda \in (0, 2^{1-2\varrho_0})$ such that $\alpha(n)$ satisfies (2.2), (2.4), and the condition that $\alpha(p^\nu) = O((\lambda p^{\varrho_0+\sigma_0-1})^\nu)$ for all prime powers p^ν. If $\beta = 1$ and $0 < h_0 < (3/2)^{2/3}$ is arbitrary, and if $B(x) \to \infty$ as $x \to \infty$, then*

$$M(x; m) = C_m B(x)^{m/2}\left(\chi_m + O\left(\frac{Mm^{\kappa+3/2}}{\sqrt{B(x)}}\right)\right)$$

uniformly for all sufficiently large x and all $m \in \mathbb{N}$ satisfying

$$m \leq h_0(B(x)/M^2)^{1/3} \quad and \quad m \ll B(x)^{1/(2\kappa+3)}.$$

If $\beta \neq 1$ and if $B(x)/(\log\log\log x)^2 \to \infty$ as $x \to \infty$, then we have

$$M(x; m) = C_m B(x)^{m/2}\left(\chi_m + O\left(\frac{Mm^{3/2}(\log\log\log x + m^\kappa)}{\sqrt{B(x)}}\right)\right)$$

uniformly for all sufficiently large x and all

$$1 \leq m \ll \min\left(B(x)^{1/(2\kappa+3)}, \frac{B(x)^{1/3}}{(\log\log\log x)^{2/3}}\right).$$

The implicit constants in both asymptotic formulas above depend at most on the explicit and implicit constants in the hypotheses except for M.

It is easy to see that if $\alpha(p^\nu) = O((\lambda p^{r_0+\sigma_0-1})^\nu)$ for all prime powers p^ν, where $\sigma_0 > 0$, $r_0 \in [0, 1/2)$ and $\lambda \in (0, 2^{1-2r_0})$ are given constants, then conditions (2.1) and (2.3) are automatically fulfilled with any fixed $\max(r_0 + \log_2 \lambda, 0) \leq \varrho_0 < 1$, $r_0 + \max(1/2, \log_2 \lambda) < r < 1$, and the same parameter σ_0. Indeed, we shall derive Theorem 2.2 as a corollary of Theorem 2.1.

Let $g \in \mathbb{Z}[x]$ be a nonconstant irreducible polynomial, and recall that for every $n \in \mathbb{N}$, $\rho_g(n)$ denotes the number of zeros of g in $\mathbb{Z}/n\mathbb{Z}$. More generally, if $g \in \mathbb{Q}[x]$ is a nonconstant irreducible polynomial, we may extend the definition above by setting $\rho_g(n) = 0$ if $\gcd(n, c_g) > 1$, where $c_g \in \mathbb{N}$ is the least positive integer such that $c_g g(x) \in \mathbb{Z}[x]$, and insisting that $\rho_g(n)$ be the number of zeros of $g(x)$ (or equivalently, $c_g g(x)$) in $\mathbb{Z}/n\mathbb{Z}$ when $\gcd(n, c_g) = 1$. Extended this way with the convention that $\rho_g(1) = 1$, the function $\rho_g(n)$ remains multiplicative. By [20, Lemma 1], ρ_g is bounded on prime powers

and satisfies
$$\sum_{p \leq x} \frac{\rho_g(p)}{p} = \log \log x + M_{\rho_g} + O\left(\frac{1}{\log x}\right).$$

Given a strongly additive function $f \colon \mathbb{N} \to \mathbb{R}$, we define
$$A_{f,g}(x) := \sum_{p \leq x} \rho_g(p) \frac{f(p)}{p}, \quad B_{f,g}(x) := \sum_{p \leq x} \rho_g(p) \frac{f(p)^2}{p}.$$

For simplicity, suppose that $g(\mathbb{N}) \subseteq \mathbb{N}$. In the case $g \in \mathbb{Z}[x]$, Halberstam [19, Theorem 3] showed that if $B_{f,g}(x) \to \infty$ as $x \to \infty$, and if $f(p) = o(\sqrt{B_{f,g}(p)})$ ([1]), then given $m \in \mathbb{N}$,
$$\frac{1}{x} \sum_{n \leq x} (f(g(n)) - A_{f,g}(x))^m = (\mu_m + o(1)) B_{f,g}(x)^{m/2}.$$

Under the stronger condition $f(p) = O(1)$, Theorem 2.1 leads to a weighted version of this result in the case $g(n) = n$. The remaining cases are captured by the following theorem.

THEOREM 2.3. *Let $f \colon \mathbb{N} \to \mathbb{R}$ be a strongly additive function with $|f(p)| \leq M$ for all p, where $M > 0$ is constant, and let $g \in \mathbb{Q}[x]$ be a nonconstant irreducible polynomial such that $g(0) \neq 0$ and $g(\mathbb{N}) \subseteq \mathbb{N}$. Let $\alpha \in \mathcal{M}^*$ with parameters $A_0, \beta, \sigma_0, \vartheta_0, \varrho_0, r$, and fix $0 < h_0 < (3/2)^{2/3}$ and $h_0' > 0$. For any $q \in \mathbb{N}$ and $a \in \mathbb{Z}$ coprime to q, define*
$$\Delta_\alpha(x; q, a) := \sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, q)}} \alpha(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \leq x \\ \gcd(n,q)=1}} \alpha(n).$$

If there exist a constant $\epsilon_0 > 0$ and a function $\delta(x) \in (0,1]$ with $\delta(x)^2 B_{f,g}(x) \to \infty$ as $x \to \infty$ such that
$$(2.5) \qquad \sum_{\substack{q: \omega(q) \leq m \\ P^+(q) \leq x^{\delta(x)/m}}} \mu(q)^2 \sum_{a \in \mathcal{Z}_g^*(q)} |\Delta_\alpha(x; q, a)|$$
$$\ll S(x) \exp(-(\log \log x)^{1/3 + \epsilon_0})$$

uniformly for all sufficiently large x and all
$$(2.6) \qquad 1 \leq m \leq \min(h_0(B_{f,g}(x)/M^2)^{1/3}, h_0'(\delta(x)^2 B_{f,g}(x)/M^2)^{1/3}),$$

([1]) Halberstam [19] wrote that for $g(x) = x$ this pair of conditions contain the condition that $f(p) = o((\log p)^\epsilon)$ for every given $\epsilon > 0$. However, this claim is incorrect. In fact, a simple counterexample may be constructed as follows. Let $\mathcal{P}$ be an arbitrary infinite subset of odd primes such that $\sum_{p \in \mathcal{P}} 1/p < \infty$, and put $\mathcal{P}(x) := \mathcal{P} \cap [3, x]$. Define $f(p) = \sqrt{\log \log p}$ for $p \in \mathcal{P}$ and $f(p) = 1$ for $p \notin \mathcal{P}$. From partial summation it follows that $\sum_{p \in \mathcal{P}(x)} f(p)^2/p = o(\log \log x)$. Then one sees readily that $f(p) = o((\log p)^\epsilon)$ for any given $\epsilon > 0$, while $f(p) \sim \sqrt{B(p)}$ for large $p \in \mathcal{P}$.

where $\mathcal{Z}_g^*(q) := \{n \in (\mathbb{Z}/q\mathbb{Z})^\times : g(n) \equiv 0 \,(\mathrm{mod}\, q)\}$ *is the zero locus of* g *in* $(\mathbb{Z}/q\mathbb{Z})^\times$, *then*

$$S(x)^{-1} \sum_{n \leq x} \alpha(n)(f(g(n)) - A_{f,g}(x))^m$$
$$= C_m B_{f,g}(x)^{m/2} \left(\chi_m + O\left(\frac{Mm^{3/2}}{\delta(x)\sqrt{B_{f,g}(x)}} \right) \right)$$

uniformly for all sufficiently large x *and all* $m \in \mathbb{N}$ *in the same range* (2.6), *where the implicit constant depends at most on the explicit and implicit constants in the hypotheses except for* M.

Theorem 2.3 is applicable to a large class of nonnegative multiplication functions $\alpha(n)$, including $d_k(n)$ for $k \in \mathbb{N}$ and $r_2(n)/4$ (see [4, 26]). Despite the great generality of (2.5), it is oftentimes more convenient to work with the stronger variant

$$(2.7) \qquad \sum_{\substack{q \leq x^{\delta(x)} \\ \omega(q) \ll (\delta(x)^2 \log\log x)^{1/3}}} \mu(q)^2 \rho_g(q) \max_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} |\Delta_\alpha(x; q, a)|$$
$$\ll S(x) \exp(-(\log\log x)^{1/3 + \epsilon_0}).$$

This condition may be viewed as an inequality of the *Bombieri–Vinogradov type*, which ensures that the values of $\alpha(n)$ are well distributed as n varies over the reduced residue classes $a \,(\mathrm{mod}\, q)$ for most values of q and a. In view of [17, Proposition 4], such a condition arises naturally from a sieving process for the sequence $\{g(n)\}_{n \geq 1}$. For this process to work, we need information about the average size of $\alpha(n)$ subject to the constraint $d \mid g(n)$ for smooth squarefree $d \in \mathbb{N} \cap [1, x^\delta]$. If d is also free of small prime factors up to some constant depending on g, then this constraint amounts to the congruences $n \equiv a \,(\mathrm{mod}\, d)$ for $a \in \mathcal{Z}_g^*(d)$. So, (2.7) reduces the sieving of $\{g(n)\}_{n \geq 1}$ to that of $\mathbb{N}$.

We shall only sketch the proof of Theorem 2.3, since it is similar to, and in fact much easier than, that of Theorem 2.1. The argument used in the proof may also be modified to study the joint distribution of $f(n + h_1)$ and $f(n + h_2)$ with any fixed integers $h_1 \neq h_2$.

It is not hard to see that the condition $f(p) = O(1)$ in Theorem 2.1 can be relaxed, especially when we do not pursue uniformity in m in the asymptotics for the mth moment. For instance, in the case $\alpha = 1$ Delange and Halberstam showed [8, Theorem 1] that if $f \colon \mathbb{N} \to \mathbb{R}$ is a strongly additive function such that $B(x) \to \infty$ as $x \to \infty$, $f(p) = O(\sqrt{B(p)})$ for all primes p, and

$$(2.8) \qquad \sum_{\substack{p \leq x \\ |f(p)| > \epsilon \sqrt{B(x)}}} \frac{f(p)^2}{p} = o(B(x))$$

for any given $\epsilon > 0$, then

$$\frac{1}{x} \sum_{n \leq x} (f(n) - A(x))^m = (\mu_m + o(1)) B(x)^{m/2}$$

for every fixed $m \in \mathbb{N}$. The implication of this result on the distribution of f is slightly weaker than the Kubilius–Shapiro theorem [29, Theorem A] in that the latter asserts that the distribution of an additive function $f \colon \mathbb{N} \to \mathbb{R}$ with an unbounded variance $B(x)$ which satisfies (2.8) for every given $\epsilon > 0$ is necessarily Gaussian with mean $A(x)$ and variance $B(x)$. On the other hand, Delange and Halberstam noted that their result no longer holds if one removes the assumption $f(p) = O(\sqrt{B(p)})$, which incidentally exposes the limitation of the method of moments compared to the method developed by Erdős and Kac. Regardless, it will be clear in the sequel that the proof of Theorem 2.1 makes it possible for us to obtain the following natural extension of the result of Delange and Halberstam.

THEOREM 2.4. *Let* $f \colon \mathbb{N} \to \mathbb{R}$ *be a strongly additive function, and let* $\alpha \in \mathcal{M}^*$ *with parameters* $A_0, \beta, \sigma_0, \vartheta_0, \varrho_0, r$. *Define*

$$B^*(x) := \begin{cases} B(x) & \text{if } \beta = 1, \\ B(x)/(\log\log\log x)^2 & \text{if } \beta \neq 1, \end{cases}$$

and suppose $B^*(x) \to \infty$ *as* $x \to \infty$. *If there exists a constant* $K > 0$ *such that* $f(n) = o(\sqrt{B(x)})$ *for all squarefree* $n \in \mathbb{N} \cap [1, x]$ *composed of prime factors* p *with* $|f(p)| > K\sqrt{B^*(x)}$, *and if*

$$\sum_{\substack{p \leq x \\ |f(p)| > \epsilon\sqrt{B^*(x)}}} \alpha(p) \frac{f(p)^2}{p^{\sigma_0}} = o(B^*(x))$$

for any given $\epsilon > 0$, *then* $M(x; m) = (\mu_m + o(1)) B(x)^{m/2}$ *for every fixed* $m \in \mathbb{N}$.

The proof of Theorem 2.4, which we shall only outline, is based on the proofs of Theorem 2.1 and [8, Theorem 1]. We shall also obtain as a corollary the following analogue of the Kubilius–Shapiro theorem [29, Theorem C].

COROLLARY 2.5. *Under the notation and hypotheses of Theorem 2.4,*

$$\lim_{x \to \infty} S(x)^{-1} \sum_{\substack{n \leq x \\ f(n) \leq A(x) + V\sqrt{B(x)}}} \alpha(n) = \Phi(V)$$

for any given $V \in \mathbb{R}$. *The same is true if* f *is merely additive.*

It is clear that Theorem 2.4 implies Corollary 2.5 when f is strongly additive. To handle the general case where f is merely additive, we shall establish a weighted version of [29, Theorem B] which shows that when it

comes to the distribution problem, there is no essential difference between strongly additive functions and general additive functions, and thus the distribution of an additive function f is determined solely by its values at primes.

Corollary 2.5 has many interesting applications. For instance, it implies at once that if $h \colon \mathbb{N} \to \mathbb{R}$ is any *completely additive* function, i.e., $h(mn) = h(m) + h(n)$ for all $m, n \in \mathbb{N}$, and if $k > 1$ is a positive integer, then the distribution of $h(d_k(n))$ weighted by $\alpha(n)$ is Gaussian with mean $h(k)\beta \log \log x$ and variance $h(k)^2 \beta \log \log x$, provided $h(k) \neq 0$. In [12] Elliott proved a weighted Erdős–Kac theorem concerning Ramanujan's τ-function. In Remark 9.1 we describe how his result may be derived from Corollary 2.5. Analogues on elliptic holomorphic newforms of weight at least 2 can be obtained in the same way. In a similar fashion, one can also show that if the weight α in Corollary 2.5 satisfies the additional condition $\alpha(p) \sim \beta p^{\sigma_0 - 1}$ for all but a subset E of primes p, where $\#(E \cap [2, x]) = o(x(\log \log x)^{2 - \vartheta_0}/(\log x)^3)$ as $x \to \infty$, then the distribution of $\Omega(\varphi(n))$ weighted by $\alpha(n)$ is Gaussian with mean $\beta(\log \log x)^2/2$ and variance $\beta(\log \log x)^3/3$, generalizing an old result of Erdős and Pomerance [16, Theorem 3.1] in an easy manner.

REMARK 2.2. The condition that $f(p) = o((\log p)^\epsilon)$ for any given $\epsilon > 0$, mentioned by Halberstam [19], does not imply (2.8) in general. To see this, assume for the moment that there exists an infinite subset $\mathcal{P}$ of primes such that

$$(2.9) \qquad s_{\mathcal{P}}(x) := \sum_{p \in \mathcal{P} \cap [17, x]} \frac{1}{p} = \frac{\log \log x}{\log \log \log x} + c + o(1)$$

for sufficiently large x, where $c \in \mathbb{R}$ is some constant. Next, define $f(p) = (\log p)^{1/(2 \log \log \log p)}$ for $p \in \mathcal{P}$ and $f(p) = 1$ for $p \notin \mathcal{P}$. Clearly, $f(p) = o((\log p)^\epsilon)$ for any given $\epsilon > 0$. It is easily seen by partial summation that

$$\sum_{p \in \mathcal{P} \cap [17, x]} \frac{f(p)^2}{p} = \int_{17^-}^{x} (\log t)^{1/\log \log \log t}\, ds_{\mathcal{P}}(t) = (1 + o(1))(\log x)^{1/\log \log \log x},$$

which implies that

$$B(x) = \sum_{p \in \mathcal{P} \cap [17, x]} \frac{f(p)^2}{p} + O(\log \log x) = (1 + o(1))(\log x)^{1/\log \log \log x}.$$

Let $y = x^{1/\log \log x}$ and $\epsilon = 1/2$. Since

$$\log \log y = \log \log x - \log \log \log x,$$

$$\log\log\log y = \left(1 + O\left(\frac{1}{\log\log x}\right)\right)\log\log\log x,$$

we have

$$(\log y)^{1/\log\log\log y} = \exp\left(\frac{\log\log x}{\log\log\log x} - 1 + O\left(\frac{1}{\log\log\log x}\right)\right).$$

It follows that

$$f(p)^2 > (\log y)^{1/\log\log\log y} > \frac{1}{3}(\log x)^{1/\log\log\log x} > \epsilon^2 B(x)$$

for $p \in \mathcal{P} \cap (y, x]$ when x is sufficiently large. Hence,

$$\sum_{\substack{p \leq x \\ |f(p)| > \epsilon\sqrt{B(x)}}} \frac{f(p)^2}{p} \geq \sum_{p \in \mathcal{P} \cap (y,x]} \frac{f(p)^2}{p} > \frac{1}{2}(\log x)^{1/\log\log\log x} > \frac{1}{3}B(x).$$

It remains to construct a set $\mathcal{P}$ with the desired property (2.9). The following inductive approach was suggested by Prof. Pomerance. Note first that $\sum_{p \leq x} 1/p = \log\log x + O(1)$ grows slightly faster than our target $u(x) := \log\log x/\log\log\log x$, according to Mertens' second theorem [22, Theorem 427]. Moreover, if $p < p'$ are large consecutive primes, then $u(p') - u(p) = o(1/\log p)$, by Bertrand's postulate. Let 17 be the first prime in $\mathcal{P}$. Suppose that we have already selected for $\mathcal{P}$ the primes up to q, where q is prime. We put the next prime q' in $\mathcal{P}$ if $s_{\mathcal{P}}(q) < u(q)$ and leave it out otherwise. Then the running sum $s_{\mathcal{P}}(x)$ changes by at most $1/q$ as x moves from q to q', while the target $u(x)$ changes by at most $o(1/\log q)$ as x moves from q to q'. Thus, the difference $s_{\mathcal{P}}(x) - u(x)$ can be kept within $o(1/\log x)$. In particular, (2.9) holds for $\mathcal{P}$ with $c = 0$.

Overview of the proof of Theorem 2.1. Before embarking on the proofs of our results, we briefly describe the main steps in the proof of Theorem 2.1. The starting point is an approximation to moments used by Granville, Soundararajan, Khan, Milinovich and Subedi. Though the underlying idea is the same, we need a more complicated version of this approximation (see Lemma 4.1) due to the more general nature of our multiplicative weights α. To utilize it, we first need to develop an asymptotic formula for the mean value of $\alpha(n)$ with $n \leq x$ restricted to the multiples of a squarefree integer $a \in \mathbb{N} \cap [1, x]$ (see Lemma 3.3). An important feature of this formula is that it holds uniformly for all squarefree integers $a \in \mathbb{N} \cap [1, x]$, which is key to both applying the moment approximation and making the moment estimates uniform. This formula will serve as a substitute for the one on d_k developed by Khan, Milinovich and Subedi. Unlike their proof, which is based on Perron's formula and the complete submultiplicativity of d_k, our proof uses the mean value estimate for α supplied by [5, Theorem 2.1] and is completely elementary. It is carried out in the next section.

After applying the moment approximation, we find that the estimation of the main contribution can be worked out as in [17, 24]. It is the estimation of the error terms that is more involved in our case. In particular, the estimation of the error term in the moment approximation provided by Lemma 4.1 requires separate treatments of $\beta = 1$ and $\beta \neq 1$. Moreover, since the error term in our asymptotic formula for the mean value of $\alpha(n)$ over $a \mid n$ supplied by Lemma 3.3 is weaker than what one can obtain for the special weight $d_k(n)$ by complex-analytic approaches, we need to handle the case $\beta \in (0, 1)$ with some special care and make a careful selection of parameters accordingly in order to minimize the error terms. With these new technicalities taken care of, we obtain the desired uniform estimates for moments stated in Theorem 2.1.

3. Mean values of multiplicative functions. Without loss of generality, we may assume $A_0 \in (0, 1)$. In addition, we shall also make use of the asymptotic formula

$$(3.1) \qquad \sum_{p \leq x} \frac{\alpha(p)}{p^{\sigma_0}} = \beta \log \log x + M_\alpha + O((\log x)^{-A_0})$$

with some constant $M_\alpha \in \mathbb{R}$, which follows immediately from (2.2) via partial summation. In view of our assumption that $f(p) = O(1)$, this formula implies trivially that $B(x) \ll \log \log x$. Moreover, if we define, for every prime p,

$$\psi_0(p) := \sum_{\nu \geq 2} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}},$$

then we infer from (2.1), (2.3) and (2.4) that

$$\psi_0(p) \ll \frac{(\log \log(p + 1))^{\vartheta_0}}{p}$$

and $\sum_p \psi_0(p) < \infty$.

LEMMA 3.1. *Let* $\alpha \colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ *be a multiplicative function satisfying* (2.1) *and* (2.4) *with some* $\sigma_0, \vartheta_0 > 0$ *and* $\varrho_0 \in [0, 1)$. *Fix* $h \in \mathbb{R}$, $\epsilon_0 \in (0, 1)$ *and* $c_0 \in [1, \epsilon_0^{-1})$, *and define*

$$I_{\alpha,h}(x; a) := \sum_{\substack{q \leq x \\ R_q = a}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{3x}{q} \right)^h,$$

where $a \in \mathbb{N} \cap [1, x]$ *is squarefree. Then there exists a constant* $\delta_0 > 0$ *such that uniformly for all sufficiently large* x, *any* $\delta \in [\delta_0 \log \log x / \log x, 1]$, *and any squarefree* $a \in \mathbb{N} \cap [1, x]$ *with* $\omega(a) \leq (1 - \varrho_0)\epsilon_0 \delta^{-1}$, *we have*

$$I_{\alpha,h}(x;a)$$
$$= \left(\tilde{\lambda}_\alpha(a) + O\left(\frac{2^{O(\omega(a))}}{\log x}\left(\frac{1}{x^{c_0\delta\omega(a)}} + \frac{\epsilon_{h,0}L(a)\log P^+(a)}{a}\right)\right)\right)(\log x)^h,$$

where

$$\tilde{\lambda}_\alpha(a) := \prod_{p|a}\sum_{\nu=1}^{\infty}\frac{\alpha(p^\nu)}{p^{\sigma_0\nu}}, \qquad L(a) := \prod_{p|a}(\log\log(p+1))^{\vartheta_0}.$$

Proof. Let $\delta \in (0,1]$ and fix $c_1 \in (c_0, \epsilon_0^{-1})$. Put $\delta_1 := (1-\varrho_0)^{-1}c_1\delta$ and $y := x^{k\delta_1}$. For any squarefree $a = p_1\cdots p_k \in \mathbb{N}\cap[1,x]$ with $p_1 < \cdots < p_k \le x$ and $k \le (1-\varrho_0)\epsilon_0\delta^{-1}$, we have $k\delta_1 \le c_1\epsilon_0 < 1$ and

$$I_{\alpha,h}(x;a) = \sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le x \\ \nu_1,\ldots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_k^{\nu_k}}\right)^h.$$

On the one hand,

$$\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y \\ \nu_1,\ldots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}\left(\log\frac{3x}{p_1^{\nu_1}\cdots p_k^{\nu_k}}\right)^h$$

$$= \sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y \\ \nu_1,\ldots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}(\log 3x)^h\left(1 + O\left(\frac{\epsilon_{h,0}}{\log 3x}\sum_{i=1}^{k}\nu_i\log p_i\right)\right)$$

$$= \sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y \\ \nu_1,\ldots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}(\log x)^h + O\left(\frac{2^{O(k)}\epsilon_{h,0}L(a)\log p_k}{a}(\log x)^{h-1}\right),$$

by (2.4). From (2.1) it follows that

$$\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}\le y \\ \nu_1,\ldots,\nu_k\ge 1}}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_k^{\nu_k})}{p_1^{\sigma_0\nu_1}\cdots p_k^{\sigma_0\nu_k}}$$

$$= \tilde{\lambda}_\alpha(a) + O\left(2^{O(k)}\sum_{\substack{p_1^{\nu_1}\cdots p_k^{\nu_k}>y \\ \nu_1,\ldots,\nu_k\ge 1}}\frac{1}{p_1^{(1-\varrho_0)\nu_1}\cdots p_k^{(1-\varrho_0)\nu_k}}\right).$$

The sum in the error term above may be split into two sums according as $p_2^{\nu_2}\cdots p_k^{\nu_k} \le y$ or $p_2^{\nu_2}\cdots p_k^{\nu_k} > y$. In the first sum we must have $p_1^{\nu_1} > y/(p_2^{\nu_2}\cdots p_k^{\nu_k})$. Thus summing over ν_1 and then over $\nu_2,\ldots,\nu_k$, we see that the first sum is

$$\ll \frac{1}{y^{1-\varrho_0}} \sum_{\substack{p_2^{\nu_2} \cdots p_k^{\nu_k} \leq y \\ \nu_2, \ldots, \nu_k \geq 1}} 1 \leq \frac{2^{O(k)}(\log x)^{k-1}}{x^{c_1 k \delta}(\log p_2) \cdots (\log p_k)}.$$

The second sum is simply

$$\sum_{\substack{p_2^{\nu_2} \cdots p_k^{\nu_k} > y \\ \nu_2, \ldots, \nu_k \geq 1}} \frac{1}{p_2^{(1-\varrho_0)\nu_2} \cdots p_k^{(1-\varrho_0)\nu_k}} \sum_{\nu_1 \geq 1} \frac{1}{p_1^{(1-\varrho_0)\nu_1}}$$

$$\ll \sum_{\substack{p_2^{\nu_2} \cdots p_k^{\nu_k} > y \\ \nu_2, \ldots, \nu_k \geq 1}} \frac{1}{p_2^{(1-\varrho_0)\nu_2} \cdots p_k^{(1-\varrho_0)\nu_k}}.$$

It follows that

$$\sum_{\substack{p_1^{\nu_1} \cdots p_k^{\nu_k} > y \\ \nu_1, \ldots, \nu_k \geq 1}} \frac{1}{p_1^{(1-\varrho_0)\nu_1} \cdots p_k^{(1-\varrho_0)\nu_k}}$$

$$\ll \sum_{\substack{p_2^{\nu_2} \cdots p_k^{\nu_k} > y \\ \nu_2, \ldots, \nu_k \geq 1}} \frac{1}{p_2^{(1-\varrho_0)\nu_2} \cdots p_k^{(1-\varrho_0)\nu_k}} + \frac{2^{O(k)}(\log x)^{k-1}}{x^{c_1 k \delta}(\log p_2) \cdots (\log p_k)}.$$

Repeating this argument, we obtain

$$\sum_{\substack{p_1^{\nu_1} \cdots p_k^{\nu_k} > y \\ \nu_1, \ldots, \nu_k \geq 1}} \frac{1}{p_1^{(1-\varrho_0)\nu_1} \cdots p_k^{(1-\varrho_0)\nu_k}} \leq \frac{2^{O(k)}(\log x)^{k-1}}{x^{c_1 k \delta}(\log p_2) \cdots (\log p_k)},$$

from which we deduce

$$(3.2) \qquad \sum_{\substack{p_1^{\nu_1} \cdots p_k^{\nu_k} \leq y \\ \nu_1, \ldots, \nu_k \geq 1}} \frac{\alpha(p_1^{\nu_1}) \cdots \alpha(p_k^{\nu_k})}{p_1^{\sigma_0 \nu_1} \cdots p_k^{\sigma_0 \nu_k}} = \tilde{\lambda}_\alpha(a) + O\left(\frac{2^{O(k)}(\log x)^{k-1}}{x^{c_1 k \delta}(\log p_2) \cdots (\log p_k)} \right).$$

On the other hand,

$$\sum_{x_1 < p^\nu \leq x_2} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \left(\log \frac{3x_2}{p^\nu} \right)^h \ll \sum_{\substack{\log_p x_1 < \nu \leq \log_p x_2 \\ \nu \in \mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)\nu}} \left(\log \frac{3x_2}{p^\nu} \right)^h$$

$$= -\int\limits_{\log_p x_1}^{\log_p x_2} \left(\log \frac{3x_2}{p^t} \right)^h d\left(\sum_{\substack{t < \nu \leq \log_p x_2 \\ \nu \in \mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)t}} \right)$$

uniformly for all primes p and all $0 < x_1 \leq x_2$. Using integration by parts,

we see that the integral above is equal to

$$-(\log(3x_2/x_1))^h \sum_{\substack{\log_p x_1 < \nu \le \log_p x_2 \\ \nu \in \mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)\nu}}$$

$$-\int_{\log_p x_1}^{\log_p x_2} \left(\sum_{\substack{t < \nu \le \log_p x_2 \\ \nu \in \mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)t}} \right) d\left(\log \frac{3x_2}{p^t} \right)^h.$$

Since

$$\sum_{\substack{t < \nu \le \log_p x_2 \\ \nu \in \mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)t}} < \frac{1}{p^{(1-\varrho_0)(\lfloor t \rfloor+1)}} \cdot \frac{p^{1-\varrho_0}}{p^{1-\varrho_0}-1} \ll \frac{1}{p^{(1-\varrho_0)t}},$$

we have

$$(\log(3x_2/x_1))^h \sum_{\substack{\log_p x_1 < \nu \le \log_p x_2 \\ \nu \in \mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)\nu}} \ll \frac{(\log(3x_2/x_1))^h}{x_1^{1-\varrho_0}}$$

and

$$\int_{\log_p x_1}^{\log_p x_2} \left(\sum_{\substack{t < \nu \le \log_p x_2 \\ \nu \in \mathbb{Z}}} \frac{1}{p^{(1-\varrho_0)t}} \right) d\left(\log \frac{3x_2}{p^t} \right)^h$$

$$\ll \epsilon_{h,0} \log p \int_{\log_p x_1}^{\log_p x_2} \frac{1}{p^{(1-\varrho_0)t}} \left(\log \frac{3x_2}{p^t} \right)^{h-1} dt$$

$$= \frac{\epsilon_{h,0}}{(3x_2)^{1-\varrho_0}} \int_{\log 3}^{\log(3x_2/x_1)} t^{h-1} e^{(1-\varrho_0)t} \, dt$$

$$\ll \frac{\epsilon_{h,0}}{(3x_2)^{1-\varrho_0}} (\log(3x_2/x_1))^{h-1} \left(\frac{3x_2}{x_1} \right)^{1-\varrho_0}$$

$$= \frac{\epsilon_{h,0} (\log(3x_2/x_1))^{h-1}}{x_1^{1-\varrho_0}}.$$

Hence,

$$(3.3) \qquad \sum_{x_1 < p^\nu \le x_2} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} \left(\log \frac{3x_2}{p^\nu} \right)^h \ll \frac{(\log(3x_2/x_1))^h}{x_1^{1-\varrho_0}}$$

uniformly for all primes p and all $0 < x_1 \le x_2$. This inequality implies immediately

$$\sum_{\substack{y < p_1^{\nu_1} \cdots p_k^{\nu_k} \leq x \\ \nu_1, \ldots, \nu_k \geq 1}} \frac{\alpha(p_1^{\nu_1}) \cdots \alpha(p_k^{\nu_k})}{p_1^{\sigma_0 \nu_1} \cdots p_k^{\sigma_0 \nu_k}} \left(\log \frac{3x}{p_1^{\nu_1} \cdots p_k^{\nu_k}} \right)^h \leq \frac{2^{O(k)} (\log x)^h}{y^{1-\varrho_0}} \sum_{\substack{p_2^{\nu_2} \cdots p_k^{\nu_k} \leq x \\ \nu_2, \ldots, \nu_k \geq 1}} 1$$

$$\leq \frac{2^{O(k)} (\log x)^{k+h-1}}{x^{c_1 k \delta} (\log p_2) \cdots (\log p_k)}.$$

Lemma 3.1 now follows upon combining the above with (3.2) and taking $\delta_0 = 1/(c_1 - c_0)$ with the range $\delta \geq \delta_0 \log \log x / \log x$ in mind. ∎

Let α be a multiplicative function as in Theorem 2.1 with $A_0 \in (0,1)$. Suppose first that (2.3) holds with the restricted sum $\sum_p'$ replaced by the full sum $\sum_p$. For $\sigma_0 = 1$ de la Bretèche and Tenenbaum [5, Theorem 2.1] showed

$$\sum_{n \leq x} \alpha(n) = \frac{1}{\Gamma(\beta)} \prod_p \left(1 - \frac{1}{p} \right)^\beta \sum_{\nu=0}^\infty \frac{\alpha(p^\nu)}{p^\nu} x (\log x)^{\beta-1} \left(1 + O\left(\frac{1}{(\log x)^{A_0}} \right) \right),$$

where the implicit constant depends at most on the explicit and implicit constants in the hypotheses. For the general case where $\sigma_0 > 0$ is arbitrary, it is easy to show, by applying the above to $\alpha(n)/n^{\sigma_0-1}$ and employing partial summation as in [11, proof of Corollary 3.2], that

$$(3.4) \qquad S(x) = \sum_{n \leq x} \alpha(n) = \lambda_\alpha x^{\sigma_0} (\log x)^{\beta-1} \left(1 + O\left(\frac{1}{(\log x)^{A_0}} \right) \right),$$

where

$$(3.5) \qquad \lambda_\alpha := \frac{1}{\sigma_0 \Gamma(\beta)} \prod_p \left(1 - \frac{1}{p} \right)^\beta \sum_{\nu=0}^\infty \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}}.$$

Suppose now that (2.3) holds with the restricted sum $\sum_p'$ being the sum $\sum_{p > Q_0}$, where $Q_0 \geq 1$ is some constant. Let $P_0 := \prod_{p \leq Q_0} p$ and let $\mathbf{1}_{P_0}(n)$ be the indicator function of the set $\{n \in \mathbb{N} : \gcd(n, P_0) = 1\}$. Then $\alpha(n)\mathbf{1}_{P_0}(n)$ is a nonnegative multiplicative function satisfying (2.1)–(2.4) with the sum $\sum_p'$ in (2.3) replaced by the full sum $\sum_p$. In particular, (3.4) is applicable to $\alpha(n)\mathbf{1}_{P_0}(n)$. Thus, we obtain

$$(3.6) \qquad \sum_{\substack{n \leq x \\ \gcd(n, P_0)=1}} \alpha(n) = \lambda_\alpha(P_0) x^{\sigma_0} (\log 3x)^{\beta-1} \left(1 + O\left(\frac{1}{(\log 3x)^{A_0}} \right) \right),$$

where

$$\lambda_\alpha(P_0) := \frac{1}{\sigma_0 \Gamma(\beta)} \prod_{p \leq Q_0} \left(1 - \frac{1}{p} \right)^\beta \cdot \prod_{p > Q_0} \left(1 - \frac{1}{p} \right)^\beta \sum_{\nu=0}^\infty \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}}.$$

Examining the proof of Lemma 3.1, we find that for every given $h \in \mathbb{R}$,

$$\sum_{\substack{q \leq x \\ R_q | P_0}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{3x}{q} \right)^h = \prod_{p \leq Q_0} \sum_{\nu=0}^{\infty} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}} (\log x)^h \left(1 + O\left(\frac{1}{\log x} \right) \right)$$

for all sufficiently large x. Combining this with (3.6) gives

$$S(x) = \sum_{\substack{q \leq x \\ R_q | P_0}} \alpha(q) \sum_{\substack{n' \leq x/q \\ \gcd(n', P_0)=1}} \alpha(n') = \lambda_\alpha x^{\sigma_0} (\log x)^{\beta-1} \left(1 + O\left(\frac{1}{(\log x)^{A_0}} \right) \right),$$

which is the same as (3.4).

For our applications, we will need an asymptotic formula for

$$S(x; a) = S_\alpha(x; a) := \sum_{\substack{n \leq x \\ \gcd(n,a)=1}} \alpha(n)$$

uniform in $a \in \mathbb{N} \cap [1, x]$. One may be tempted to apply (3.4) to the function $\alpha(n) \mathbf{1}_a(n)$, where $\mathbf{1}_a(n)$ is the indicator function of the set $\{n \in \mathbb{N}: \gcd(n, a) = 1\}$. However, it is not immediately clear whether the implied constant in the error term obtained via this naive approach is independent of $a \in \mathbb{N} \cap [1, x]$. Fortunately, the following lemma provides the desired estimate for $S(x; a)$ under the hypotheses (2.1)–(2.4).

LEMMA 3.2. *For any $\alpha \in \mathcal{M}^*$ with parameters $A_0, \beta, \sigma_0, \vartheta_0, \varrho_0, r$, we have*

$$S(x; a) = x^{\sigma_0} (\log x)^{\beta-1} \left(\lambda_\alpha(a) + O\left(\frac{1}{(\log x)^{A_0}} \right) \right)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$, where

$$\lambda_\alpha(a) := \frac{1}{\sigma_0 \Gamma(\beta)} \prod_{p|a} \left(1 - \frac{1}{p} \right)^\beta \cdot \prod_{p \nmid a} \left(1 - \frac{1}{p} \right)^\beta \sum_{\nu=0}^{\infty} \frac{\alpha(p^\nu)}{p^{\sigma_0 \nu}},$$

The implicit constant depends at most on the explicit and implicit constants in the hypotheses.

Proof. Let $a \in \mathbb{N} \cap [1, x]$. For simplicity of notation, we write $\sum^a$ for sums in which the indices take values coprime to a. As we have demonstrated above, there is no loss of generality in assuming that $\sigma_0 = 1$ and that (2.3) holds with the restricted sum $\sum'_p$ replaced by the full sum $\sum_p$. Note that

$$0 < \lambda_\alpha(a) = \lambda_\alpha \prod_{p|a} \left(\sum_{\nu=0}^{\infty} \frac{\alpha(p^\nu)}{p^\nu} \right)^{-1} \leq \lambda_\alpha.$$

To estimate $S(x; a)$, we start by connecting it with

$$T(x; a) = T_\alpha(x; a) := \sum_{n \leq x}^a \frac{\alpha(n)}{n}.$$

It is clear from (3.4) that $S(x; a) \leq S(x; 1) \ll x(\log x)^{\beta-1}$ and $T(x; a) \leq T(x; 1) \ll (\log x)^\beta$. Moreover, it is shown in [5, proof of Theorem 2.1] that

(3.7) $$T(x; 1) = \left(1 + O\left(\frac{1}{\log x}\right)\right) \frac{\lambda_\alpha}{\beta} (\log x)^\beta.$$

Following that proof, we find

(3.8) $\quad S(x; a) \log x$

$$= \sum_{n \leq x}^a \alpha(n) \log n + \sum_{n \leq x}^a \alpha(n) \log \frac{x}{n}$$

$$= \sum_{k \leq x}^a \alpha(k) \sum_{\substack{p^\nu \leq x/k \\ p \nmid k}}^a \alpha(p^\nu) \log p^\nu + \int_{1^-}^x \frac{S(t; a)}{t} \, dt$$

$$= \sum_{k \leq x}^a \alpha(k) \sum_{p \leq x/k}^a \alpha(p) \log p + O\left(\sum_{k \leq x} \alpha(k) \sum_{\substack{p \leq x/k \\ p|k}} \alpha(p) \log p\right)$$

$$+ O\left(\sum_{k \leq x} \alpha(k) \sum_{\substack{p^\nu \leq x/k \\ \nu \geq 2}} \alpha(p^\nu) \log p^\nu\right) + O(x(\log x)^{\beta-1})$$

$$= \beta x T(x, a) - \sum_{k \leq x}^a \alpha(k) \sum_{\substack{p \leq x/k \\ p|a}} \alpha(p) \log p$$

$$+ O\left(x \sum_{k \leq x} \frac{\alpha(k)}{k(\log(3x/k))^{A_0}}\right) + O(x(\log x)^{\beta-1}).$$

By partial summation we have

(3.9) $\quad \displaystyle\sum_{k \leq x} \frac{\alpha(k)}{k(\log(3x/k))^{A_0}} = \frac{S(x)}{x(\log 3)^{A_0}} + \int_{1^-}^x \frac{\log(3x/t) - A_0}{t^2(\log(3x/t))^{A_0+1}} S(t) \, dt$

$$\ll (\log x)^{\beta-1} + \int_1^x \frac{(\log 3t)^{\beta-1}}{t(\log(3x/t))^{A_0}} \, dt$$

$$= (\log x)^{\beta-1} + \int_0^{\log x} \frac{(\log 3 + t)^{\beta-1}}{(\log 3x - t)^{A_0}} \, dt$$

$$\ll (\log x)^{\beta-1} + \frac{1}{(\log x)^{A_0}} \int\limits_{0}^{(\log x)/2} (\log 3 + t)^{\beta-1} \, dt$$

$$+ (\log x)^{\beta-1} \int\limits_{(\log x)/2}^{\log x} \frac{1}{(\log 3x - t)^{A_0}} \, dt$$

$$\ll (\log x)^{\beta-A_0}.$$

Let $x_1 := x/(\log x)^2$. For $k \le x_1$ we see that

$$\sum_{\substack{p \le x/k \\ p|a}} \alpha(p) \log p \ll (\log \log x)^{\vartheta_0} \log a \ll \frac{x}{k(\log(x/k))^{A_0}},$$

so that

$$(3.10) \quad \sum_{k \le x_1}^{a} \alpha(k) \sum_{\substack{p \le x/k \\ p|a}} \alpha(p) \log p \ll x \sum_{k \le x} \frac{\alpha(k)}{k(\log(3x/k))^{A_0}} \ll x(\log x)^{\beta-A_0}.$$

On the other hand, by (2.2) we have

$$(3.11) \quad \sum_{x_1 < k \le x}^{a} \alpha(k) \sum_{\substack{p \le x/k \\ p|a}} \alpha(p) \log p$$

$$\ll x \sum_{x_1 < k \le x} \frac{\alpha(k)}{k} \ll x\big((\log x)^{\beta} - (\log x_1)^{\beta} + O((\log x)^{\beta-1})\big)$$

$$\ll x(\log x)^{\beta-1} \log \log x,$$

where we have used (3.7) to estimate the sum over k and the mean value theorem to get

$$(\log x)^{\beta} - (\log x_1)^{\beta} = \beta \xi^{\beta-1} \log \frac{x}{x_1} \ll (\log x)^{\beta-1} \log \log x$$

for some $\xi \in (\log x_1, \log x)$. Combining (3.10) with (3.11), we obtain

$$\sum_{k \le x}^{a} \alpha(k) \sum_{\substack{p \le x/k \\ p|a}} \alpha(p) \log p \ll x(\log x)^{\beta-A_0}.$$

Inserting this and (3.9) into (3.8) yields

$$(3.12) \qquad S(x; a) = \frac{\beta x}{\log x} T(x; a) + O\big(x(\log x)^{\beta-1-A_0}\big)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$.

It remains to estimate $T(x; a)$. To this end, we repeat the argument above with $\alpha(n)$ replaced by $\alpha(n)/n$. From (2.2) it follows that

$$(3.13) \qquad \sum_{p \leq x} \frac{\alpha(p)}{p} \log p = \beta \log x + O\big((\log x)^{1-A_0}\big).$$

Thus,

$$
\begin{aligned}
T(x;a) \log x &= \sum_{n \leq x}^a \frac{\alpha(n)}{n} \log n + \sum_{n \leq x}^a \frac{\alpha(n)}{n} \log \frac{x}{n} \\
&= \sum_{k \leq x}^a \frac{\alpha(k)}{k} \sum_{\substack{p^\nu \leq x/k \\ p \nmid k}}^a \frac{\alpha(p^\nu)}{p^\nu} \log p^\nu + U(x;a) \\
&= \sum_{k \leq x}^a \frac{\alpha(k)}{k} \sum_{p \leq x/k}^a \frac{\alpha(p)}{p} \log p + O\bigg(\sum_{k \leq x}^a \frac{\alpha(k)}{k} \sum_{\substack{p \leq x/k \\ p|k}} \frac{\alpha(p)}{p} \log p \bigg) \\
&\quad + O\bigg(\sum_{k \leq x}^a \frac{\alpha(k)}{k} \sum_{\substack{p^\nu \leq x/k \\ \nu \geq 2}} \frac{\alpha(p^\nu)}{p^\nu} \log p^\nu \bigg) + U(x;a) \\
&= (\beta+1)U(x;a) - \sum_{k \leq x}^a \frac{\alpha(k)}{k} \sum_{\substack{p \leq x/k \\ p|a}} \frac{\alpha(p)}{p} \log p \\
&\quad + O\big((\log x)^{1-A_0} T(x;a)\big),
\end{aligned}
$$

where

$$(3.14) \qquad U(x;a) := \sum_{n \leq x}^a \frac{\alpha(n)}{n} \log \frac{x}{n} = \int_{1^-}^{x} \frac{T(t;a)}{t}\, dt.$$

In view of (3.13), we have

$$
\begin{aligned}
\sum_{\substack{p \leq x/k \\ p|a}} \frac{\alpha(p)}{p} \log p &\leq \sum_{p \leq (\log x)^2} \frac{\alpha(p)}{p} \log p + \sum_{\substack{(\log x)^2 < p \leq x \\ p|a}} \frac{\alpha(p)}{p} \log p \\
&\ll \log\log x + (\log\log x)^{\vartheta_0} \sum_{\substack{(\log x)^2 < p \leq x \\ p|a}} \frac{\log p}{p} \\
&\ll \log\log x + (\log\log x)^{\vartheta_0} \omega(a) \frac{\log\log x}{(\log x)^2} \ll \log\log x,
\end{aligned}
$$

so that

$$\sum_{k \leq x}^a \frac{\alpha(k)}{k} \sum_{\substack{p \leq x/k \\ p|a}} \frac{\alpha(p)}{p} \log p \ll (\log\log x) T(x;a).$$

It follows that

$$T(x; a) \log x = (\beta + 1)U(x; a) + O\big((\log x)^{1-A_0} T(x; a)\big).$$

Hence, there exists a function $\epsilon(x; a)$ such that $\epsilon(x; a) = O((\log x)^{-A_0})$ and

$$(3.15) \qquad T(x; a) = \frac{1}{1 - \epsilon(x; a)} \cdot \frac{\beta + 1}{\log x} U(x; a)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$.

Finally, we estimate $U(x; a)$ and $T(x; a)$ by following [30, proof of Theorem A]. For $y \geq 2$ and $a \in \mathbb{N} \cap [1, y]$, let

$$V(y; a) := \log\Big(\frac{\beta + 1}{(\log y)^{\beta+1}} U(y; a) \Big).$$

In light of (3.14) and (3.15), we have

$$\frac{d}{dy} V(y; a) = -\frac{\beta + 1}{y \log y} + \frac{1}{U(y; a)} \cdot \frac{d}{dy} U(y; a)$$

$$= -\frac{\beta + 1}{y \log y} + \frac{T(y; a)}{U(y; a) y} = \frac{\beta + 1}{y \log y} \cdot \frac{\epsilon(y; a)}{1 - \epsilon(y; a)} \ll \frac{1}{y (\log y)^{A_0+1}}$$

uniformly for all sufficiently large y and all $a \in \mathbb{N} \cap [1, y]$, which implies that

$$V_a := \int_2^\infty \frac{d}{dy} V(y; a) \, dy < \infty.$$

Since

$$V(x; a) - V(2; a) = V_a - \int_x^\infty \frac{d}{dy} V(y; a) \, dy = V_a + O((\log x)^{-A_0})$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$, it follows that

$$\frac{\beta + 1}{(\log x)^{\beta+1}} U(x; a) = \exp(V(x; a)) = \exp(V_a + V(2; a))\big(1 + O((\log x)^{-A_0})\big).$$

Combining this estimate with (3.15), we infer that

$$(3.16) \qquad T(x; a) = \exp(V_a + V(2; a))(\log x)^\beta \big(1 + O((\log x)^{-A_0})\big)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1, x]$. The leading coefficient can be made explicit by arguing as in [30, proof of Theorem A]. Alternatively, we can also take advantage of (3.6). Fixing $a \in \mathbb{N}$, we deduce by (3.6) with $\sigma_0 = 1$ that

$$T(x; a) = \frac{\lambda_\alpha(a)}{\beta} (\log x)^\beta \big(1 + O((\log x)^{-A_0})\big)$$

for all sufficiently large x. Comparing this with (3.16) shows $\exp(V_a + V(2; a))$

$= \lambda_\alpha(a)/\beta$. Carrying this back into (3.16), we obtain

$$T(x;a) = \frac{\lambda_\alpha(a)}{\beta}(\log x)^\beta\big(1 + O((\log x)^{-A_0})\big)$$

uniformly for all sufficiently large x and all $a \in \mathbb{N} \cap [1,x]$. Inserting the above into (3.12) completes the proof of Lemma 3.2. ∎

The next result, which is key to the computation of moments, is a direct corollary of Lemmas 3.1 and 3.2.

LEMMA 3.3. *Fix* $\epsilon_0 \in (0,1)$, *and let* $\alpha \in \mathcal{M}^*$ *with parameters* A_0, β, σ_0, ϑ_0, ϱ_0, r. *Then there exist constants* $\delta_0 > 0$ *and* $Q_0 \geq 2$ *such that, uniformly for all sufficiently large* x, *any* $\delta \in [\delta_0 \log\log x/\log x, 1]$, *and any squarefree* $a \in \mathbb{N} \cap [1,x]$ *with* $\omega(a) \leq (1-\rho_0)\epsilon_0\delta^{-1}$, $P^-(a) > Q_0$ *and* $P^+(a) \leq x^\delta$, *we have*

$$\sum_{\substack{n \leq x \\ a|n}} \alpha(n) = \lambda_\alpha\bigg(F(\sigma_0, a)$$

$$+ O\bigg(\frac{2^{O(\omega(a))}L(a)}{a}\bigg(\frac{1}{(\log x)^{A_0}} + \frac{\epsilon_{\beta,1}\log P^+(a)}{\log x}\bigg)\bigg)\bigg) x^{\sigma_0}(\log x)^{\beta-1},$$

where $L(a)$ *is defined as in Lemma 3.1,*

$$F(\sigma_0, a) := \prod_{p|a}\bigg(1 - \bigg(\sum_{\nu=0}^\infty \alpha(p^\nu)p^{-\sigma_0\nu}\bigg)^{-1}\bigg),$$

and λ_α *is defined by* (3.5).

Proof. Suppose that $\delta_0 > 0$ is a constant for which Lemma 3.1 holds when $c_0 = 1$ and $h \in \{\beta - 1, \beta - 1 - A_0\}$. Let $Q_0 \geq 2$ be such that

$$\sum_{\nu=1}^\infty \frac{\alpha(p^\nu)}{p^{\sigma_0\nu}} \leq \frac{1}{2}$$

for all $p > Q_0$. Then

$$(3.17) \qquad F(\sigma_0, p) = \sum_{\nu=1}^\infty \frac{\alpha(p^\nu)}{p^{\sigma_0\nu}} + O\bigg(\bigg(\sum_{\nu=1}^\infty \frac{\alpha(p^\nu)}{p^{\sigma_0\nu}}\bigg)^2\bigg)$$

$$= \frac{\alpha(p)}{p^{\sigma_0}} + O\bigg(\psi_0(p) + \frac{\alpha(p)^2}{p^{2\sigma_0}}\bigg)$$

for all $p > Q_0$. For any squarefree integer $a \in [1,x]$ with $\omega(a) \leq (1-\varrho_0)\epsilon_0\delta^{-1}$, $P^-(a) > Q_0$ and $P^+(a) \leq x^\delta$, by Lemma 3.2 we have

$$(3.18) \qquad \sum_{\substack{n \leq x \\ \gcd(n,a)=1}} \alpha(n) = x^{\sigma_0}(\log 3x)^{\beta-1}\bigg(\lambda_\alpha(a) + O\bigg(\frac{1}{(\log 3x)^{A_0}}\bigg)\bigg).$$

Note that

$$\sum_{\substack{n \le x \\ a \mid n}} \alpha(n) = \sum_{\substack{q \le x \\ R_q = a}} \alpha(q) \sum_{\substack{n' \le x/q \\ \gcd(n',a)=1}} \alpha(n').$$

By (3.18), the main term of the inner sum contributes

$$\lambda_\alpha(a) x^{\sigma_0} \sum_{\substack{q \le x \\ R_q = a}} \alpha(q) \left(\log \frac{3x}{q} \right)^{\beta - 1},$$

which, by Lemma 3.1, is equal to

$$\lambda_\alpha(a) x^{\sigma_0} \left(\tilde{\lambda}_\alpha(a) + O\left(\frac{2^{O(\omega(a))}}{\log x} \left(\frac{1}{x^{\delta\omega(a)}} + \frac{\epsilon_{\beta,1} L(a) \log P^+(a)}{a} \right) \right) \right) (\log x)^{\beta - 1}$$

$$= \lambda_\alpha \left(F(\sigma_0, a) + O\left(\frac{2^{O(\omega(a))}}{\log x} \left(\frac{1}{a} + \frac{\epsilon_{\beta,1} L(a) \log P^+(a)}{a} \right) \right) \right) x^{\sigma_0} (\log x)^{\beta - 1},$$

since $a \le x^{\delta\omega(a)}$. Analogously, the contribution from the error term of the inner sum is

$$\ll \lambda_\alpha \left(F(\sigma_0, a) + \frac{2^{O(\omega(a))} L(a) \log P^+(a)}{a \log x} \right) x^{\sigma_0} (\log x)^{\beta - 1 - A_0}$$

$$\ll \frac{\lambda_\alpha 2^{O(\omega(a))} L(a)}{a} x^{\sigma_0} (\log x)^{\beta - 1 - A_0},$$

where we have used the estimate $F(\sigma_0, a) \ll 2^{O(\omega(a))} L(a)/a$, which follows directly from (2.4) and (3.17). Combining these estimates completes the proof of Lemma 3.3. ∎

REMARK 3.1. We point out that the lower bound Q_0 for $\omega(a)$ in the lemma above is by and large an artificial thing, whose value is insignificant for our applications. However, we need it because (3.17) may not hold for small primes. As we shall see later, having such a lower bound also frees us from dealing with minor contributions from small primes.

4. Computing moments. By rescaling the strongly additive function f in Theorem 2.1, we may assume, without loss of generality, that $|f(p)| \le 1$ for all primes p. Note that $0 \le F(\sigma_0, p) < 1$ for all primes p. For every p we define $f_p \colon \mathbb{N} \to \mathbb{R}$ by

$$f_p(n) := \begin{cases} f(p)(1 - F(\sigma_0, p)) & \text{if } p \mid n, \\ -f(p) F(\sigma_0, p) & \text{otherwise.} \end{cases}$$

Given any $q \in \mathbb{N}$ we may also extend f_p via complete multiplicativity by setting

$$f_q(n) := \prod_{p^\nu \| q} f_p(n)^\nu.$$

It is clear that $|f_q(n)| \leq 1$. The following result provides an approximation of the moments of f in terms of those of f_p.

LEMMA 4.1. *Let $\alpha \in \mathcal{M}^*$ with parameters $A_0, \beta, \sigma_0, \vartheta_0, \varrho_0, r$, and let $f \colon \mathbb{N} \to \mathbb{R}$ be a strongly additive function with $|f(p)| \leq 1$ for all p. Then there exists a constant $Q_0 \geq 2$ such that*

$$\sum_{n \leq y} \alpha(n)(f(n) - A(x))^m = \sum_{n \leq y} \alpha(n) \Big(\sum_{Q_0 < p \leq z} f_p(n) \Big)^m + O(E(y, z, w; m))$$

uniformly for all sufficiently large $x \geq z$, any $y \geq 1$, and all $m \in \mathbb{N}$, where

$$E(y, z, w; m)$$
$$:= \sum_{\substack{a+b+c=m \\ 0 \leq a < m \\ b,c \geq 0}} \binom{m}{a, b, c} 2^{O(m-a)} (\log(v+2))^c \sum_{n \leq y} \alpha(n) \Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a \omega(n; z, w)^b,$$

$v := \log x / \log z$, $w := x^{1/\log(v+2)}$, *and*

$$\omega(n; z, w) := \sum_{\substack{z < p \leq w \\ p \mid n}} 1.$$

Proof. Let $Q_0 \geq 2$ be a constant for which (3.17) holds. Suppose that $z > Q_0$ is sufficiently large. By (2.4), (3.17) and the fact that $\sum_p \psi_0(p) < \infty$, we find

$$\sum_{Q_0 < p \leq x} f(p) F(\sigma_0, p) = A(x) + O(1).$$

We compute

$$f(n) - A(x) = \sum_{\substack{p \mid n \\ p > Q_0}} f(p) - \sum_{Q_0 < p \leq x} f(p) F(\sigma_0, p) + O(1)$$

$$= \sum_{\substack{Q_0 < p \leq z \\ p \mid n}} f(p) + \sum_{\substack{p > z \\ p \mid n}} f(p)$$

$$\quad - \sum_{Q_0 < p \leq z} f(p) F(\sigma_0, p) - \sum_{z < p \leq x} f(p) F(\sigma_0, p) + O(1)$$

$$= \sum_{Q_0 < p \leq z} f_p(n) + \sum_{\substack{p > z \\ p \mid n}} f(p) - \sum_{z < p \leq x} f(p) F(\sigma_0, p) + O(1).$$

By (3.1) we have

$$\left| \sum_{z < p \le x} f(p) F(\sigma_0, p) \right| \le \sum_{z < p \le x} \frac{\alpha(p)}{p^{\sigma_0}} + O(1) \le \beta \log(v+2) + O(1).$$

Since

$$\sum_{\substack{p > z \\ p | n}} |f(p)| \le \sum_{\substack{z < p \le x \\ p | n}} 1 < \omega(n; z, w) + \log(v+2),$$

it follows that

$$f(n) - A(x) = \sum_{Q_0 < p \le z} f_p(n) + O(\omega(n; z, w) + \log(v+2)).$$

We have thus proved

$$\sum_{n \le y} \alpha(n)(f(n) - A(x))^m$$
$$= \sum_{n \le y} \alpha(n) \Big(\sum_{Q_0 < p \le z} f_p(n) + O(\omega(n; z, w) + \log(v+2)) \Big)^m.$$

Opening the mth power on the right-hand side by means of the multinomial theorem completes the proof of Lemma 4.1. ∎

Let $z = x^{1/v}$ and $w = x^{1/\log(v+2)}$ be as in Lemma 4.1, where $v \ge 1$ is a function of x and m to be chosen later. Fix $\epsilon_0 \in (0, 1)$ and $\eta_0 \in (0, 1]$, and suppose that $y \in [x^{\eta_0}, x]$. Under the hypotheses in Theorem 2.1, we seek to estimate the weighted moments

$$\sum_{n \le y} \alpha(n) \Big(\sum_{Q_0 < p \le z} f_p(n) \Big)^m,$$

appearing in Lemma 4.1. Expanding out the mth power we see that

$$(4.1) \qquad \sum_{n \le y} \alpha(n) \Big(\sum_{Q_0 < p \le z} f_p(n) \Big)^m = \sum_{Q_0 < p_1, \ldots, p_m \le z} \sum_{n \le y} \alpha(n) f_{p_1 \cdots p_m}(n).$$

This suggests studying the sum

$$\sum_{n \le y} \alpha(n) f_q(n)$$

for $q \in \mathbb{N}$ with $\omega(q) \le m$, $P^-(q) > Q_0$ and $P^+(q) \le z$. A key observation is that $f_q(n) = f_q(\gcd(n, R_q))$. From this we deduce

$$\sum_{n \le y} \alpha(n) f_q(n) = \sum_{a | R_q} f_q(a) \sum_{\substack{n \le y \\ \gcd(n, R_q) = a}} \alpha(n) = \sum_{ab | R_q} f_q(a) \mu(b) \sum_{\substack{n \le y \\ ab | n}} \alpha(n).$$

Note that $\log y / \log z \in [\eta_0 v, v]$. By Lemma 3.3, there exists a constant

$v_0 > 0$, independent of Q_0 and η_0, such that

$$(4.2) \qquad \sum_{n \leq y} \alpha(n) f_q(n) = \lambda_\alpha \big(G(\sigma_0, q) + O(2^{O(m)} E_y(q)) \big) y^{\sigma_0} (\log y)^{\beta - 1}$$

uniformly for all sufficiently large x, any $y \in [x^{\eta_0}, x]$ and

$$v \in [\eta_0^{-1}, v_0 \log x / \log \log x],$$

and all $m \leq (1 - \varrho_0)\epsilon_0 \log y / \log z$, where

$$G(\sigma_0, q) := \sum_{ab | R_q} f_q(a) \mu(b) F(\sigma_0, ab),$$

$$E_y(q) := \sum_{ab | R_q} \frac{|f_q(a)| L(ab)}{ab} \left(\frac{1}{(\log y)^{A_0}} + \frac{\epsilon_{\beta,1} \log P^+(ab)}{\log y} \right).$$

Combining (4.2) with (4.1) gives

$$(4.3) \qquad \sum_{n \leq y} \alpha(n) \Big(\sum_{Q_0 < p \leq z} f_p(n) \Big)^m$$
$$= \lambda_\alpha \big(G(z) + O(2^{O(m)} D(y, z)) \big) y^{\sigma_0} (\log y)^{\beta - 1},$$

where

$$G(z) := \sum_{Q_0 < p_1, \ldots, p_m \leq z} G(\sigma_0, p_1 \cdots p_m),$$

$$D(y, z) := \sum_{Q_0 < p_1, \ldots, p_m \leq z} E_y(p_1 \cdots p_m).$$

5. Estimation of $G(z)$ and $D(y, z)$. It is easy to see that $G(\sigma_0, q)$ is multiplicative as a function of q. Indeed, given any $q_1, q_2 \in \mathbb{N}$ with $\gcd(q_1, q_2) = 1$, we have

$$G(\sigma_0, q_1) G(\sigma_0, q_2) = \sum_{\substack{a_1 b_1 | R_{q_1} \\ a_2 b_2 | R_{q_2}}} f_{q_1}(a_1) f_{q_2}(a_2) \mu(b_1) \mu(b_2) F(\sigma_0, a_1 b_1) F(\sigma_0, a_2 b_2)$$

$$= \sum_{\substack{a_1 b_1 | R_{q_1} \\ a_2 b_2 | R_{q_2}}} f_{q_1}(a_1 a_2) f_{q_2}(a_1 a_2) \mu(b_1 b_2) F(\sigma_0, a_1 a_2 b_1 b_2)$$

$$= \sum_{\substack{a_1 b_1 | R_{q_1} \\ a_2 b_2 | R_{q_2}}} f_{q_1 q_2}(a_1 a_2) \mu(b_1 b_2) F(\sigma_0, a_1 a_2 b_1 b_2)$$

$$= \sum_{ab | R_{q_1 q_2}} f_{q_1 q_2}(a) \mu(b) F(\sigma_0, ab) = G(\sigma_0, q_1 q_2).$$

Furthermore,

$$
\begin{aligned}
G(\sigma_0, p^\nu) &= f_{p^\nu}(1) + f_{p^\nu}(p)F(\sigma_0, p) - f_{p^\nu}(1)F(\sigma_0, p) \\
&= (-f(p)F(\sigma_0, p))^\nu + \big(f(p)(1 - F(\sigma_0, p))\big)^\nu F(\sigma_0, p) \\
&\quad - (-f(p)F(\sigma_0, p))^\nu F(\sigma_0, p) \\
&= f(p)^\nu F(\sigma_0, p)(1 - F(\sigma_0, p)) \\
&\quad \times \big((-1)^\nu F(\sigma_0, p)^{\nu-1} + (1 - F(\sigma_0, p))^{\nu-1}\big)
\end{aligned}
$$

for all prime powers p^ν. Note that $G(\sigma_0, p) = 0$, $|G(\sigma_0, p^\nu)| \le 1/4$, and $G(\sigma_0, p^\nu) \ge 0$ when $2 \mid \nu$. In addition, by (3.17),

$$
\begin{aligned}
(5.1) \qquad G(\sigma_0, p^2) &= f(p)^2 F(\sigma_0, p)(1 - F(\sigma_0, p)) \\
&= \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} + O\left(\psi_0(p) + \frac{\alpha(p)^2}{p^{2\sigma_0}}\right)
\end{aligned}
$$

and

$$
(5.2) \qquad |G(\sigma_0, p^\nu)| \le |f(p)|^\nu F(\sigma_0, p) \le \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} + O\left(\psi_0(p) + \frac{\alpha(p)^2}{p^{2\sigma_0}}\right)
$$

for all p^ν with $p > Q_0$ and $\nu \ge 2$.

Now we proceed to estimate $G(z)$ in the main term of (4.3). Recall that $y \in [x^{\eta_0}, x]$ and $z = x^{1/v}$. We shall suppose in this section that $1 \le m \le \min(v, h_0 B(x)^{1/3})$, $\log(v+2) = o(B(x))$, and $m\log(v+2) \ll B(x)$, where $0 < h_0 < (3/2)^{2/3}$ is any given constant, and obtain a uniform treatment for $G(z)$ and $D(y, z)$ under this more general assumption. Since $G(\sigma_0, q)$ is multiplicative in q and $G(\sigma_0, p) = 0$ for all $p > Q_0$, we have

$$
(5.3) \qquad G(z) = \sum_{\substack{Q_0 < p_1, \dots, p_m \le z \\ p_1 \cdots p_m \text{ squareful}}} G(\sigma_0, p_1 \cdots p_m).
$$

When $2 \mid m$, the main contribution arises from

$$
(5.4) \qquad \frac{m!}{(m/2)!2^{m/2}} \sum_{\substack{Q_0 < p_1, \dots, p_{m/2} \le z \\ p_1, \dots, p_{m/2} \text{ distinct}}} G(\sigma_0, p_1^2 \cdots p_{m/2}^2)
$$

$$
= C_m \sum_{\substack{Q_0 < p_1, \dots, p_{m/2} \le z \\ p_1, \dots, p_{m/2} \text{ distinct}}} \prod_{i=1}^{m/2} G(\sigma_0, p_i^2),
$$

since the number of ways to partition a set of m elements into $m/2$ two-element equivalence classes is

$$
\frac{m!}{(m/2)!2^{m/2}} = \frac{m!}{m!!} = C_m.
$$

The sum on the right-hand side of (5.4) can be rewritten as

$$\sum_{\substack{Q_0 < p_1,\ldots,p_{m/2-1} \leq z \\ p_1,\ldots,p_{m/2-1} \text{ distinct}}} \prod_{i=1}^{m/2-1} G(\sigma_0, p_i^2) \sum_{\substack{Q_0 < p_{m/2} \leq z \\ p_{m/2} \neq p_1,\ldots,p_{m/2-1}}} G(\sigma_0, p_{m/2}^2).$$

By (5.1) and (3.1), the inner sum over $p_{m/2}$ is equal to

$$\sum_{Q_0 < p \leq z} G(\sigma_0, p^2) - \sum_{i=1}^{m/2-1} G(\sigma_0, p_i^2) = B(z) + O\left(\sum_{Q_0 < p \leq q_N} \frac{\alpha(p)}{p^{\sigma_0}}\right)$$

$$= B(z) + O(\log\log(m+2)),$$

where $N = m/2 + \pi(Q_0)$ and q_N is the Nth prime. Repeating this argument we obtain

$$\sum_{\substack{Q_0 < p_1,\ldots,p_{m/2} \leq z \\ p_1,\ldots,p_{m/2} \text{ distinct}}} \prod_{i=1}^{m/2} G(\sigma_0, p_i^2) = \big(B(z) + O(\log\log(m+2))\big)^{m/2}.$$

But

$$B(x) - B(z) = \sum_{z < p \leq x} \alpha(p) \frac{f(p)^2}{p^{\sigma_0}} \leq \beta \log(v+2) + O(1).$$

Hence when m is even, the main contribution to $G(z)$ is given by

$$C_m \big(B(x) + O(\log(v+2))\big)^{m/2} = C_m B(x)^{m/2}\big(1 + O(m B(x)^{-1} \log(v+2))\big).$$

The remaining contribution to $G(z)$ comes from

$$(5.5) \qquad \sum_{s < m/2} \sum_{Q_0 < p_1 < \cdots < p_s \leq z} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1,\ldots,k_s \geq 2}} \binom{m}{k_1,\ldots,k_s} \prod_{i=1}^{s} G(\sigma_0, p_i^{k_i}).$$

Since (5.5) vanishes when $m \leq 2$, we may suppose $m \geq 3$. By (5.2),

$$\prod_{i=1}^{s} |G(\sigma_0, p_i^{k_i})| \leq \prod_{i=1}^{s} \left(\alpha(p_i) \frac{f(p_i)^2}{p_i^{\sigma_0}} + O\left(\psi_0(p_i) + \frac{\alpha(p_i)^2}{p_i^{2\sigma_0}}\right)\right).$$

Thus,

$$\sum_{Q_0 < p_1 < \cdots < p_s \leq z} \prod_{i=1}^{s} |G(\sigma_0, p_i^{k_i})|$$

$$\leq \frac{1}{s!}(B(x) + O(1))^s = \frac{1}{s!} B(x)^s\big(1 + O(s B(x)^{-1})\big) \ll \frac{B(x)^s}{s!}.$$

Since

$$\sum_{\substack{k_1+\cdots+k_s=m \\ k_1,\ldots,k_s \geq 2}} \binom{m}{k_1,\ldots,k_s} \leq \frac{m!}{2^s} \sum_{\substack{k_1+\cdots+k_s=m \\ k_1,\ldots,k_s \geq 2}} 1 = \frac{m!}{2^s}\binom{m-s-1}{s-1}, \quad (^2)$$

the sum (5.5) is

$$\ll m! \sum_{s<m/2} \frac{1}{s!2^s}\binom{m-s-1}{s-1}B(x)^s.$$

To further estimate the latter sum, we put $m_1 := \lfloor (m-1)/2 \rfloor$ and observe that

$$\sum_{s<m/2} \frac{1}{s!2^s}\binom{m-s-1}{s-1}B(x)^s$$

$$= B(x)^{m_1} \sum_{s\leq m_1} \frac{1}{s!2^s}\binom{m-s-1}{s-1}B(x)^{s-m_1}$$

$$\leq B(x)^{m_1} m^{-3m_1} \sum_{s\leq m_1} \frac{1}{s!2^s}\binom{m-s-1}{s-1}h_0^{3(m_1-s)}m^{3s},$$

where we have used the assumption that $B(x) \geq m^3/h_0^3$ with some $0 < h_0 < (3/2)^{2/3}$. Let

$$e_m := \begin{cases} 1 & \text{if } 2\,|\,m, \\ 1/2 & \text{otherwise.} \end{cases}$$

Then $m_1 = m/2 - e_m$. Note that

$$m^{-3m_1} \sum_{s\leq m/4} \frac{1}{s!2^s}\binom{m-s-1}{s-1}h_0^{3(m_1-s)}m^{3s}$$

$$\leq m^{-3m_1} \sum_{s\leq m/4} \frac{1}{s!(s-1)!}\left(\frac{9}{4}\right)^{m_1-s}\left(\frac{m^4}{2}\right)^s$$

$$\ll m^{-3m_1}\left(\frac{9}{4}\right)^{m_1}\left(\frac{m^4}{2}\right)^{m/4} \ll \frac{C_m}{m!}m^{3e_m},$$

since

$$\frac{C_m}{m!} = \frac{1}{2^{m/2}\Gamma(m/2+1)} \asymp m^{-(m+1)/2}e^{m/2}.$$

$(^2)$ We have corrected the binomial coefficient in [17, (11)].

by Stirling's formula. Next,

$$m^{-3m_1} \sum_{m/4 < s \leq m/3} \frac{1}{s! 2^s} \binom{m-s-1}{s-1} h_0^{3(m_1-s)} m^{3s}$$

$$\leq 2^{O(m)} m^{-3m_1} \sum_{m/4 < s \leq m/3} \frac{1}{s!(s-1)!} \left(\frac{m^4}{2} \right)^s$$

$$\leq \frac{2^{O(m)} m^{-3m_1}}{m^{m/2}} \sum_{m/4 < s \leq m/3} \left(\frac{m^4}{2} \right)^s$$

$$\leq \frac{2^{O(m)} m^{-3m_1}}{m^{m/2}} m^{4m/3} = 2^{O(m)} m^{-2m/3 + 3e_m}$$

$$\ll \frac{C_m}{m!} m^{3e_m}.$$

Finally, we observe that

$$m^{-3m_1} \sum_{m/3 < s \leq m_1} \frac{1}{s! 2^s} \binom{m-s-1}{s-1} h_0^{3(m_1-s)} m^{3s}$$

$$= m^{-3m_1} \sum_{m/3 < s \leq m_1} \frac{1}{s! 2^s} \binom{m-s-1}{m-2s} h_0^{3(m_1-s)} m^{3s}$$

$$\leq m^{-3m_1} \sum_{m/3 < s \leq m_1} \frac{1}{s! 2^s} (m-s)^{m-2s} h_0^{3(m_1-s)} m^{3s}$$

$$\leq \frac{m^{-3m_1}}{m_1!} \sum_{m/3 < s \leq m_1} \frac{m_1!}{s! 2^s} \left(\frac{2m}{3} \right)^{m-2s} h_0^{3(m_1-s)} m^{3s}$$

$$\leq \frac{m^{-3m_1}}{m_1!} \sum_{m/3 < s \leq m_1} \frac{1}{2^s} \left(\frac{m}{2} \right)^{m_1-s} \left(\frac{2m}{3} \right)^{m-2s} h_0^{3(m_1-s)} m^{3s}$$

$$\ll \frac{m^{m-2m_1}}{m_1! 2^{m/2}} \sum_{m/3 < s \leq m_1} \left(\frac{2 h_0^{3/2}}{3} \right)^{m-2s} \ll \frac{C_m}{m!} m^{3e_m}.$$

Collecting the estimates above, we deduce that the contribution to $G(z)$ from (5.5) is

$$\ll C_m m^{3e_m} B(x)^{m_1} = C_m B(x)^{m/2} \left(\frac{m^3}{B(x)} \right)^{e_m} \leq C_m B(x)^{m/2} \frac{m^{3/2}}{\sqrt{B(x)}}.$$

We can therefore conclude that

$$(5.6) \quad G(z) = C_m B(x)^{m/2} \left(\chi_m \left(1 + O \left(\frac{m \log(v+2)}{B(x)} \right) \right) + O \left(\frac{m^{3/2}}{\sqrt{B(x)}} \right) \right).$$

Next, we estimate $D(y, z)$ in the error term of (4.3). By definition,

$$D(y, z) = \sum_{s \le m} \sum_{Q_0 < p_1 < \cdots < p_s \le z} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} E_y(p_1^{k_1} \cdots p_s^{k_s}).$$

Let

$$H(\sigma_0, q) := \sum_{ab | R_q} \frac{|f_q(a)| L(ab)}{ab}.$$

Then $H(\sigma_0, q)$ is multiplicative in q. Moreover,

$$E_y(q) \le H(\sigma_0, q) \left(\frac{1}{(\log y)^{A_0}} + \frac{\epsilon_{\beta,1} \log P^+(q)}{\log y} \right).$$

It follows that $D(y, z) \le D_1(y, z) + \epsilon_{\beta,1} D_2(y, z)$, where

$$D_1(y, z) := \frac{1}{(\log y)^{A_0}}$$

$$\times \sum_{s \le m} \sum_{Q_0 < p_1 < \cdots < p_s \le z} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} \prod_{i=1}^{s} H(\sigma_0, p_i^{k_i}),$$

$$D_2(y, z) := \frac{1}{\log y}$$

$$\times \sum_{s \le m} \sum_{Q_0 < p_1 < \cdots < p_s \le z} \log p_s \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} \prod_{i=1}^{s} H(\sigma_0, p_i^{k_i}).$$

By Mertens' theorems [22, Theorems 425, 427], for any $t \ge 3$,

$$(5.7) \qquad \sum_{p \le t} \frac{(\log \log(p+1))^{\vartheta_0}}{p} = \frac{1}{\vartheta_0 + 1} (\log \log t)^{\vartheta_0 + 1} + O(1)$$

and

$$(5.8) \qquad \sum_{p \le t} \frac{(\log \log(p+1))^{\vartheta_0} \log p}{p}$$

$$= \left(1 + O\left(\frac{1}{\log t} + \frac{\vartheta_0}{\log \log t} \right) \right) (\log \log t)^{\vartheta_0} \log t.$$

Furthermore, let

$$T_n(t) := \sum_{k=0}^{n} \left\{ {n \atop k} \right\} t^k$$

denote the nth Touchard polynomial, where

$$\begin{Bmatrix} n \\ k \end{Bmatrix} := \frac{1}{k!} \sum_{\substack{n_1+\cdots+n_k=n \\ n_1,\ldots,n_k\in\mathbb{N}}} \binom{n}{n_1,\ldots,n_k}$$

is the kth Stirling number of the second kind of size n. The sequence $\{T_n(t)\}_{n=0}^{\infty}$ of the Touchard polynomials is known to satisfy the recurrence relation

$$T_{n+1}(t) = t \sum_{i=0}^{n} \binom{n}{i} T_i(t),$$

from which one verifies readily by induction that

(5.9) $$T_n(t) \le \left(t + \frac{n-1}{2}\right)^n$$

for all $n \ge 1$ and $t \ge 0$. Since

$$H(\sigma_0, p^\nu) = |f(p)|F(\sigma_0,p)\left(1 + \frac{L(p)}{p}\right) + \frac{|f(p)|L(p)}{p}(1 - F(\sigma_0,p))$$

$$= |f(p)|\left(F(\sigma_0,p) + \frac{(\log\log(p+1))^{\vartheta_0}}{p}\right)$$

for any prime powers p^ν with $p > Q_0$, we deduce from (3.17), (5.7), (5.8) and (5.9) that

$$D_1(y,z) \le \frac{2^{O(m)}}{(\log x)^{A_0}} \sum_{s\le m} \frac{1}{s!} (\log\log z)^{s(\vartheta_0+1)} \sum_{\substack{k_1+\cdots+k_s=m \\ k_1,\ldots,k_s\in\mathbb{N}}} \binom{m}{k_1,\ldots,k_s}$$

$$\le \frac{2^{O(m)}}{(\log x)^{A_0}} T_m((\log\log z)^{\vartheta_0+1}) \le \frac{2^{O(m)}}{(\log x)^{A_0}} (\log\log x)^{m(\vartheta_0+1)},$$

and

$$D_2(y,z) \le \frac{2^{O(m)}\log z}{\log x} \sum_{s\le m} \frac{1}{(s-1)!} (\log\log z)^{s(\vartheta_0+1)-1} \sum_{\substack{k_1+\cdots+k_s=m \\ k_1,\ldots,k_s\in\mathbb{N}}} \binom{m}{k_1,\ldots,k_s}$$

$$\le \frac{2^{O(m)}}{v\log\log z} T_m((\log\log z)^{\vartheta_0+1}) \le \frac{2^{O(m)}}{v}(\log\log x)^{m(\vartheta_0+1)-1}.$$

Hence, we conclude that

(5.10) $$D(y,z) \le 2^{O(m)}(\log\log x)^{m(\vartheta_0+1)-1}\left(\frac{\log\log x}{(\log x)^{A_0}} + \frac{\epsilon_{\beta,1}}{v}\right).$$

6. Estimation of $E(y,z,w;m)$. In this section, we seek to bound the function $E(y,z,w;m)$ introduced in Lemma 4.1 under the assumptions of Theorem 2.1. We start with the case $\beta = 1$. Suppose that $1 \le m \le h_0 B(x)^{1/3}$,

where $0 < h_0 < (3/2)^{2/3}$ is any given constant. Recall that $y \in [x^{\eta_0}, x]$, $z = x^{1/v}$ and $w = x^{1/\log(v+2)}$. With the choice $v = (1 - \varrho_0)^{-1}\epsilon_0^{-1}\eta_0^{-1}m$, we clearly have $v \in [\eta_0^{-1}, v_0 \log x/\log\log x]$ and $m \leq (1 - \varrho_0)\epsilon_0 \log y/\log z$. Inserting (5.6) and (5.10) into (4.3), we obtain

$$(6.1) \quad \sum_{n\leq y} \alpha(n)\Big(\sum_{Q_0 < p \leq z} f_p(n) \Big)^m = \lambda_\alpha C_m B(x)^{m/2}\Big(\chi_m + O\Big(\frac{m^{3/2}}{\sqrt{B(x)}}\Big)\Big)y^{\sigma_0}.$$

The key lies in the estimation of the sum

$$(6.2) \quad \sum_{n\leq y} \alpha(n)\Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a \omega(n; z, w)^b.$$

In the present case, we may simply use the trivial bound $\omega(n; z, w) \ll v \ll m$, so that (6.2) is bounded above by

$$2^{O(b)}m^b \sum_{n\leq y} \alpha(n)\Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a.$$

It is clear that we can use (6.1) to handle the sum above. If a is even, then this sum is $\ll \lambda_\alpha C_a B(x)^{a/2}y^{\sigma_0}$; if a is odd, then it is

$$\leq \Big(\sum_{n\leq y} \alpha(n)\Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^{a-1}\Big)^{1/2}\Big(\sum_{n\leq y} \alpha(n)\Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^{a+1}\Big)^{1/2}$$
$$\ll \lambda_\alpha \sqrt{C_{a-1}C_{a+1}}B(x)^{a/2}y^{\sigma_0}$$

by the Cauchy–Schwarz inequality. The sequence $\{C_\ell\}_{\ell=1}^\infty$ is strictly increasing, which can be easily seen from the identity

$$\frac{C_{\ell+1}}{C_\ell} = \frac{\ell+1}{\sqrt{2}} \cdot \frac{\Gamma(\ell/2+1)}{\Gamma((\ell+1)/2+1)} = \sqrt{2} \cdot \frac{\Gamma(\ell/2+1)}{\Gamma((\ell/2+1/2)}$$

and the fact that $\Gamma(y)$ is strictly increasing on $[3/2, \infty)$. Moreover, by Stirling's formula,

$$\frac{C_\ell}{C_{\ell+1}} \ll \frac{1}{\ell+1} \cdot \frac{((\ell+1)/2)^{\ell/2+1}e^{-(\ell+1)/2}}{(\ell/2)^{(\ell+1)/2}e^{-\ell/2}} \ll \frac{1}{\sqrt{\ell+1}},$$

which implies that

$$C_a \leq 2^{O(m-a)}C_m\sqrt{\frac{a!}{m!}} \leq 2^{O(m-a)}C_m\sqrt{\frac{a^a}{m^m}} \leq \frac{2^{O(m-a)}C_m}{(\sqrt{m})^{m-a}}$$

for all $0 \leq a \leq m$. Hence, (6.2) is bounded above by

$$\frac{2^{O(m-a)}\lambda_\alpha C_m m^b}{(\sqrt{m})^{m-a}}B(x)^{a/2}y^{\sigma_0} \leq 2^{O(m-a)}\lambda_\alpha C_m(\sqrt{m})^{m-a}B(x)^{a/2}y^{\sigma_0}.$$

Inserting this inequality into the definition of $E(y, z, w; m)$, we conclude that

$$(6.3) \qquad E(y, z, w; m) \leq \lambda_\alpha C_m y^{\sigma_0} \sum_{a=0}^{m-1} \binom{m}{a} B(x)^{a/2} (O(\sqrt{m}))^{m-a}$$

$$\ll \lambda_\alpha C_m m^{3/2} B(x)^{(m-1)/2} y^{\sigma_0}.$$

Now we consider the case $\beta \neq 1$. Suppose that

$$1 \leq m \ll B(x)^{1/3} / (\log\log\log x)^{2/3}$$

and $B(x)/(\log\log\log x)^2 \to \infty$ as $x \to \infty$. In this case we take

$$v = (\log\log x)^{m(\vartheta_0 + 2)},$$

so that $v \in [2\eta_0^{-1}, v_0 \log x / \log\log x]$ and $m \leq (1 - \varrho_0)\epsilon_0 \log t / \log z$ for any $t \in [x^{\eta_0/2}, x]$ when x is sufficiently large. Inserting (5.6) and (5.10) into (4.3) leads to

$$(6.4) \qquad \sum_{n \leq t} \alpha(n) \Big(\sum_{Q_0 < p \leq z} f_p(n) \Big)^m$$

$$= \lambda_\alpha C_m B(x)^{m/2} \Big(\chi_m + O\Big(\frac{m^{3/2}}{\sqrt{B(x)}} \Big) \Big) t^{\sigma_0} (\log t)^{\beta-1}$$

uniformly for all $t \in [x^{\eta_0/2}, x]$. Again, we need to estimate (6.2) uniformly for $y \in [x^{\eta_0}, x]$. Note that (6.2) can be rewritten as

$$\sum_{k=1}^{b} \sum_{z < p_1 < \cdots < p_k \leq w} \sum_{\substack{l_1 + \cdots + l_k = b \\ l_1, \ldots, l_k \geq 1}} \binom{b}{l_1, \ldots, l_k} \sum_{\substack{n \leq y \\ p_1 \cdots p_k | n}} \alpha(n) \Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a.$$

Observe that

$$\sum_{\substack{n \leq y \\ p_1 \cdots p_k | n}} \alpha(n) \Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a = \sum_{\substack{q \leq y \\ R_q = p_1 \cdots p_k}} \alpha(q) \sum_{\substack{n' \leq y/q \\ \gcd(n', q) = 1}} \alpha(n') \Big| \sum_{Q_0 < p \leq z} f_p(n') \Big|^a$$

$$\leq \sum_{\substack{q \leq y \\ R_q = p_1 \cdots p_k}} \alpha(q) \sum_{n \leq y/q} \alpha(n) \Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a,$$

since $p_1, \ldots, p_k > p$.

If $q = p_1^{\nu_1} \cdots p_k^{\nu_k} > \sqrt{y}$ with given $z < p_1 < \cdots < p_k \leq w$, then we have the trivial estimate

$$\sum_{n \leq y/q} \alpha(n) \Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a \leq \pi(z)^a \sum_{n \leq 3y/q} \alpha(n)$$

$$\ll \lambda_\alpha \pi(z)^a \Big(\frac{y}{q} \Big)^{\sigma_0} \Big(\log \frac{3y}{q} \Big)^{\beta-1}$$

by (3.4) and the fact that $|f_p(n)| \leq 1$. By the proof of Lemma 3.1, and in particular by (3.3), we find that

$$
\sum_{\substack{\sqrt{y} < q \leq y \\ R_q = p_1 \cdots p_k}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{3y}{q} \right)^{\beta - 1}
$$

$$
= \sum_{\substack{\sqrt{y} < p_1^{\nu_1} \cdots p_k^{\nu_k} \leq y \\ \nu_1, \ldots, \nu_k \geq 1}} \frac{\alpha(p_1^{\nu_1}) \cdots \alpha(p_k^{\nu_k})}{p_1^{\sigma_0 \nu_1} \cdots p_k^{\sigma_0 \nu_k}} \left(\log \frac{3y}{p_1^{\nu_1} \cdots p_k^{\nu_k}} \right)^{\beta - 1}
$$

$$
\ll \frac{2^{O(k)} (\log y)^{k + \beta - 2}}{(\sqrt{y})^{1 - \varrho_0}},
$$

from which it follows that

$$
\sum_{\substack{\sqrt{y} < q \leq y \\ R_q = p_1 \cdots p_k}} \alpha(q) \sum_{n \leq y/q} \alpha(n) \Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a \ll \lambda_\alpha \pi(z)^a \frac{2^{O(k)} y^{\sigma_0} (\log y)^{k + \beta - 2}}{(\sqrt{y})^{1 - \varrho_0}}.
$$

Summing the above over all $z < p_1 < \cdots < p_k \leq w$ yields immediately

$$
(6.5) \qquad \sum_{z < p_1 < \cdots < p_k \leq w} \sum_{\substack{\sqrt{y} < q \leq y \\ R_q = p_1 \cdots p_k}} \alpha(q) \sum_{n \leq y/q} \alpha(n) \Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a
$$

$$
\leq \lambda_\alpha \pi(z)^a \pi(w)^k \frac{2^{O(k)} y^{\sigma_0} (\log y)^{k + \beta - 2}}{k! (\sqrt{y})^{1 - \varrho_0}} \leq \frac{\lambda_\alpha y^{\sigma_0} (\log y)^{\beta - 1}}{k! (\sqrt[3]{y})^{1 - \varrho_0}}
$$

for sufficiently large x, since $y \in [x^{\eta_0}, x]$ and

$$
a + k \leq m \ll (\log \log x)^{1/3} / (\log \log \log x)^{2/3},
$$

and

$$
\pi(z)^a \pi(w)^k \leq \left(\frac{w}{\log w} + O\left(\frac{w}{(\log w)^2} \right) \right)^m
$$

$$
\ll \left(\frac{w}{\log w} \right)^m \leq \frac{x^{1/\log \log \log x} (m \log \log \log x)^m}{(\log x)^m}.
$$

If $q = p_1^{\nu_1} \cdots p_k^{\nu_k} \leq \sqrt{y}$, then $x^{\eta_0/2} \leq \sqrt{y} \leq y/q \leq y \leq x$. Thus, we can apply (6.4) with $t = y/q$ to handle

$$
\sum_{n \leq y/q} \alpha(n) \Big| \sum_{Q_0 < p \leq z} f_p(n) \Big|^a.
$$

If a is even, then this sum is

$$\ll \lambda_\alpha C_a B(x)^{a/2} \left(\frac{y}{q}\right)^{\sigma_0} \left(\log \frac{y}{q}\right)^{\beta-1}$$

$$\le \frac{2^{O(m-a)} \lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{a/2} \left(\frac{y}{q}\right)^{\sigma_0} \left(\log \frac{y}{q}\right)^{\beta-1};$$

if a is odd, then it is

$$\le \left(\sum_{n\le y/q} \alpha(n) \Big| \sum_{Q_0<p\le z} f_p(n) \Big|^{a-1}\right)^{1/2} \left(\sum_{n\le y/q} \alpha(n) \Big| \sum_{Q_0<p\le z} f_p(n) \Big|^{a+1}\right)^{1/2}$$

$$\ll \lambda_\alpha \sqrt{C_{a-1} C_{a+1}} B(x)^{a/2} \left(\frac{y}{q}\right)^{\sigma_0} \left(\log \frac{y}{q}\right)^{\beta-1}$$

$$\le \frac{2^{O(m-a)} \lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{a/2} \left(\frac{y}{q}\right)^{\sigma_0} \left(\log \frac{y}{q}\right)^{\beta-1}$$

by Cauchy–Schwarz. It follows that

$$\sum_{\substack{q\le\sqrt{y}\\R_q=p_1\cdots p_k}} \alpha(q) \sum_{n\le y/q} \alpha(n) \Big| \sum_{Q_0<p\le z} f_p(n) \Big|^a$$

$$\le \frac{2^{O(m-a)} \lambda_\alpha C_m y^{\sigma_0}}{(\sqrt{m})^{m-a}} B(x)^{a/2} \sum_{\substack{q\le\sqrt{y}\\R_q=p_1\cdots p_k}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{y}{q}\right)^{\beta-1}$$

$$\le \frac{2^{O(m-a)} \lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{a/2} y^{\sigma_0} (\log y)^{\beta-1} \prod_{i=1}^{k} \sum_{\nu=1}^{\infty} \frac{\alpha(p_i^\nu)}{p_i^{\sigma_0 \nu}}$$

$$= \frac{2^{O(m-a)} \lambda_\alpha C_m}{(\sqrt{m})^{m-a}} B(x)^{a/2} y^{\sigma_0} (\log y)^{\beta-1} \prod_{i=1}^{k} \left(\frac{\alpha(p_i)}{p_i^{\sigma_0}} + \psi_0(p_i)\right)$$

for all $0 \le a < m$. Since (3.1) implies that

$$\sum_{z<p_1<\cdots<p_k\le w} \prod_{i=1}^{k} \left(\frac{\alpha(p_i)}{p_i^{\sigma_0}} + \psi_0(p_i)\right)$$

$$\le \frac{1}{k!} \left(\sum_{z<p\le w} \left(\frac{\alpha(p)}{p^{\sigma_0}} + \psi_0(p)\right)\right)^k \le \frac{2^{O(k)}}{k!} (\log v)^k,$$

we obtain

$$(6.6) \quad \sum_{z<p_1<\cdots<p_k\le w} \sum_{\substack{q\le\sqrt{y}\\R_q=p_1\cdots p_k}} \alpha(q) \sum_{n\le y/q} \alpha(n) \Big| \sum_{Q_0<p\le z} f_p(n) \Big|^a$$

$$\le \frac{2^{O(m-a)} \lambda_\alpha C_m}{k!(\sqrt{m})^{m-a}} (\log v)^k B(x)^{a/2} y^{\sigma_0} (\log y)^{\beta-1}.$$

Combining (6.6) with (6.5) and extending the inner sum over q to the entire range, we conclude that

$$\sum_{\substack{z<p_1<\cdots<p_k\leq w}} \sum_{\substack{q\leq y \\ R_q=p_1\cdots p_k}} \alpha(q) \sum_{n\leq y/q} \alpha(n)\left| \sum_{Q_0<p\leq z} f_p(n)\right|^a$$

$$\leq \frac{2^{O(m-a)}\lambda_\alpha C_m}{k!(\sqrt{m})^{m-a}}(\log v)^k B(x)^{a/2}y^{\sigma_0}(\log y)^{\beta-1}.$$

Hence, (6.2) is bounded above by

$$\frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{a/2}y^{\sigma_0}(\log y)^{\beta-1}\sum_{k=1}^{b}\frac{(\log v)^k}{k!}\sum_{\substack{l_1+\cdots+l_k=b \\ l_1,\ldots,l_k\geq 1}}\binom{b}{l_1,\ldots,l_k}$$

$$=\frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{a/2}y^{\sigma_0}(\log y)^{\beta-1}\sum_{k=1}^{b}\begin{Bmatrix}b\\k\end{Bmatrix}(\log v)^k$$

$$\leq \frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{a/2}T_b(\log v)y^{\sigma_0}(\log y)^{\beta-1}.$$

It follows by (5.9) that the above does not exceed

$$\frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{a/2}(\log v)^b y^{\sigma_0}(\log y)^{\beta-1},$$

where we have used the observation that $\log v > m\log\log\log x > m\geq b$. In other words, we have shown that

$$\sum_{n\leq y}\alpha(n)\left|\sum_{Q_0<p\leq z}f_p(n)\right|^a\omega(n;z,w)^b$$

$$\leq \frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{a/2}(\log v)^b y^{\sigma_0}(\log y)^{\beta-1}.$$

Inserting this inequality into the definition of $E(y,z,w;m)$, we conclude that

$$(6.7)\quad E(y,z,w;m)$$

$$\leq \lambda_\alpha C_m y^{\sigma_0}(\log y)^{\beta-1}\sum_{a=0}^{m-1}\binom{m}{a}B(x)^{a/2}\left(O\left(\frac{\log v}{\sqrt{m}}\right)\right)^{m-a}$$

$$\ll \lambda_\alpha C_m\sqrt{m}\,(\log v)B(x)^{(m-1)/2}y^{\sigma_0}(\log y)^{\beta-1}$$

$$\ll \lambda_\alpha C_m m^{3/2}(\log\log\log x)B(x)^{(m-1)/2}y^{\sigma_0}(\log y)^{\beta-1}.$$

7. Deduction of Theorems 2.1 and 2.2. Theorem 2.1 now follows immediately upon combining (6.1) and (6.4) with (6.3) and (6.7) and invoking Lemma 4.1 and (3.4). In fact, we have shown that the same asymptotic

formulas which hold for $M(x; m)$ also hold for

$$(7.1) \qquad S(y)^{-1} \sum_{n \leq y} \alpha(n)(f(n) - A(x))^m$$

uniformly in the range $y \in [x^{\eta_0}, x]$, where $\eta_0 \in (0, 1]$ is any fixed constant.

Now we prove Theorem 2.2. Recall that under the hypotheses in Theorem 2.2, the multiplicative function $\alpha(n)$ satisfies conditions (2.1)–(2.4). We shall again suppose $A_0 \in (0, 1)$ throughout the proof. Define the strongly additive function $\tilde{f} \colon \mathbb{N} \to \mathbb{R}$, called the *strongly additive contraction* of f, by $\tilde{f}(p) = f(p)$ for all primes p. Then

$$(7.2) \qquad \sum_{n \leq x} \alpha(n)(f(n) - A(x))^m$$

$$= \sum_{k=0}^{m} \binom{m}{k} \sum_{n \leq x} \alpha(n)(\tilde{f}(n) - A(x))^k (f(n) - \tilde{f}(n))^{m-k}$$

for every $m \in \mathbb{N}$. The term corresponding to $k = m$ can be estimated directly using Theorem 2.1. Hence, it remains to deal with

$$(7.3) \qquad \sum_{n \leq x} \alpha(n)(\tilde{f}(n) - A(x))^k (f(n) - \tilde{f}(n))^l$$

for $0 \leq k < m$ and $l = m - k$.

Note that

$$\left| \sum_{n \leq x} \alpha(n)(\tilde{f}(n) - A(x))^k (f(n) - \tilde{f}(n))^l \right|$$

$$\leq \sum_{n \leq x} \alpha(n)|\tilde{f}(n) - A(x)|^k \left| \sum_{p^\nu \| n, \nu \geq 2} (f(p^\nu) - f(p)) \right|^l$$

$$\leq \sum_{\substack{p_1, \ldots, p_l \leq \sqrt{x}}} \sum_{\substack{p_1^{\nu_1}, \ldots, p_l^{\nu_l} \leq x \\ \nu_1, \ldots, \nu_l \geq 2}} |f(p_1^{\nu_1}) - f(p_1)| \cdots |f(p_l^{\nu_l}) - f(p_l)|$$

$$\times \sum_{\substack{n \leq x \\ p_1^{\nu_1}, \ldots, p_l^{\nu_l} \| n}} \alpha(n)|\tilde{f}(n) - A(x)|^k.$$

Since $f(p^\nu) = O(\nu^\kappa)$ for all p^ν, the last expression above does not exceed

$$2^{O(l)} \sum_{s \leq l} \sum_{p_1 < \cdots < p_s \leq \sqrt{x}} \sum_{\substack{l_1 + \cdots + l_s = l \\ l_1, \ldots, l_s \in \mathbb{N}}} \binom{l}{l_1, \ldots, l_s}$$

$$\times \sum_{\substack{p_1^{\nu_1} \cdots p_s^{\nu_s} \leq x \\ \nu_1, \ldots, \nu_s \geq 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \sum_{\substack{n \leq x \\ p_1^{\nu_1}, \ldots, p_s^{\nu_s} \| n}} \alpha(n)|\tilde{f}(n) - A(x)|^k.$$

If we write $n = p_1^{\nu_1} \cdots p_s^{\nu_s} n'$ with $\gcd(n', p_1 \cdots p_s) = 1$, then it is clear that

$$|\tilde{f}(n) - A(x)|^k = \left| \tilde{f}(n') - A(x) + \sum_{i=1}^{s} f(p_i) \right|^k$$

$$\leq \sum_{a=0}^{k} \binom{k}{a} |\tilde{f}(n') - A(x)|^a \left| \sum_{i=1}^{s} f(p_i) \right|^{k-a}.$$

Thus, the innermost sum of $\alpha(n)|\tilde{f}(n) - A(x)|^k$ is not greater than

$$(7.4) \quad \alpha(p_1^{\nu_1}) \cdots \alpha(p_s^{\nu_s}) \sum_{a=0}^{k} \binom{k}{a} \left| \sum_{i=1}^{s} f(p_i) \right|^{k-a} \sum_{n \leq x/(p_1^{\nu_1} \cdots p_s^{\nu_s})} \alpha(n)|\tilde{f}(n) - A(x)|^a,$$

where we have dropped the superscript of n for simplicity of notation. Since the right-hand side of the above clearly vanishes if $p_1 \cdots p_s > \sqrt{x}$, we may assume $p_1 \cdots p_s \leq \sqrt{x}$ instead. Let $\lambda' := 1 - \varrho_0 - \log_2 \lambda > \rho_0$, and choose a constant $\max(1/2, \sqrt{\varrho_0/\lambda'}) < \delta_0 < 1$ such that $1 - \varrho_0 + \delta_0^2 \lambda' > 1$. Let $x_s := x/(p_1 \cdots p_s)$ and $y_s := x_s^{\delta_0}$. Then $x_s \geq \sqrt{x} \geq p_1 \cdots p_s$. If $p_1^{\nu_1} \cdots p_s^{\nu_s} > p_1 \cdots p_s y_s$ with given $p_1 < \cdots < p_s$, then we use the trivial estimate

$$\sum_{n \leq x/(p_1^{\nu_1} \cdots p_s^{\nu_s})} \alpha(n)|\tilde{f}(n) - A(x)|^a$$

$$\ll 2^{O(a)} (\log x)^a \sum_{n \leq x/(p_1^{\nu_1} \cdots p_s^{\nu_s})} \alpha(n)$$

$$\ll \lambda_\alpha 2^{O(a)} (\log x)^a \left(\frac{x}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\sigma_0} \left(\log \frac{3x}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\beta-1}.$$

Thus, (7.4) is

$$\ll \frac{\alpha(p_1^{\nu_1}) \cdots \alpha(p_s^{\nu_s})}{p_1^{\sigma_0 \nu_1} \cdots p_s^{\sigma_0 \nu_s}} \left(\log \frac{3x}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\beta-1} \lambda_\alpha 2^{O(k)} x^{\sigma_0} (\log x)^k.$$

Since $\alpha(p^\nu) = O((\lambda p^{\varrho_0 + \sigma_0 - 1})^\nu)$ for all p^ν, we have

$$\sum_{\substack{p_1 \cdots p_s y_s < p_1^{\nu_1} \cdots p_s^{\nu_s} \leq x \\ \nu_1, \ldots, \nu_s \geq 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \frac{\alpha(p_1^{\nu_1}) \cdots \alpha(p_s^{\nu_s})}{p_1^{\sigma_0 \nu_1} \cdots p_s^{\sigma_0 \nu_s}} \left(\log \frac{3x}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\beta-1}$$

$$\leq 2^{O(l)} \sum_{\substack{p_1 \cdots p_s y_s < p_1^{\nu_1} \cdots p_s^{\nu_s} \leq x \\ \nu_1, \ldots, \nu_s \geq 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s}$$

$$\times \left(\frac{\lambda}{p_1^{1-\varrho_0}} \right)^{\nu_1} \cdots \left(\frac{\lambda}{p_s^{1-\varrho_0}} \right)^{\nu_s} \left(\log \frac{3x}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\beta-1}$$

$$\leq \frac{2^{O(l)}}{(p_1 \cdots p_s)^{1-\varrho_0}} \sum_{\substack{y_s < p_1^{\nu_1} \cdots p_s^{\nu_s} \leq x_s \\ \nu_1, \ldots, \nu_s \geq 1}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s}$$

$$\times \left(\frac{\lambda}{p_1^{1-\varrho_0}} \right)^{\nu_1} \cdots \left(\frac{\lambda}{p_s^{1-\varrho_0}} \right)^{\nu_s} \left(\log \frac{3x_s}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\beta-1}.$$

It is not hard to see that the proof of (3.3) also gives

$$\sum_{z_1 < p^{\nu} \leq z_2} \left(\frac{\lambda}{p^{1-\varrho_0}} \right)^{\nu} \left(\log \frac{3z_2}{p^{\nu}} \right)^{\beta-1} \ll \frac{(\log(3z_2/z_1))^{\beta-1}}{z_1^{1-\varrho_0 - \log_p \lambda}}$$

uniformly for all primes p and all $0 < z_1 \leq z_2$. Thus, we have

$$\sum_{\substack{y_s < p_1^{\nu_1} \cdots p_s^{\nu_s} \leq x_s \\ \nu_1, \ldots, \nu_s \geq 1}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \left(\frac{\lambda}{p_1^{1-\varrho_0}} \right)^{\nu_1} \cdots \left(\frac{\lambda}{p_s^{1-\varrho_0}} \right)^{\nu_s} \left(\log \frac{3x_s}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\beta-1}$$

$$\leq 2^{O(l)} (\log x)^{\kappa l} \sum_{\substack{y_s < p_1^{\nu_1} \cdots p_s^{\nu_s} \leq x_s \\ \nu_1, \ldots, \nu_s \geq 1}} \left(\frac{\lambda}{p_1^{1-\varrho_0}} \right)^{\nu_1} \cdots \left(\frac{\lambda}{p_s^{1-\varrho_0}} \right)^{\nu_s} \left(\log \frac{3x_s}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\beta-1}$$

$$\leq 2^{O(l)} (\log x)^{\kappa l} \sum_{\substack{p_2^{\nu_2} \cdots p_s^{\nu_s} \leq x_s \\ \nu_2, \ldots, \nu_s \geq 1}} \left(\frac{\lambda}{p_2^{\log_{p_1} \lambda}} \right)^{\nu_2} \cdots \left(\frac{\lambda}{p_s^{\log_{p_1} \lambda}} \right)^{\nu_s} \frac{(\log(3x_s/y_s))^{\beta-1}}{y_s^{1-\varrho_0 - \log_{p_1} \lambda}}$$

$$\leq \frac{2^{O(l)} (\log x)^{(\kappa+1)m + \beta - 2}}{x^{\delta_0(1-\delta_0)\lambda'/2} (p_1 \cdots p_s)^{\delta_0^2 \lambda'}} \leq \frac{2^{O(l)} (\log x)^{\beta-1}}{x^{(1-\delta_0)\lambda'/5} (p_1 \cdots p_s)^{\delta_0^2 \lambda'}},$$

where the penultimate inequality follows from the previous line together with the observations that $p_i^{\log_{p_1} \lambda} > \lambda$ for all $2 \leq i \leq s$, that $x^{(1+\delta_0)/2} \geq (p_1 \cdots p_s)^{1+\delta_0}$, and that

$$y_s^{1-\varrho_0 - \log_{p_1} \lambda} \geq y_s^{\lambda'} = \left(x^{(1-\delta_0)/2} \cdot \frac{x^{(1+\delta_0)/2}}{p_1 \cdots p_s} \right)^{\delta_0 \lambda'} \geq x^{\delta_0(1-\delta_0)\lambda'/2} (p_1 \cdots p_s)^{\delta_0^2 \lambda'}.$$

It follows that

$$\sum_{\substack{p_1 \cdots p_s y_s < p_1^{\nu_1} \cdots p_s^{\nu_s} \leq x \\ \nu_1, \ldots, \nu_s \geq 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \frac{\alpha(p_1^{\nu_1}) \cdots \alpha(p_s^{\nu_s})}{p_1^{\sigma_0 \nu_1} \cdots p_s^{\sigma_0 \nu_s}} \left(\log \frac{3x}{p_1^{\nu_1} \cdots p_s^{\nu_s}} \right)^{\beta-1}$$

$$\leq \frac{2^{O(l)}}{(p_1 \cdots p_s)^{1-\varrho_0 + \delta_0^2 \lambda'}} x^{-(1-\delta_0)\lambda'/5} (\log x)^{\beta-1},$$

from which we deduce that

$$(7.5) \quad \sum_{\substack{p_1<\cdots<p_s\le\sqrt{x}}} \sum_{\substack{p_1\cdots p_s y_s < p_1^{\nu_1}\cdots p_s^{\nu_s}\le x \\ \nu_1,\ldots,\nu_s\ge 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s} \sum_{\substack{n\le x \\ p_1^{\nu_1},\ldots,p_s^{\nu_s}\|n}} \alpha(n)|\tilde{f}(n)-A(x)|^k$$

$$\le 2^{O(m)}\lambda_\alpha x^{\sigma_0-(1-\delta_0)\lambda'/5}(\log x)^{k+\beta-1} \sum_{p_1<\cdots<p_s\le\sqrt{x}} \frac{1}{(p_1\cdots p_s)^{1-\varrho_0+\delta_0^2\lambda'}}$$

$$\le \frac{1}{s!}\lambda_\alpha x^{\sigma_0-(1-\delta_0)\lambda'/6}(\log x)^{\beta-1}.$$

On the other hand, if $p_1^{\nu_1}\cdots p_s^{\nu_s}\le p_1\cdots p_s y_s$, then $x^{(1-\delta_0)/2}\le x/(p_1^{\nu_1}\cdots p_s^{\nu_s})$ $\le x$. Therefore, we can apply the asymptotic formulas for (7.1) with $\eta_0=(1-\delta_0)/2$ and $y=x/(p_1^{\nu_1}\cdots p_s^{\nu_s})$, in conjunction with the Cauchy–Schwarz inequality, to estimate the inner sum in (7.4). As a consequence,

$$\sum_{n\le x/(p_1^{\nu_1}\cdots p_s^{\nu_s})} \alpha(n)|\tilde{f}(n)-A(x)|^a$$

$$\ll \frac{2^{O(m-a)}\lambda_\alpha C_m}{(\sqrt{m})^{m-a}}B(x)^{a/2}\left(\frac{x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\sigma_0}\left(\log\frac{x}{p_1^{\nu_1}\cdots p_s^{\nu_s}}\right)^{\beta-1}.$$

Inserting this into (7.4) shows that

$$\sum_{\substack{n\le x \\ p_1^{\nu_1},\ldots,p_s^{\nu_s}\|n}} \alpha(n)|\tilde{f}(n)-A(x)|^k$$

$$\le \frac{2^{O(m-k)}\lambda_\alpha C_m}{(\sqrt{m})^{m-k}}\cdot\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}$$

$$\times\left(\sqrt{B(x)}+O\left(\frac{1}{\sqrt{m}}\sum_{i=1}^s|f(p_i)|\right)\right)^k x^{\sigma_0}(\log x)^{\beta-1}$$

$$= \frac{2^{O(l)}\lambda_\alpha C_m}{m^{l/2}}\cdot\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}B(x)^{k/2}\left(1+O\left(\sqrt{\frac{m}{B(x)}}\right)\right)x^{\sigma_0}(\log x)^{\beta-1}$$

$$\le \frac{2^{O(l)}\lambda_\alpha C_m}{m^{l/2}}\cdot\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}B(x)^{k/2}x^{\sigma_0}(\log x)^{\beta-1}.$$

Note that

$$\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\le p_1\cdots p_s y_s \\ \nu_1,\ldots,\nu_s\ge 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\frac{\alpha(p_1^{\nu_1})\cdots\alpha(p_s^{\nu_s})}{p_1^{\sigma_0\nu_1}\cdots p_s^{\sigma_0\nu_s}}$$

$$\le 2^{O(l)}\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\le p_1\cdots p_s y_s \\ \nu_1,\ldots,\nu_s\ge 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s}\left(\frac{\lambda}{p_1^{1-\varrho_0}}\right)^{\nu_1}\cdots\left(\frac{\lambda}{p_s^{1-\varrho_0}}\right)^{\nu_s}$$

$$\leq \frac{2^{O(l)}}{(p_1 \cdots p_s)^{1-\varrho_0}} \sum_{\substack{p_1^{\nu_1} \cdots p_s^{\nu_s} \leq y_s \\ \nu_1, \ldots, \nu_s \geq 1}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \left(\frac{\lambda}{p_1^{1-\varrho_0}} \right)^{\nu_1} \cdots \left(\frac{\lambda}{p_s^{1-\varrho_0}} \right)^{\nu_s}$$

$$\leq \frac{2^{O(l)}}{(p_1 \cdots p_s)^{1-\varrho_0}} \prod_{i=1}^{s} \mathrm{Li}_{-\lceil \kappa l_i \rceil}(\lambda/p_i^{1-\varrho_0}),$$

where

$$\mathrm{Li}_{-\ell}(\zeta) := \sum_{n=1}^{\infty} n^{\ell} \zeta^n$$

is the polylogarithm function of order $-\ell$ and complex argument ζ with $|\zeta| < 1$, where $\ell \geq 0$ is any integer. For example, $\mathrm{Li}_0(\zeta) = \zeta/(1-\zeta)$ and $\mathrm{Li}_{-1}(\zeta) = \zeta/(1-\zeta)^2$. The function $\mathrm{Li}_{-\ell}(\zeta)$ can be expressed in terms of the Eulerian polynomial $A_\ell(\zeta)$:

$$\mathrm{Li}_{-\ell}(\zeta) = \frac{\zeta A_\ell(\zeta)}{(1-\zeta)^{\ell+1}},$$

where

$$A_\ell(\zeta) := \sum_{j=0}^{\ell} \left\langle \begin{matrix} \ell \\ j \end{matrix} \right\rangle \zeta^j$$

is the ℓth Eulerian polynomial, and

$$\left\langle \begin{matrix} \ell \\ j \end{matrix} \right\rangle := \sum_{a=0}^{j} (-1)^a \binom{\ell+1}{a} (j+1-a)^\ell$$

is the jth Eulerian number of size ℓ. Combinatorially, it is known that, for every $\ell \geq 1$,

$$\left\langle \begin{matrix} \ell \\ j \end{matrix} \right\rangle = \#\{\tau \in S_\ell \colon \tau \text{ has exactly } j \text{ ascents}\},$$

where S_ℓ is the set of all permutations of $\{1, \ldots, \ell\}$. Using this combinatorial intepretation one finds that $A_\ell(1) = \#S_\ell = \ell!$. Since $l_1 + \cdots + l_s = l \leq m$, we have

$$\prod_{i=1}^{s} \mathrm{Li}_{-\lceil \kappa l_i \rceil}(\lambda/p_i^{1-\varrho_0}) \leq \frac{2^{O(l)} \lceil \kappa l_1 \rceil! \cdots \lceil \kappa l_s \rceil!}{(p_1 \cdots p_s)^{1-\varrho_0}} = \frac{2^{O(l)}(l_1^{l_1} \cdots l_s^{l_s})^\kappa}{(p_1 \cdots p_s)^{1-\varrho_0}}$$

$$\leq \frac{2^{O(l)} m^{\kappa l}}{(p_1 \cdots p_s)^{1-\varrho_0}},$$

by Stirling's formula. Hence,

$$\sum_{\substack{p_1^{\nu_1} \cdots p_s^{\nu_s} \leq p_1 \cdots p_s y_s \\ \nu_1, \ldots, \nu_s \geq 2}} \nu_1^{\kappa l_1} \cdots \nu_s^{\kappa l_s} \frac{\alpha(p_1^{\nu_1}) \cdots \alpha(p_s^{\nu_s})}{p_1^{\sigma_0 \nu_1} \cdots p_s^{\sigma_0 \nu_s}} \leq \frac{2^{O(l)} m^{\kappa l}}{(p_1 \cdots p_s)^{2(1-\varrho_0)}}.$$

It follows that

$$
\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\leq p_1\cdots p_s y_s \\ \nu_1,\ldots,\nu_s\geq 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s} \sum_{\substack{n\leq x \\ p_1^{\nu_1},\ldots,p_s^{\nu_s}\|n}} \alpha(n)|\tilde{f}(n)-A(x)|^k
$$
$$
\leq \frac{2^{O(l)}\lambda_\alpha C_m m^{\kappa l}}{m^{l/2}(p_1\cdots p_s)^{2(1-\varrho_0)}} B(x)^{k/2}x^{\sigma_0}(\log x)^{\beta-1}.
$$

Summing the above over $p_1<\cdots<p_s\leq\sqrt{x}$, we arrive at

$$
\sum_{p_1<\cdots<p_s\leq\sqrt{x}}\;\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\leq p_1\cdots p_s y_s \\ \nu_1,\ldots,\nu_s\geq 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s} \sum_{\substack{n\leq x \\ p_1^{\nu_1},\ldots,p_s^{\nu_s}\|n}} \alpha(n)|\tilde{f}(n)-A(x)|^k
$$
$$
\leq 2^{O(l)}\lambda_\alpha C_m m^{(\kappa-1/2)l}B(x)^{k/2}x^{\sigma_0}(\log x)^{\beta-1}\sum_{p_1<\cdots<p_s\leq\sqrt{x}}\frac{1}{(p_1\cdots p_s)^{2(1-\varrho_0)}}
$$
$$
\leq \frac{2^{O(l)}}{s!}\lambda_\alpha C_m m^{(\kappa-1/2)l}B(x)^{k/2}x^{\sigma_0}(\log x)^{\beta-1},
$$

since $\varrho_0\in[0,1/2)$. Combining this estimate with (7.5), we obtain

$$
\sum_{p_1<\cdots<p_s\leq\sqrt{x}}\;\sum_{\substack{p_1^{\nu_1}\cdots p_s^{\nu_s}\leq x \\ \nu_1,\ldots,\nu_s\geq 2}} \nu_1^{\kappa l_1}\cdots\nu_s^{\kappa l_s} \sum_{\substack{n\leq x \\ p_1^{\nu_1},\ldots,p_s^{\nu_s}\|n}} \alpha(n)|\tilde{f}(n)-A(x)|^k
$$
$$
\leq \frac{2^{O(l)}}{s!}\lambda_\alpha C_m m^{(\kappa-1/2)l}B(x)^{k/2}x^{\sigma_0}(\log x)^{\beta-1}.
$$

Therefore, (7.3) is bounded above by

$$
2^{O(l)}\lambda_\alpha C_m m^{(\kappa-1/2)l}B(x)^{k/2}x^{\sigma_0}(\log x)^{\beta-1}\sum_{s\leq l}\frac{1}{s!}\sum_{\substack{l_1+\cdots+l_s=l \\ l_1,\ldots,l_s\in\mathbb{N}}}\binom{l}{l_1,\ldots,l_s}
$$
$$
\leq 2^{O(l)}\lambda_\alpha C_m m^{(\kappa-1/2)l}B(x)^{k/2}x^{\sigma_0}(\log x)^{\beta-1}T_l(1)
$$
$$
\leq 2^{O(l)}C_m m^{(\kappa+1/2)l}B(x)^{k/2}S(x),
$$

which allows us to conclude that

$$
\sum_{k=0}^{m-1}\binom{m}{k}\sum_{n\leq x}\alpha(n)(\tilde{f}(n)-A(x))^k(f(n)-\tilde{f}(n))^{m-k}
$$
$$
\ll C_m m^{\kappa+3/2}B(x)^{(m-1)/2}S(x),
$$

provided that in addition $1\leq m\ll B(x)^{1/(2\kappa+3)}$. Inserting the above estimate and the estimate for the term corresponding to $k=m$ into (7.2) completes the proof of Theorem 2.2.

8. Proof of Theorem 2.3 (sketch). Now we outline the proof of Theorem 2.3. The first step is to redefine $f_q(n)$ introduced in Section 4. Again, let us suppose that $A_0 \in (0,1)$ and that $|f(p)| \leq 1$ for all primes p. For every $q \in \mathbb{N}$ we define

$$\overline{F}(\sigma_0, q) := \prod_{p \mid q}(1 - F(\sigma_0, p)), \quad \widetilde{F}(\sigma_0, q) := \frac{\rho_g(q)}{\varphi(q)}\overline{F}(\sigma_0, q).$$

For each prime p we put

$$f_p(n) := \begin{cases} f(p)(1 - \widetilde{F}(\sigma_0, p)) & \text{if } p \mid n, \\ -f(p)\widetilde{F}(\sigma_0, p) & \text{otherwise,} \end{cases}$$

and as before we set

$$f_q(n) := \prod_{p^\nu \| q} f_p(n)^\nu$$

for any $q \in \mathbb{N}$. In addition, let $c_g \in \mathbb{N}$ be the least positive integer such that $c_g g(x) \in \mathbb{Z}[x]$, and let $Q_0 > c_g|g(0)| \geq 1$ be such that (3.17) holds. Then for each $q \in \mathbb{N}$ with $P^-(q) > Q_0$ we have $\mathcal{Z}_g(q) = \mathcal{Z}_g^*(q) \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$ and $\rho_g(q) = \#\mathcal{Z}_g^*(q)$, where $\mathcal{Z}_g(q)$ denotes the zero locus of g in $\mathbb{Z}/q\mathbb{Z}$. In particular, we have $0 \leq \rho_g(q) \leq \varphi(q)$, which implies that $0 \leq \widetilde{F}(\sigma_0, q) \leq 1$ and $|f_q(n)| \leq 1$ for all $n \in \mathbb{N}$.

Next, we need an analogue of Lemma 4.1. Let x be sufficiently large and set $z := x^{\delta(x)/m} > Q_0$. Then

$$\sum_{Q_0 < p \leq x} f(p)\widetilde{F}(\sigma_0, p) = A_{f,g}(x) + O(1)$$

by (2.1), (3.17), and the facts that ρ_g is bounded on prime powers and that $\sum_p \psi_0(p) < \infty$. It is easily seen that

$$f(g(n)) - A_{f,g}(x) = \sum_{\substack{Q_0 < p \leq z \\ p \mid g(n)}} f_p(g(n)) + \sum_{p > z} f(p) - \sum_{z < p \leq x} f(p)\widetilde{F}(\sigma_0, p) + O(1).$$

Note that

$$\sum_{z < p \leq x} f(p)\widetilde{F}(\sigma_0, p) = A_{f,g}(x) - A_{f,g}(z) + O(1) \ll \log\left(\frac{m}{\delta(x)} + 1\right).$$

Since $1 \leq g(n) \ll n^{d_g}$ uniformly for all $n \in \mathbb{N}$, where $d_g := \deg g \geq 1$, we have

$$\sum_{\substack{p > z \\ p \mid g(n)}} f(p) \ll \frac{m}{\delta(x)}.$$

It follows that
$$\sum_{n\le x}\alpha(n)(f(g(n))-A_{f,g}(x))^m = \sum_{n\le x}\alpha(n)\Big(\sum_{Q_0<p\le z}f_p(g(n))\Big)^m+O(E_g(x;m)),$$
where
$$E_g(x;m):=\sum_{k=0}^{m-1}\binom{m}{k}2^{O(m-k)}(m\delta(x)^{-1})^{m-k}\sum_{n\le x}\alpha(n)\Big|\sum_{p\le z}f_p(g(n))\Big|^k.$$

Now we turn to the estimation of
$$\sum_{n\le x}\alpha(n)\Big(\sum_{Q_0<p\le z}f_p(g(n))\Big)^m = \sum_{Q_0<p_1,\dots,p_m\le z}\sum_{n\le x}\alpha(n)f_{p_1\cdots p_m}(g(n)).$$

Let $q\in\mathbb{N}\cap[1,x^{\delta(x)}]$ with $\omega(q)\le m$, $P^-(q)>Q_0$ and $P^+(q)\le z$. Then
$$\sum_{n\le x}\alpha(n)f_q(g(n)) = \sum_{ab|R_q}f_q(a)\mu(b)\sum_{\substack{n\le x\\ab|g(n)}}\alpha(n)$$
$$= \sum_{ab|R_q}f_q(a)\mu(b)\sum_{c\in\mathcal{Z}_g^*(ab)}\sum_{\substack{n\le x\\n\equiv c\,(\mathrm{mod}\,ab)}}\alpha(n).$$

Thus in place of Lemma 3.3, we need to input in our analysis the information about the distribution of values of $\alpha(n)$ with n restricted to reduced residue classes. By Lemma 3.2, the innermost sum differs from
$$\frac{1}{\varphi(ab)}\sum_{\substack{n\le x\\\gcd(n,ab)=1}}\alpha(n) = \frac{1}{\varphi(ab)}x^{\sigma_0}(\log x)^{\beta-1}\Big(\lambda_\alpha(ab)+O\Big(\frac{1}{(\log x)^{A_0}}\Big)\Big)$$
$$= \frac{1}{\varphi(ab)}x^{\sigma_0}(\log x)^{\beta-1}\Big(\lambda_\alpha\overline{F}(\sigma_0,ab)+O\Big(\frac{1}{(\log x)^{A_0}}\Big)\Big)$$
$$= \frac{1}{\varphi(ab)}\lambda_\alpha x^{\sigma_0}(\log x)^{\beta-1}\Big(\overline{F}(\sigma_0,ab)+O\Big(\frac{1}{(\log x)^{A_0}}\Big)\Big)$$
by the amount $\Delta_\alpha(x;ab,c)$. Hence,
$$\sum_{n\le x}\alpha(n)f_q(g(n)) = \lambda_\alpha\Big(\widetilde{G}_1(\sigma_0,q)+O\Big(\frac{\widetilde{G}_2(\sigma_0,q)}{(\log x)^{A_0}}\Big)\Big)x^{\sigma_0}(\log x)^{\beta-1}+J_q(x),$$
where
$$\widetilde{G}_1(\sigma_0,q):=\sum_{ab|R_q}f_q(a)\mu(b)\widetilde{F}(\sigma_0.ab),$$
$$\widetilde{G}_2(\sigma_0,q):=\sum_{ab|R_q}\frac{\rho_g(ab)}{\varphi(ab)}|f_q(a)|,$$
$$J_q(x):=\sum_{ab|R_q}f_q(a)\mu(b)\sum_{c\in\mathcal{Z}_g^*(ab)}\Delta_\alpha(x;ab,c).$$

It is clear that $\widetilde{G}_1$ and $\widetilde{G}_2$ are both multiplicative in q. Easy calculation shows that

$$\widetilde{G}_1(\sigma_0, p^\nu)$$
$$= f(p)^\nu \widetilde{F}(\sigma_0, p)(1 - \widetilde{F}(\sigma_0, p))\big((-1)^\nu \widetilde{F}(\sigma_0, p)^{\nu-1} + (1 - \widetilde{F}(\sigma_0, p))^{\nu-1}\big)$$

for any prime power p^ν. In particular, $\widetilde{G}_1(\sigma_0, p) = 0$, $|\widetilde{G}_1(\sigma_0, p^\nu)| \leq 1/4$, and $\widetilde{G}_1(\sigma_0, p^\nu) \geq 0$ when $2 \mid \nu$. Moreover,

$$\widetilde{G}_1(\sigma_0, p^2) = f(p)^2 \widetilde{F}(\sigma_0, p)(1 - \widetilde{F}(\sigma_0, p)) = \rho_g(p)\frac{f(p)^2}{p} + O\left(\frac{F(\sigma_0, p)}{p} + \frac{1}{p^2}\right),$$

and

$$|\widetilde{G}_1(\sigma_0, p^\nu)| \leq |f(p)|^\nu \widetilde{F}(\sigma_0, p) \leq \rho_g(p)\frac{f(p)^2}{\varphi(p)} = \rho_g(p)\frac{f(p)^2}{p} + O\left(\frac{1}{p^2}\right)$$

for all p^ν with $p > Q_0$ and $\nu \geq 2$. In addition,

$$\left|\sum_{Q_0 < p_1, \ldots, p_m \leq z} J_{p_1 \cdots p_m}(x)\right|$$

$$\leq \sum_{\substack{q:\,\omega(q) \leq m \\ P^+(q) \leq z}} \mu(q)^2 \sum_{c \in \mathcal{Z}_g^*(q)} |\Delta_\alpha(x; q, c)|$$

$$\times \sum_{a \mid q} \sum_{s \leq m} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} \sum_{\substack{Q_0 < p_1 < \cdots < p_s \leq z \\ q \mid p_1 \cdots p_s}} |f_{p_1^{k_1} \cdots p_s^{k_s}}(a)|.$$

The inner sum over s is

$$\leq \sum_{s=\omega(q)}^{m} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} \frac{1}{(s - \omega(q))!} \Big(\sum_{\substack{Q_0 < p \leq z \\ p \nmid q}} |f_p(a)|\Big)^{s - \omega(q)}$$

$$\leq 2^{O(m)} \sum_{s=\omega(q)}^{m} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} \frac{1}{(s - \omega(q))!} (\log \log x)^{s - \omega(q)}$$

$$\leq 2^{O(m)} T_m(\log \log x) \leq 2^{O(m)} (\log \log x)^m.$$

It follows by (2.5) that

$$\left|\sum_{Q_0 < p_1, \ldots, p_m \leq z} J_{p_1 \cdots p_m}(x)\right|$$

$$\leq 2^{O(m)} (\log \log x)^m \sum_{\substack{q:\,\omega(q) \leq m \\ P^+(q) \leq z}} \mu(q)^2 \sum_{c \in \mathcal{Z}_g^*(q)} |\Delta_\alpha(x; q, c)|$$

$$\leq 2^{O(m)} (\log \log x)^m S(x) \exp(-(\log \log x)^{1/3 + \epsilon_0}),$$

which is $o(S(x))$. These observations allow us to conclude the proof of Theorem 2.3 by arguing as in Sections 5 and 6. It is also clear from the last inequality above that the bound $S(x)\exp(-(\log\log x)^{1/3+\epsilon_0})$ in (2.5) can be weakened to a complicated one involving $\delta(x)$ and $B_{f,g}(x)$.

9. Proofs of Theorem 2.4 and Corollary 2.5 (sketch). Now we outline the proof of Theorem 2.4, which borrows the ideas from the proofs of Theorem 2.1 and [8, Theorem 1] with proper modifications. Let $0 < \epsilon < \min(1, K)$, and take $z := x^{1/v}$ and

$$w := \begin{cases} x^{1/\log(v+2)} & \text{if } \beta = 1, \\ x^{1/(\epsilon \log(v+2))} & \text{if } \beta \neq 1, \end{cases}$$

where we recall that $v \asymp m$ when $\beta = 1$, and $v = (\log\log x)^{m(\vartheta_0+2)}$ when $\beta \neq 1$ as chosen in Section 6. Having made these choices, we deduce that $\epsilon \log(v+2) \to \infty$ as $x \to \infty$ in the case $\beta \neq 1$. Let

$$\mathcal{P}_\epsilon^-(x) := \big\{p \leq x \colon |f(p)| \leq \epsilon\sqrt{B^*(x)}\big\},$$
$$\mathcal{P}_\epsilon^+(x) := \big\{p \leq x \colon \epsilon\sqrt{B^*(x)} < |f(p)| \leq K\sqrt{B^*(x)}\big\},$$
$$\mathcal{P}_\infty(x) := \big\{p \leq x \colon |f(p)| > K\sqrt{B^*(x)}\big\},$$

and put $\mathcal{P}_K(x) := \mathcal{P}_\epsilon^-(x) \cup \mathcal{P}_\epsilon^+(x)$. We consider the strongly additive function

$$f_\epsilon(n;x) := \sum_{\substack{p|n \\ p\in\mathcal{P}_\epsilon^-(x)}} f(p) + \epsilon_{\beta,1} \sum_{\substack{p|n \\ p\in\mathcal{P}_\epsilon^+(x)\cap(z,x]}} f(p) + \sum_{\substack{p|n \\ p\in\mathcal{P}_\infty(x)}} f(p),$$

where we recall that $\epsilon_{\beta,1}$ takes value 0 if $\beta = 1$ and 1 otherwise, and define

$$A_\epsilon(x) := \sum_{p\in\mathcal{P}_\epsilon^-(x)} \alpha(p)\frac{f(p)}{p^{\sigma_0}}, \quad B_\epsilon(x) := \sum_{p\in\mathcal{P}_\epsilon^-(x)} \alpha(p)\frac{f(p)^2}{p^{\sigma_0}}.$$

By hypothesis,

$$B(x) - B_\epsilon(x) = \sum_{\substack{p\leq x \\ |f(p)|>\epsilon\sqrt{B^*(x)}}} \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} = o(B^*(x)),$$

and so

$$|A_\epsilon(x) - A(x)| \leq \frac{1}{\epsilon\sqrt{B^*(x)}} \sum_{\substack{p\leq x \\ |f(p)|>\epsilon\sqrt{B^*(x)}}} \alpha(p)\frac{f(p)^2}{p^{\sigma_0}} = o(\epsilon^{-1}\sqrt{B^*(x)}).$$

We expect that the distribution of $f_\epsilon(n;x)$ is close to being Gaussian with mean $A(x)$ and variance $B(x)$ when x gets sufficiently large. In what follows, we shall restrict our attention to the case $\beta \neq 1$, since the opposite

case is similar and easier. Looking back at the proof of Lemma 4.1, we find, for sufficiently large x, that

$$\sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (Q_0, x]} f(p) F(\sigma_0, p) = A_\epsilon(x) + O(\epsilon \sqrt{B^*(x)}) = A(x) + O(\epsilon \sqrt{B^*(x)}),$$

so that

$$(9.1) \quad f_\epsilon(n; x) - A(x)$$
$$= \sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (Q_0, z]} f_p(n) + \sum_{\substack{p | n \\ p \in \mathcal{P}_K(x) \cap (z, w]}} f(p) + O(\epsilon \sqrt{B(x)}),$$

where we have used the hypothesis that $f(n) = o(\sqrt{B(x)})$ for all $n \le x$ whose prime factors p satisfy $|f(p)| > K\sqrt{B^*(x)}$. This leads to an analogue of Lemma 4.1 in which the second sum above plays the same role as $\omega(n; z, w)$. To estimate the moments of $f_\epsilon(n; x)$, one only needs to recycle the arguments used in the proof of Theorem 2.1 and make suitable modifications. For instance, the estimation of

$$\sum_{n \le y} \alpha(n) \Big(\sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (Q_0, z]} f_p(n) \Big)^m$$

is essentially the same as that of (4.1) given in Sections 4 and 5, except that we use the inequality $|f(p)| \le \epsilon \sqrt{B^*(x)}$ for $p \in \mathcal{P}_\epsilon^-(x)$ in place of the bound $f(p) = O(1)$ throughout the argument. This way, we obtain

$$(9.2) \quad \sum_{n \le y} \alpha(n) \Big(\sum_{p \in \mathcal{P}_\epsilon^-(x) \cap (Q_0, z]} f_p(n) \Big)^m$$
$$= \lambda_\alpha \Big(\mu_m + O\Big(\frac{\epsilon \log v}{\log \log \log x} \Big) \Big) B(x)^{m/2} y^{\sigma_0} (\log y)^{\beta - 1}$$
$$= \lambda_\alpha (\mu_m + O(\epsilon)) B(x)^{m/2} y^{\sigma_0} (\log y)^{\beta - 1}$$

uniformly for $y \in [x^{\eta_0}, x]$, where $\eta_0 \in (0, 1]$ is any given constant. On the other hand, the estimation of the error involving the second sum in (9.1) is essentially the same as that of $E(y, z, w; m)$ in the case $\beta \ne 1$ given in Section 6. The only difference is that we now make use of the estimates $|f(p)| \le K\sqrt{B^*(x)}$ for all $p \in \mathcal{P}_K(x)$ and

$$\sum_{p \in \mathcal{P}_K(x) \cap (z, w]} \alpha(p) \frac{|f(p)|^\nu}{p^{\sigma_0}} \ll B^*(x)^{(\nu - 1)/2} \sum_{p \in \mathcal{P}_K(x) \cap (z, w]} \alpha(p) \frac{|f(p)|}{p^{\sigma_0}}$$
$$\ll \epsilon B^*(x)^{\nu/2} \log v \ll \epsilon B(x)^{\nu/2}$$

for all $\nu \ge 1$, which can be easily seen by considering $p \in \mathcal{P}_\epsilon^-(x)$ and

$p \in \mathcal{P}_\epsilon^+(x)$ separately, in place of the estimates $f(p) = O(1)$ and

$$\sum_{z < p \leq w} \alpha(p) \frac{|f(p)|^\nu}{p^{\sigma_0}} = O(\log v) = O(\log \log \log x),$$

respectively. One shows in this way that the error involving the second sum in (9.1) is $O(\epsilon \lambda_\alpha B(x)^{m/2} y^{\sigma_0} (\log y)^{\beta-1})$. Combining this estimate with (9.2) and taking $y = x$ yields

$$S(x)^{-1} \sum_{n \leq x} \alpha(n)(f_\epsilon(n; x) - A(x))^m = (\mu_m + O(\epsilon)) B(x)^{m/2}$$

for every fixed $m \in \mathbb{N}$ and all sufficiently large x, where the implied constant is independent of ϵ.

To complete the proof of Theorem 2.4 for $\beta \neq 1$, it is sufficient to show

$$(9.3) \qquad S(x)^{-1} \sum_{n \leq x} \alpha(n) |f(n) - f_\epsilon(n; x)|^m = O(\epsilon B^*(x)^{m/2})$$

for every given $\epsilon \in (0, 1)$ and $m \in \mathbb{N}$, where the implicit constant is independent of ϵ. Since the case where m is odd follows from the case where m is even by Cauchy–Schwarz, we only need to consider the latter. The proof of this case is largely the same as that of [8, Lemma 2], except for the slight complication for $\beta \in (0, 1)$. When m is even, we have

$$S(x)^{-1} \sum_{n \leq x} \alpha(n) |f(n) - f_\epsilon(n; x)|^m$$

$$= S(x)^{-1} \sum_{n \leq x} \alpha(n) \sum_{\substack{p_1, \ldots, p_m | n \\ p_1, \ldots, p_m \in \mathcal{P}_\epsilon^+(x) \cap [2, z]}} f(p_1) \cdots f(p_m),$$

which, after grouping terms according to the distinct primes among $p_1, \ldots, p_m$, becomes

(9.4)

$$S(x)^{-1} \sum_{\substack{s \leq m \\ }} \sum_{\substack{p_1 < \cdots < p_s \leq z \\ p_1, \ldots, p_s \in \mathcal{P}_\epsilon^+(x)}} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} f(p_1)^{k_1} \cdots f(p_s)^{k_s} \sum_{\substack{n \leq x \\ p_1 \cdots p_s | n}} \alpha(n).$$

By (3.4) we have

$$\sum_{\substack{n \leq x \\ p_1 \cdots p_s | n}} \alpha(n) = \sum_{\substack{q \leq x \\ R_q = p_1 \cdots p_s}} \alpha(q) \sum_{\substack{n' \leq x/q \\ \gcd(n', q) = 1}} \alpha(n')$$

$$\ll \lambda_\alpha x^{\sigma_0} \sum_{\substack{q \leq x \\ R_q = p_1 \cdots p_s}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{3x}{q} \right)^{\beta-1}.$$

Appealing to (3.3) we derive

$$
\sum_{\substack{q \le x \\ R_q = p_1 \cdots p_s}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{3x}{q} \right)^{\beta-1}
$$

$$
\ll (\log x)^{\beta-1} \sum_{\substack{q \le \sqrt{x} \\ R_q = p_1 \cdots p_s}} \frac{\alpha(q)}{q^{\sigma_0}} + \sum_{\substack{\sqrt{x} < q \le x \\ R_q = p_1 \cdots p_s}} \frac{\alpha(q)}{q^{\sigma_0}} \left(\log \frac{3x}{q} \right)^{\beta-1}
$$

$$
\ll (\log x)^{\beta-1} \prod_{i=1}^{s} \sum_{\nu=1}^{\infty} \frac{\alpha(p_i^{\nu})}{p_i^{\sigma_0 \nu}} + \frac{(\log x)^{s+\beta-2}}{(\sqrt{x})^{1-\varrho_0}}
$$

$$
= (\log x)^{\beta-1} \prod_{i=1}^{s} \left(\frac{\alpha(p_i)}{p_i^{\sigma_0}} + \psi_0(p_i) \right) + \frac{(\log x)^{s+\beta-2}}{(\sqrt{x})^{1-\varrho_0}}.
$$

These estimates together with (3.4) imply that (9.4) is $\ll \Sigma_1 + \Sigma_2$, where

$$
\Sigma_1 := \sum_{s \le m} \sum_{\substack{p_1 < \cdots < p_s \le z \\ p_1, \ldots, p_s \in \mathcal{P}_\epsilon^+(x)}} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} |f(p_1)^{k_1} \cdots f(p_s)^{k_s}|
$$

$$
\times \prod_{i=1}^{s} \left(\frac{\alpha(p_i)}{p_i^{\sigma_0}} + \psi_0(p_i) \right),
$$

$$
\Sigma_2 := \frac{(\log x)^{m-1}}{(\sqrt{x})^{1-\varrho_0}}
$$

$$
\times \sum_{s \le m} \sum_{\substack{p_1 < \cdots < p_s \le z \\ p_1, \ldots, p_s \in \mathcal{P}_\epsilon^+(x)}} \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s} |f(p_1)^{k_1} \cdots f(p_s)^{k_s}|.
$$

Since $f(p) \le K \sqrt{B^*(x)}$ for all $p \in \mathcal{P}_\epsilon^+(x)$, we have

$$
\Sigma_2 \ll \frac{(\log x)^{m-1}}{(\sqrt{x})^{1-\varrho_0}} \pi(z)^m B^*(x)^{m/2} = o(B^*(x)^{m/2}) \ll \epsilon B^*(x)^{m/2}.
$$

To bound Σ_1, we observe

$$
|f(p_1)^{k_1} \cdots f(p_s)^{k_s}| \ll B^*(x)^{(m-s)/2} |f(p_1) \cdots f(p_s)|.
$$

Thus,

$$
\Sigma_1 \le \sum_{s \le m} B^*(x)^{(m-s)/2} \frac{1}{s!} \left(\sum_{\substack{p \le z \\ p \in \mathcal{P}_\epsilon^+(x)}} \left(\alpha(p) \frac{|f(p)|}{p^{\sigma_0}} + \psi_0(p) \right) \right)^s
$$

$$
\times \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s}
$$

$$= \sum_{s \leq m} B^*(x)^{(m-s)/2} \frac{1}{s!} (o(\epsilon^{-1} \sqrt{B^*(x)}))^s \sum_{\substack{k_1 + \cdots + k_s = m \\ k_1, \ldots, k_s \in \mathbb{N}}} \binom{m}{k_1, \ldots, k_s}$$

$$\ll \epsilon B^*(x)^{m/2}.$$

Combining these estimates completes the proof of (9.3) in the case $\beta \neq 1$.

As mentioned in Section 2, Corollary 2.5 is an immediate consequence of Theorem 2.4 when f is strongly additive. The transition to the general additive case is then accomplished by applying the following analogue of [29, Theorem B]. This is the only place where we need to make use of characteristic functions.

LEMMA 9.1. *Let* $f \colon \mathbb{N} \to \mathbb{R}$ *be an additive function, and let* $\alpha \in \mathcal{M}^*$ *with parameters* $A_0, \beta, \sigma_0, \vartheta_0, \varrho_0, r$. *Denote by* $\tilde{f}$ *the strongly additive contraction of* f. *Furthermore, suppose that* $B(x) \to \infty$ *as* $x \to \infty$. *Then* $X_N(n) := (f(n) - A(N))/\sqrt{B(N)}$ *admits a limiting distribution function with respect to the natural probability measure induced by* α *if and only if* $\widetilde{X}_N(n) := (\tilde{f}(n) - A(N))/\sqrt{B(N)}$ *does, in which case they share the same limiting distribution function.*

Proof. As before, we shall assume $A_0 \in (0, 1)$. For each $N \in \mathbb{N}$, the distribution functions of $X_N(n)$ and $\widetilde{X}_N(n)$ are given by

$$\Phi_N(V) = S(N)^{-1} \sum_{\substack{n \leq N \\ X_N \leq V}} \alpha(n), \quad \widetilde{\Phi}_N(V) = S(N)^{-1} \sum_{\substack{n \leq N \\ \widetilde{X}_N \leq V}} \alpha(n),$$

respectively. We have to show that $\Phi_N(V)$ converges weakly to a distribution function as $N \to \infty$ if and only if $\widetilde{\Phi}_N(V)$ does, in which case they converge weakly to the same limit. Note that the characteristic functions of $X_N(n)$ and $\widetilde{X}_N(n)$ are

$$\varphi_N(t) = S(N)^{-1} \sum_{n \leq N} \alpha(n) e^{it X_N(n)}, \quad \widetilde{\varphi}_N(t) = S(N)^{-1} \sum_{n \leq N} \alpha(n) e^{it \widetilde{X}_N(n)},$$

respectively. By Lévy's continuity theorem [31, Theorem III.2.6], it suffices to show that

(9.5) $$\lim_{N \to \infty} (\varphi_N(t) - \widetilde{\varphi}_N(t)) = 0$$

for any given $t \in \mathbb{R}$.

To prove this, let us fix $t \in \mathbb{R}$ and let $\epsilon \in (0, 1/(2|t| + 1))$ be arbitrary. Denote by $J_\epsilon(N)$ the greatest integer not exceeding $\sqrt{N}$ such that $|f(n)| \leq \epsilon \sqrt{B(N)}$ for all $1 \leq n \leq J_\epsilon(N)$. Since $B(N) \nearrow \infty$ as $N \to \infty$, we have $J_\epsilon(N) \nearrow \infty$ as $N \to \infty$. By (3.4),

$$\left|\varphi_N(t) - \widetilde{\varphi}_N(t)\right| \le S(N)^{-1} \sum_{n \le N} \alpha(n) \left|\exp\left(it\frac{f(n) - \tilde{f}(n)}{\sqrt{B(N)}}\right) - 1\right|$$

$$= S(N)^{-1} \sum_{\substack{a \le N \\ a\,\text{squareful}}} \alpha(a) \left|\exp\left(it\frac{f(a) - f(R_a)}{\sqrt{B(N)}}\right) - 1\right| \sum_{\substack{b \le N/a \\ b\,\text{squarefree} \\ \gcd(b,a)=1}} \alpha(b)$$

$$\ll S(N)^{-1} \lambda_\alpha N^{\sigma_0} \sum_{\substack{a \le N \\ a\,\text{squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log\frac{3N}{a}\right)^{\beta-1} \left|\exp\left(it\frac{f(a) - f(R_a)}{\sqrt{B(N)}}\right) - 1\right|.$$

From (2.1) and (2.3) it follows that

$$\sum_{\substack{a=1 \\ a\,\text{squareful}}}^{\infty} \frac{\alpha(a)}{a^s} = \prod_p \left(1 + \sum_{\nu \ge 2} \frac{\alpha(p^\nu)}{p^{\nu s}}\right)$$

is absolutely convergent for $s \in \mathbb{C}$ with $\Re(s) > \max(\varrho_0, r) + \sigma_0 - 1$. Thus

$$c(\delta) := \sum_{\substack{a=1 \\ a\,\text{squareful}}}^{\infty} \frac{\alpha(a)}{a^{\sigma_0-\delta}} = \prod_p \left(1 + \sum_{\nu \ge 2} \frac{\alpha(p^\nu)}{p^{\nu(\sigma_0-\delta)}}\right) < \infty$$

for any $\delta < 1 - \max(\varrho_0, r)$. Since

$$\left|it\frac{f(a) - f(R_a)}{\sqrt{B(N)}}\right| \le 2\epsilon|t| < 1$$

for all $a \le J_\epsilon(N)$, this implies

$$\sum_{\substack{a \le J_\epsilon(N) \\ a\,\text{squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log\frac{3N}{a}\right)^{\beta-1} \left|\exp\left(it\frac{f(a) - f(R_a)}{\sqrt{B(N)}}\right) - 1\right| \ll \epsilon|t|(\log N)^{\beta-1}.$$

Now fix $0 < \delta < 1 - \max(\varrho_0, r)$. By partial summation we have

$$\sum_{\substack{a \le x \\ a\,\text{squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} = c(0) - \int_x^\infty \frac{1}{t^\delta}\, d\left(\sum_{\substack{a \le t \\ a\,\text{squareful}}} \frac{\alpha(a)}{a^{\sigma_0-\delta}}\right) = c(0) + o(x^{-\delta})$$

when x is sufficiently large. It follows that

$$\sum_{\substack{J_\epsilon(N) < a \le N \\ a\,\text{squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log\frac{3N}{a}\right)^{\beta-1} \left|\exp\left(it\frac{f(a) - f(R_a)}{\sqrt{B(N)}}\right) - 1\right|$$

$$\le 2 \sum_{\substack{J_\epsilon(N) < a \le N \\ a\,\text{squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log\frac{3N}{a}\right)^{\beta-1}$$

$$= 2 \int_{J_\epsilon(N)}^{N} \left(\log \frac{3N}{t} \right)^{\beta-1} d\left(\sum_{\substack{a \leq t \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \right)$$

$$= o(N^{-\delta}) + o((\log N)^{\beta-1} J_\epsilon(N)^{-\delta}) + o\left(\int_{J_\epsilon(N)}^{N} t^{-1-\delta} \left(\log \frac{3N}{t} \right)^{\beta-2} dt \right)$$

for sufficiently large N. By a change of variable we see that

$$\int_{J_\epsilon(N)}^{N} t^{-1-\delta} \left(\log \frac{3N}{t} \right)^{\beta-2} dt = (3N)^{-\delta} \int_{\log 3}^{\log(3N/J_\epsilon(N))} e^{\delta t} t^{\beta-2} dt$$

$$\ll (3N)^{-\delta} \left(\frac{3N}{J_\epsilon(N)} \right)^{\delta} \left(\log \frac{3N}{J_\epsilon(N)} \right)^{\beta-2}$$

$$\ll (\log N)^{\beta-2} J_\epsilon(N)^{-\delta}.$$

Hence,

$$\sum_{\substack{J_\epsilon(N)<a\leq N \\ a \text{ squareful}}} \frac{\alpha(a)}{a^{\sigma_0}} \left(\log \frac{3N}{a} \right)^{\beta-1} \left| \exp\left(it\frac{f(a)-f(R_a)}{\sqrt{B(N)}} \right) - 1 \right|$$

$$= o((\log N)^{\beta-1} J_\epsilon(N)^{-\delta})$$

for sufficiently large N. Gathering the above estimates, we obtain

$$\varphi_N(t) - \widetilde{\varphi}_N(t) \ll \epsilon|t| + o(J_\epsilon(N)^{-\delta})$$

for sufficiently large N, where the implicit constants are independent of t, ϵ and N. From this estimate we infer that

$$\limsup_{N\to\infty} |\varphi_N(t) - \widetilde{\varphi}_N(t)| = O(\epsilon|t|),$$

where the implicit constant is independent of t and ϵ. As $\epsilon \in (0, 1/(2|t|+1))$ is arbitrary, we obtain (9.5) as desired. ∎

REMARK 9.1. Let $\alpha(n) = \tau(n)^2/n^{11}$, where τ is Ramanujan's τ-function, and define the additive function $f(n)$ by $f(p^\nu) = \log\sqrt{\alpha(p^\nu)}$ if $\alpha(p^\nu) \neq 0$ and $f(p^\nu) = 0$ otherwise, where p^ν is any prime power. Then $\alpha(n)$ satisfies conditions (2.1)–(2.4) with any fixed $A_0 > 0$, $\beta = 1$, $\sigma_0 = 1$, $\vartheta_0 = 0$, and any fixed $\varrho_0 \in (0,1)$ and $r \in (1/2, 1)$. Moreover, $\alpha(n) \leq d(n)^2$ by Deligne's bound [9]. As alluded to in Section 2, Elliott [12] showed, using ideas from probability theory, that the limiting distribution of $(f(n) - A(x))/\sqrt{B(x)}$ with respect to the natural probability measure induced by α is the standard Gaussian distribution. In fact, we can derive his result from Corollary 2.5 in combination with Lemma 9.1 and [12, Lemma 7] without difficulty. In

comparison to Elliott's probabilistic approach, our approach enables us to get around some of the complications resulting from the analysis of $\tau(n)$.

To illustrate this, let us consider the strongly additive function $f_0(n)$ defined by $f_0(p) = \log\sqrt{\alpha(p)}$ if $p \notin E_0$ and $f_0(p) = 0$ otherwise, where $E_0 := \{p > 2 \colon \alpha(p) \le \exp(-2\sqrt[3]{\log\log p})\}$. Denote by $A_0(x)$ and $B_0(x)$, respectively, the expected mean and variance of $f_0(n)$ weighted by $\alpha(n)$. It can be shown [12, Lemma 7] that $B(x) \asymp \log\log x$. Since $t|\log t| \le \sqrt{t}$ for all $t \in [0,1]$, we have

$$\sum_{\substack{p \le x \\ p \in E_0}} \alpha(p) \frac{|f(p)|}{p} \le \sum_{\substack{p \le x \\ p \in E_0}} \frac{\sqrt{\alpha(p)}}{p} \le \sum_{p > 2} \frac{1}{p} \exp\left(-\sqrt[3]{\log\log p}\right) < \infty.$$

It follows that $A_0(x) = A(x) + O(1)$. A similar argument shows that $B_0(x) = B(x) + O(1) \asymp \log\log x$. Thus, $f_0(p) = O(B_0(p)^{1/3})$ for all p, which shows that $f_0(n)$ satisfies the hypotheses in Corollary 2.5. Hence, the limiting distribution of $(f_0(n) - A(x))/\sqrt{B(x)}$ with respect to the natural probability measure induced by α is the standard Gaussian distribution.

To complete our argument, let $\widetilde{f}$ be the strongly additive contraction of f. Then $f_0(n) \ge \widetilde{f}(n)$ for all $n \in \mathbb{N}$. Moreover, Deligne's bound and the fact that $\tau(n) \in \mathbb{Z}$ for all $n \in \mathbb{N}$ imply that $-(11\log p)/2 \le f(p) \le \log 2$ whenever $\alpha(p) \ne 0$. Since

$$\sum_{\nu \ge 1} \frac{\alpha(p^\nu)}{p^\nu} \le \frac{\alpha(p)}{p} + \sum_{\nu \ge 2} \frac{(\nu+1)^2}{p^\nu} = \frac{\alpha(p)}{p} + O\left(\frac{1}{p^2}\right),$$

we have

$$S(x)^{-1} \sum_{n \le x} \alpha(n)(f_0(n) - \widetilde{f}(n)) = S(x)^{-1} \sum_{\substack{p \le x \\ p \in E_0}} |f(p)| \sum_{\substack{n \le x \\ p \mid n}} \alpha(n)$$

$$= S(x)^{-1} \sum_{\substack{p \le x \\ p \in E_0}} |f(p)| \sum_{\nu \ge 1} \alpha(p^\nu) \sum_{\substack{n' \le x/p^\nu \\ p \nmid n'}} \alpha(n')$$

$$\ll S(x)^{-1} \lambda_\alpha x \sum_{\substack{p \le x \\ p \in E_0}} |f(p)| \sum_{\nu \ge 1} \frac{\alpha(p^\nu)}{p^\nu}$$

$$\ll \sum_{\substack{p \le x \\ p \in E_0}} \alpha(p) \frac{|f(p)|}{p} + O\left(\sum_{p > 2} \frac{\log p}{p^2}\right) \ll 1.$$

This estimate is sufficient for us to conclude that the limiting distribution of $(\widetilde{f}(n) - A(x))/\sqrt{B(x)}$ with respect to the natural probability measure induced by α is also the standard Gaussian distribution. By Lemma 9.1, the same is true for $(f(n) - A(x))/\sqrt{B(x)}$.

10. Concluding remarks. Although in the present paper we only focused on the subclass $\mathcal{M}^*$ of multiplicative functions, it is also of interest to consider weight functions $\alpha(n)$ which satisfy certain Landau–Selberg–Delange type conditions. Given more information about $\alpha(n)$ and its associated Dirichlet series $F(s) = \sum_{n=1}^{\infty} \alpha(n) n^{-s}$, better results are obtainable in some circumstances. Below we give a brief description of the method in the special case where $F(s)$ is close to an integral power of the Riemann zeta-function $\zeta(s)$.

For $s \in \mathbb{C}$, we write $\sigma = \Re(s)$ and $t = \Im(s)$. Let $\alpha \colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a multiplicative function whose Dirichlet series $F(s) = \sum_{n=1}^{\infty} \alpha(n) n^{-s}$ is absolutely convergent for $s \in \mathbb{C}$ with $\sigma > \sigma_0$, where $\sigma_0 > 0$ is constant. Suppose that there exist constants $\beta \in \mathbb{N}$, $0 < \vartheta_0 < \sigma_0$, $B > 0$ and $0 < \delta < 1$ such that $H_\beta(s) := F(s)\zeta(s - \sigma_0 + 1)^{-\beta}$ has an analytic continuation in the half-plane $\sigma \geq \vartheta_0$ with

$$\lim_{s \to \sigma_0} F(s)(s - \sigma_0)^\beta > 0,$$

and such that $|H_\beta(s)| \leq B(1 + |t|)^{1-\delta}$ for all $s \in \mathbb{C}$ with $\sigma \geq \vartheta_0$. It is clear that $F(s)$ has (absolute) abscissa of convergence σ_0. Adapting the argument used in the proof of [24, Lemma 2.1] or [31, Theorem II.5.2], one can show that there exists some constant $\epsilon_0 > 0$ such that

$$(10.1) \qquad S(x) = \frac{1}{\sigma_0} \operatorname{Res}_{s = \sigma_0} \left(\frac{F(s) x^s}{s - \sigma_0 + 1} \right)$$

$$- x^{\sigma_0} (\log x)^{\beta - 1} \sum_{k=1}^{\beta - 1} \sum_{j=0}^{k-1} c_{j,k} \frac{\mu_j(\beta)}{(\log x)^k} + O(B x^\vartheta)$$

uniformly for all $x \geq 3$ and $\vartheta \in (\sigma_0 - \epsilon_0, \sigma_0)$, where

$$\mu_k(\beta) := \frac{1}{k!} \cdot \frac{d^k}{ds^k} \left(\frac{F(s)(s - \sigma_0)^\beta}{s - \sigma_0 + 1} \right) \bigg|_{s = \sigma_0}, \qquad c_{j,k} := \frac{(-1)^{k-j}(\sigma_0 - 1)}{(\beta - k - 1)! \sigma_0^{k-j+1}},$$

and the implicit constant in the error term depends at most on $\beta, \sigma_0, \vartheta_0, \delta, \epsilon_0$. Notably, one gains an asymptotic for $S(x)$ with a power-saving error term uniformly in B, in contrast to what is provided by (3.4). Furthermore, suppose that there exists a constant $\lambda > 0$ such that $\alpha(p^\nu) = O((\lambda p^{\sigma_0 - 1})^\nu)$ for all prime powers p^ν. Let

$$F(s, a) := \prod_{p \mid a} \left(1 - \left(\sum_{\nu=0}^{\infty} \alpha(p^\nu) p^{-\nu s} \right)^{-1} \right)$$

for $s \in \mathbb{C}$ with $\sigma \geq \vartheta_0$ and squarefree $a \in \mathbb{N}$. When $s = \sigma_0$, this definition coincides with the one introduced in Lemma 3.3. As in the proof of that

lemma, it is not hard to show that

$$(10.2) \qquad F(s,p) = \frac{\alpha(p)}{p^s} + O\left(\frac{\alpha(p)^2}{p^{2\sigma}} + \frac{1}{p^{2(\sigma-\sigma_0+1)}}\right)$$

for all $s \in \mathbb{C}$ with $\sigma \geq \vartheta_0$ and all sufficiently large p. In addition,

$$\sum_{\substack{n=1 \\ a|n}}^{\infty} \frac{\alpha(n)}{n^s} = \left(\prod_{p\nmid a}\sum_{\nu=0}^{\infty}\alpha(p^\nu)p^{-\nu s}\right)\left(\prod_{p|a}\sum_{\nu=1}^{\infty}\alpha(p^\nu)p^{-\nu s}\right)$$

$$= F(s)\left(\prod_{p|a}\sum_{\nu=0}^{\infty}\alpha(p^\nu)p^{-\nu s}\right)^{-1}\left(\prod_{p|a}\sum_{\nu=1}^{\infty}\alpha(p^\nu)p^{-\nu s}\right) = F(s)F(s,a)$$

for $s \in \mathbb{C}$ with $\sigma > \sigma_0$ and squarefree $a \in \mathbb{N}$. Applying (10.1) to the above Dirichlet series expansion of $F(s)F(s,a)$ and using (10.2) to obtain upper bounds for $H_\beta(s)F(s,a)$ uniformly in $\sigma \geq \vartheta_0$, we see that there exist constants $\epsilon \in (0,1)$, $Q_0 \geq 2$ and $d_{j,k} \in \mathbb{R}$, where $0 \leq j < k < \beta$, such that

$$(10.3) \qquad \sum_{\substack{n\leq x \\ a|n}}\alpha(n) = \frac{\mu_0(\beta)F(\sigma_0,a)}{(\beta-1)!\sigma_0}x^{\sigma_0}(\log x)^{\beta-1}$$

$$+ x^{\sigma_0}(\log x)^{\beta-1}\sum_{k=1}^{\beta-1}\sum_{j=0}^{k}d_{j,k}\frac{F^{(j)}(\sigma_0,a)}{(\log x)^k}$$

$$+ O(B2^{O(\omega(a))}a^{\sigma_0-1}(x/a)^\vartheta)$$

uniformly for all $x \geq 3$, $\vartheta \in (\sigma_0 - \epsilon, \sigma_0)$ and squarefree $a \in \mathbb{N}$ with $P^-(a) > Q_0$, where $F^{(j)}(\sigma_0,a)$ is the jth order derivative of $F(s,a)$ with respect to s evaluated at $s = \sigma_0$. Again, one may compare this result with Lemma 3.3.

Now, if $f \colon \mathbb{N} \to \mathbb{R}$ is a strongly additive function with $|f(p)| \leq M$ for all primes p, where $M > 0$ is constant, and if $0 < h_0 < (3/2)^{2/3}$ is fixed but arbitrary, then by using (10.3) as a substitute for Lemma 3.3 and arguing as before with the technique of [24, Section 4.2], we deduce that

$$M(x;m) = C_m B(x)^{m/2}\left(\chi_m + O\left(\frac{m^{3/2}}{\sqrt{B(x)}}\right)\right)$$

uniformly for all sufficiently large x and all $1 \leq m \leq h_0(B(x)/M^2)^{1/3}$, provided that $B(x) \to \infty$ as $x \to \infty$. Analogously, let $f \colon \mathbb{N} \to \mathbb{R}$ be strongly additive such that $f(p) = O(\sqrt{B(p)})$ for all primes p, $B(x) \to \infty$ as $x \to \infty$, and

$$\sum_{\substack{p\leq x \\ |f(p)|>\epsilon\sqrt{B(x)}}}\alpha(p)\frac{f(p)^2}{p} = o(B(x))$$

for any given $\epsilon > 0$. Then $M(x;m) = (\mu_m + o(1))B(x)^{m/2}$ for every fixed

$m \in \mathbb{N}$. These results supplement Theorems 2.1 and 2.4. It may be worth pointing out that in the proofs of these results one can simply take $z = x^{1/v}$ with v being a suitable constant multiple of m. We invite the reader to fill in the details.

One of the key ingredients in the proof of Theorem 2.1 is an asymptotic formula for

$$\sum_{\substack{n \le x \\ d \mid n}} \alpha(n),$$

which is provided by Lemma 3.3. More generally, let $\mathcal{A}(x) = \{a_n\}_{n \le x}$ be a nondecreasing sequence of positive integers, and suppose that

$$(10.4) \qquad \mathcal{A}_{d,\alpha}(x) := \sum_{\substack{n \le x \\ d \mid a_n}} \alpha(n) = \rho(d) S(x) + r_d(x)$$

for squarefree integers $d \in \mathbb{N}$, where $\rho \colon \mathbb{N} \to [0,1]$ is a multiplicative function, and $r_d(x)$ is a remainder term which is expected to be small for all d or small on average over d. Here, $\rho(d)$ can be viewed as the density of the set $\{n \in \mathbb{N} \colon d \mid a_n\}$ with respect to the probability measure induced by α. In this sieve-theoretic setting one can derive, without much difficulty, an analogue of [17, Proposition 4]. It may be of interest to determine if such an analogue can be used to obtain general weighted Erdős–Kac theorems for various interesting sequences $\{a_n\}$ studied relatively recently, including $g(p_n)$, $\varphi(n)$, the Carmichael function $\lambda(n)$, and the aliquot sum $s(n) := \sigma(n) - n$, where $g \in \mathbb{Z}[x]$ is an irreducible polynomial, p_n is the nth prime, and $\lambda(n)$ denotes the exponent of $(\mathbb{Z}/n\mathbb{Z})^\times$ (see [20, 14, 16, 27]). The same approach may also be adapted to prove results of weighted Erdős–Kac type for short intervals as well as in the function field setting. We hope to return to these and other related problems in the future.

References

[1] K. Alladi, *A study of the moments of additive functions using Laplace transforms and sieve methods*, in: Number Theory (Ootacamund, 1984), Lecture Notes in Math. 1122, Springer, Berlin, 1985, 1–37.

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

[3] P. Billingsley, *Probability and Measure*, 3rd ed., Wiley Ser. Probab. Math. Statist., Wiley-Interscience, New York, 1995.

[4] E. Bombieri, J. B. Friedlander, and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Math. 156 (1986), 203–251.

[5] R. de la Bretèche and G. Tenenbaum, *Remarks on the Selberg–Delange method*, Acta Arith. 200 (2021), 349–369.

[6] H. Delange, *Sur le nombre des diviseurs premiers de n*, C. R. Acad. Sci. Paris 237 (1953), 542–544.

[7] H. Delange, *Sur un théorème d'Erdős et Kac*, Acad. Roy. Belg. Bull. Cl. Sci. (5) 42 (1956), 130–144.

[8] H. Delange and H. Halberstam, *A note on additive functions*, Pacific J. Math. 7 (1957), 1551–1556.

[9] P. Deligne, *La conjecture de Weil. I*, Publ. Math. Inst. Hautes Études Sci. 43 (1974), 273–307.

[10] R. Durrett, *Probability: Theory and Examples*, 5th ed., Cambridge Ser. Statist. Probab. Math. 49, Cambridge Univ. Press, Cambridge, 2019.

[11] D. Elboim and O. Gorodetsky, *Multiplicative arithmetic functions and the generalized Ewens measure*, Israel J. Math. 262 (2024), 143–189.

[12] P. D. T. A. Elliott, *A central limit theorem for Ramanujan's tau function*, Ramanujan J. 29 (2012), 145–161.

[13] P. D. T. A. Elliott, *Corrigendum: central limit theorems for classical cusp forms*, Ramanujan J. 36 (2015), 99–102.

[14] P. Erdős, A. Granville, C. Pomerance, and C. Spiro, *On the normal behavior of the iterates of some arithmetic functions*, in: Analytic Number Theory, in Honor of Paul T. Bateman, Progr. Math. 85, Birkhäuser, Boston, 1990, 165–204.

[15] P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math. 62 (1940), 738–742.

[16] P. Erdős and C. Pomerance, *On the normal number of prime factors of $\varphi(n)$*, Rocky Mountain J. Math. 15 (1985), 343–352.

[17] A. Granville and K. Soundararajan, *Sieving and the Erdős–Kac theorem*, in: Equidistribution in Number Theory: An Introduction, NATO Sci. Ser. II Math. Phys. Chem. 237, Springer, Dordrecht, 2007, 15–27.

[18] H. Halberstam, *On the distribution of additive number-theoretic functions*, J. London Math. Soc. 30 (1955), 43–53.

[19] H. Halberstam, *On the distribution of additive number-theoretic functions (II)*, J. London Math. Soc. 31 (1956), 1–14.

[20] H. Halberstam, *On the distribution of additive number-theoretic functions (III)*, J. London Math. Soc. 31 (1956), 14–27.

[21] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number n*, Quart. J. Pure Appl. Math. 48 (1971), 76–92.

[22] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 6th ed., Oxford Univ. Press, Oxford, 2008.

[23] M. Kac, *Probability methods in some problems of analysis and number theory*, Bull. Amer. Math. Soc. 55 (1949), 641–665.

[24] R. Khan, M. B. Milinovich, and U. Subedi, *A weighted version of the Erdős–Kac theorem*, J. Number Theory 239 (2022), 1–20.

[25] W. J. LeVeque, *On the size of certain number-theoretic functions*, Trans. Amer. Math. Soc. 66 (1949), 440–463.

[26] D. T. Nguyen, *Generalized divisor functions in arithmetic progressions: I*, J. Number Theory 227 (2021), 30–93.

[27] P. Pollack and L. Troupe, *Sums of proper divisors follow the Erdős–Kac law*, Proc. Amer. Math. Soc. 151 (2023), 977–988.

[28] A. Rényi and P. Turán, *On a theorem of Erdős–Kac*, Acta Arith. 4 (1958), 71–84.

[29] H. N. Shapiro, *Distribution functions of additive arithmetic functions*, Proc. Nat. Acad. Sci. USA 42 (1956), 426–430.

[30] J. M. Song, *Sums of nonnegative multiplicative functions over integers without large prime factors I*, Acta Arith. 97 (2001), 329–351.

[31] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, 3rd ed., Grad. Stud. Math. 163, Amer. Math. Soc., Providence, RI, 2015.

[32] G. Tenenbaum, *Moyennes effectives de fonctions multiplicatives complexes*, Ramanujan J. 44 (2017), 641–701.

[33] G. Tenenbaum, *Correction to [32]*, Ramanujan J. 53 (2020), 243–244.

[34] P. Turán, *On a theorem of Hardy and Ramanujan*, Ramanujan J. 9 (1934), 274–276.

Kai (Steve) Fan
Max-Planck-Institut für Mathematik
53111 Bonn, Germany
E-mail: steve.fan.1024@gmail.com