

THE DISTRIBUTION OF SUMS AND PRODUCTS OF ADDITIVE FUNCTIONS

GREG MARTIN AND LEE TROUPE

ABSTRACT. The celebrated Erdős–Kac theorem says, roughly speaking, that the values of additive functions satisfying certain mild hypotheses are normally distributed. In the intervening years, similar normal distribution laws have been shown to hold for certain non-additive functions and for amenable arithmetic functions over certain subsets of the natural numbers. Continuing in this vein, we show that if $g_1(n), \dots, g_k(n)$ is a collection of functions satisfying certain mild hypotheses for which an Erdős–Kac-type normal distribution law holds, and if $Q(x_1, \dots, x_k)$ is a polynomial with nonnegative real coefficients, then $Q(g_1(n), \dots, g_k(n))$ also obeys a normal distribution law. We also show that a similar result can be obtained if the set of inputs n is restricted to certain subsets of the natural numbers, such as shifted primes. Our proof uses the method of moments. We conclude by providing examples of our theorem in action.

1. INTRODUCTION

In 1939, Erdős and Kac [3] used probabilistic methods to prove a striking fact concerning a class of strongly additive functions, establishing the following foundational result in the field of probabilistic number theory.

Theorem 1.1 (Erdős/Kac). *Let g be a strongly additive function with $|g(p)| \leq 1$ for all primes p . Define*

$$A_g(x) = \sum_{p \leq x} \frac{g(p)}{p} \quad \text{and} \quad B_g(x) = \left(\sum_{p \leq x} \frac{g^2(p)}{p} \right)^{1/2}, \quad (1)$$

and assume that $B_g(x)$ is unbounded. Then for every real number u ,

$$\lim_{x \rightarrow \infty} \frac{1}{x} \#\{n \leq x : g(n) < A_g(x) + uB_g(x)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt. \quad (2)$$

Here, as usual, a function $g: \mathbb{N} \rightarrow \mathbb{R}$ is *additive* if $g(mn) = g(m) + g(n)$ whenever $(m, n) = 1$; an additive function g is *strongly additive* if $g(p^\alpha) = g(p)$ for every prime p and positive integer α . In particular, a strongly additive function is completely determined by its values on prime inputs. For example, $\omega(n)$, the number of distinct prime factors of n , is a strongly additive function.

Theorem 1.1 tells us that the values of the normalized version $(g(n) - A_g(n))/B_g(n)$ of g are distributed, in the limit, exactly like random real numbers chosen from the standard normal distribution of mean 0 and variance 1. We codify this type of distributional law with the following terminology:

Definition 1.2. Let $g: \mathbb{N} \rightarrow \mathbb{R}$ and $A_g, B_g: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ be functions. We say that g satisfies an Erdős–Kac law with mean A_g and variance B_g if equation (2) holds for every real number u .

Different proofs and generalizations of Erdős and Kac's theorem abound. In a 1955 paper, Halberstam [5] calculated the m th moments of the quantity $(g(n) - A_g(x))/B_g(x)$ for $m \geq 1$, deducing the Erdős–Kac theorem from these calculations. In 2007, Granville and Soundararajan [4] calculated these moments using a technique both simpler and more adaptable than Halberstam's methods; as a result, one can prove Erdős–Kac laws for strongly additive functions g restricted to certain subsets of the natural numbers. An example of a result in this direction (though predating the work of Granville and Soundararajan) is due to Alladi [2], who proved an analogue of the Erdős–Kac theorem for the usual prime-factor-counting functions ω and Ω over friable integers.

In this paper, we adapt the methods of [4] to establish an Erdős–Kac law for arbitrary sums and products of strongly additive functions satisfying certain standard hypotheses:

Theorem 1.3. *Let $Q(T_1, \dots, T_\ell)$ be a polynomial with nonnegative real coefficients. Let g_1, g_2, \dots, g_ℓ be nonnegative strongly additive functions such that:*

- (a) *each $g_j(p) \ll_j 1$ uniformly for all primes p ;*
- (b) *for each $1 \leq j \leq \ell$, the series $\sum_p g_j^2(p)/p$ diverges.*

Define

$$\mu(g_j) = \sum_{p \leq x} \frac{g_j(p)}{p} \quad \text{and} \quad \kappa(g_i, g_j) = \sum_{p \leq x} \frac{g_i(p)g_j(p)}{p} \left(1 - \frac{1}{p}\right) \quad \text{and} \quad \sigma^2(g_i) = \kappa(g_i, g_i), \quad (3)$$

and suppose that $\sigma^2(g_i) \gg \mu(g_j)$ for sufficiently large x for all $1 \leq i, j \leq \ell$.

Then $Q(g_1(n), \dots, g_\ell(n))$ satisfies an Erdős–Kac law with mean

$$A_Q(x) = Q(\mu(g_1), \dots, \mu(g_\ell)) \quad (4)$$

and variance

$$B_Q(x)^2 = \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \frac{\partial Q}{\partial T_i}(\mu(g_1), \dots, \mu(g_\ell)) \frac{\partial Q}{\partial T_j}(\mu(g_1), \dots, \mu(g_\ell)) \kappa(g_i, g_j). \quad (5)$$

While all of the hypotheses of the theorem are necessary for our proof in its full generality, the assumptions on the relative sizes of the $\mu(g_j)$ and $\sigma^2(g_i)$ could usually be relaxed in specific situations; moreover, the assumption that the g_j be pointwise nonnegative is hardly used, other than to ensure that the covariances $\kappa(g_i, g_j)$ are nonnegative so that the right-hand side of equation (5) has a square root in general. In the same vein, we work with strongly additive functions for simplicity, but in principle the methods could establish Theorem 1.3 for additive functions whose values on prime powers can depend upon the power as well as the prime.

In the last section of this paper, we give examples of some special cases of this theorem, such as the Erdős–Kac law for a product of additive functions (including the integer powers of an additive function); these results, which are justified in Section 6, continue to use the notation from equation (3).

Corollary 1.4. *Let $g(n)$ be a strongly additive function satisfying the hypotheses of Theorem 1.3. For any positive integer δ , the function $g(n)^\delta$ satisfies an Erdős–Kac law with mean $\mu^\delta(g)$ and variance $\delta^2 \mu(g)^{2\delta-2} \sigma^2(g)$.*

Corollary 1.5. *Let g_1, \dots, g_ℓ be strongly additive functions satisfying the hypotheses of Theorem 1.3. Then for any $\ell \geq 1$, the function $g_1(n) \cdots g_\ell(n)$ satisfies an Erdős–Kac law with mean $\mu(g_1) \cdots \mu(g_\ell)$ and variance*

$$(\mu(g_1) \cdots \mu(g_\ell))^2 \sum_{1 \leq i, j \leq \ell} \frac{\kappa(g_i, g_j)}{\mu(g_i)\mu(g_j)}.$$

As it turns out, the proof of Theorem 1.3 is no harder if we replace $\{p \leq x\}$ with any finite set \mathcal{P} of primes. Similarly, we may replace the set of inputs $\{n \leq x\}$ with a finite set \mathcal{A} , at the cost of introducing further multiplicative functions $h(n)$ that measure the local densities of \mathcal{A} but with no other significant changes. As remarked in [4], this added flexibility allows for the derivation of several variants of the Erdős–Kac theorem. As an example, we establish the following analogue of Theorem 1.3 for polynomials evaluated at additive functions on shifted primes:

Theorem 1.6. *Let a be a fixed nonzero integer. Under the hypotheses of Theorem 1.3, the function $Q(g_1(p-a), \dots, g_\ell(p-a))$ satisfies an Erdős–Kac law with $A_Q(x)$ and $B_Q(x)$ as defined in equations (4) and (5). In other words, for every real number u ,*

$$\lim_{x \rightarrow \infty} \frac{1}{\pi(x)} \#\{p \leq x: Q(g_1(p-a), \dots, g_\ell(p-a)) < A_Q(x) + uB_Q(x)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt. \quad (6)$$

In the next section we describe the generalized setting, with arbitrary finite sets \mathcal{P} and \mathcal{A} as mentioned above and other notation used throughout the paper, and establish a general distributional limit law (Theorem 2.8) from the calculation of the appropriate m th moments (Theorem 2.7). Further, in that section, we also derive Theorems 1.3 and 1.6 from Theorem 2.8. The remainder of the paper will then be devoted to establishing Theorem 2.7, other than giving some examples in Section 6 that in particular establish Corollaries 1.4 and 1.5.

2. NOTATION AND RESTATEMENT OF MAIN THEOREM

Following the methods established in [4], we begin by setting up a sieve-theoretic framework within which we will recast Theorem 1.3.

Definition 2.1. Let \mathcal{A} denote a finite set (or multiset) of positive integers. For every positive integer d , set $\mathcal{A}_d = \{a \in \mathcal{A}: d \mid a\}$. As is usual in sieve theory, we suppose that there exists a positive real number X that is a good approximation to the cardinality of \mathcal{A} , and a multiplicative function $h(d)$ (satisfying $0 \leq h(d) \leq d$) such that $\frac{h(d)}{d}X$ is a good approximation to $\#\mathcal{A}_d$. More explicitly, we define the remainder terms E_d by

$$\#\mathcal{A}_d = \frac{h(d)}{d}X + E_d, \quad (7)$$

and we suppose that we have some suitable control over E_d , individually or on average. It would certainly suffice for our purposes to have $E_d \ll d^c$ for some fixed $c < 1$.

In the setting of Theorem 1.3, for example, we will have $E_d \ll 1$. In general, one needs only to know the behavior of the terms E_d on average; for example, in the proof of Theorem 1.6 we will handle the terms E_d via the Bombieri–Vinogradov theorem.

Definition 2.2. Let \mathcal{P} denote a finite set of primes. For any nonnegative integer m , define

$$\mathcal{D}_k(\mathcal{P}) = \bigcup_{\ell=0}^k \{p_1 \cdots p_\ell : p_1, \dots, p_\ell \text{ are distinct elements of } \mathcal{P}\};$$

note that the elements of $\mathcal{D}_k(\mathcal{P})$ are squarefree integers contained in the interval $[1, (\max \mathcal{P})^k]$. In particular, $\mathcal{D}_0(\mathcal{P}) = \{1\}$.

For any function $g: \mathbb{N} \rightarrow \mathbb{R}$ (but usually for strongly additive functions g), we write

$$g^{\mathcal{P}}(a) = \sum_{\substack{p \in \mathcal{P} \\ p|a}} g(p).$$

(We remark that this notation would be sensible even if \mathcal{P} were an infinite set of primes, since any given integer a has only finitely many prime factors.) In keeping with the terminology and notation established in [4], we define the *mean* of g over \mathcal{P} as

$$\mu^{\mathcal{P}}(g) = \sum_{p \in \mathcal{P}} g(p) \frac{h(p)}{p}; \quad (8)$$

for example, $\mu^{\mathcal{P}}(1) = \sum_{p \in \mathcal{P}} h(p)/p$. We continue to borrow from probabilistic terminology by defining the *standard deviation* of g over \mathcal{P} as

$$\sigma^{\mathcal{P}}(g) = \left(\sum_{p \in \mathcal{P}} g(p)^2 \frac{h(p)}{p} \left(1 - \frac{h(p)}{p}\right) \right)^{1/2},$$

and we call the quantity $\sigma^{\mathcal{P}}(g)^2$ the *variance* of g over \mathcal{P} . Furthermore, for two functions $g_1, g_2: \mathbb{N} \rightarrow \mathbb{R}$ we define the *covariance* of g_1 and g_2 over \mathcal{P} by

$$\kappa^{\mathcal{P}}(g_1, g_2) = \sum_{p \in \mathcal{P}} g_1(p)g_2(p) \frac{h(p)}{p} \left(1 - \frac{h(p)}{p}\right).$$

Note that $\sigma^{\mathcal{P}}(g)^2 = \kappa^{\mathcal{P}}(g, g)$.

It is perhaps worth emphasizing that here the set \mathcal{P} represents the “good” primes whose contribution to the values of $g(n)$ we want to include, for all integers $n \in \mathcal{A}$; this is as opposed to sieve theory, where one usually names the set of “bad” primes to exclude, for the goal of seeking special elements of \mathcal{A} free of such prime factors.

For example, consider the case where \mathcal{P} is the set of all primes up to x , and $g(n) = \omega(n)$. A result of Mertens then says that the mean $\mu^{\mathcal{P}}(\omega) \sim \log \log x$ and the variance $\sigma^{\mathcal{P}}(\omega)^2 \sim \log \log x$, which are the quantities that we recognize from the classical Erdős–Kac theorem.

In the following few definitions, the objects being defined will depend upon ℓ functions $g_1, \dots, g_\ell: \mathbb{N} \rightarrow \mathbb{R}$; we suppress the dependence upon these functions from the notation.

Definition 2.3. We let \mathfrak{K} denote the maximum variance among the functions g_j :

$$\mathfrak{K} = \max_{1 \leq i, j \leq \ell} |\kappa^{\mathcal{P}}(g_i, g_j)| = \max_{1 \leq i \leq \ell} \sigma^{\mathcal{P}}(g_i)^2.$$

The latter equality is justified by the Cauchy–Schwarz inequality, which implies that

$$|\kappa^{\mathcal{P}}(g_i, g_j)| \leq \max\{\sigma^{\mathcal{P}}(g_i)^2, \sigma^{\mathcal{P}}(g_j)^2\}$$

Similarly, we let \mathfrak{M} denote the maximum mean among the functions g_j :

$$\mathfrak{M} = \max_{1 \leq i \leq \ell} \mu^{\mathcal{P}}(g_i).$$

Note that if there exists a positive constant G such that $0 \leq g_j(p) \leq G$ for all $1 \leq j \leq \ell$ and all $p \in \mathcal{P}$, then trivially $\mathfrak{K} \leq G\mathfrak{M}$.

Definition 2.4. Suppose that the functions g_1, \dots, g_ℓ take nonnegative values. Given any polynomial $Q(T_1, \dots, T_\ell)$ in ℓ variables with positive coefficients, we define

$$A_Q(\mathcal{P}) = Q(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell)) \quad (9)$$

and

$$B_Q(\mathcal{P}) = \left(\sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \frac{\partial Q}{\partial T_i}(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell)) \frac{\partial Q}{\partial T_j}(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell)) \kappa^{\mathcal{P}}(g_i, g_j) \right)^{1/2}. \quad (10)$$

Definition 2.5. For any polynomial $Q(t_1, \dots, t_\ell)$ in ℓ variables and any nonnegative integer m , define the m th moment

$$M_m = \sum_{a \in \mathcal{A}} (Q(g_1^{\mathcal{P}}(a), \dots, g_\ell^{\mathcal{P}}(a)) - A_Q(\mathcal{P}))^m.$$

Definition 2.6. For an integer $m \geq 0$, let

$$C_m = \begin{cases} \frac{m!}{2^{m/2}(m/2)!}, & \text{if } m \text{ is even,} \\ 0, & \text{if } m \text{ is odd} \end{cases}$$

denote the m th moment of the standard normal distribution.

These last two definitions foreshadow our strategy of deducing a normal limiting distribution from asymptotic formulas for the m th moments of a centered version of our quantity of interest. Indeed, under all of the notation above, we establish the following formula for the M_m :

Theorem 2.7. *Let $Q(T_1, \dots, T_\ell)$ be a polynomial of degree δ with nonnegative real coefficients, let m be a positive integer, and let \mathcal{P} be a finite set of primes. Let g_1, g_2, \dots, g_ℓ be nonnegative strongly additive functions such that, for some fixed $G > 0$, we have $g_j(p) \leq G$ for each j and for all primes $p \in \mathcal{P}$. Let \mathcal{A} be a finite multiset of positive integers such that, in the notation of Definition 2.1,*

$$\mu^{\mathcal{P}}(1)^{\delta m} \sum_{d \in \mathcal{D}_{\delta m}(\mathcal{P})} |E_d| \ll X \mathfrak{K}^{\delta m/2 - 1} \quad (11)$$

If m is even, then

$$M_m = C_m X B_Q(\mathcal{P})^m + O(X \mathfrak{M}^{\delta m - (m+1)/2}),$$

while if m is odd, then

$$M_m \ll X B_Q(\mathcal{P})^{m-1} + X \mathfrak{M}^{\delta m - (m+1)/2}.$$

Here, the implied constants may depend on the polynomial Q , the moment index m , the constant G , and the finite set \mathcal{A} .

Proving this theorem is the primary goal of this paper, and we will begin the proof in the next section. For the remainder of this section, however, we expound upon the consequences of Theorem 2.7 to various Erdős–Kac laws.

Let $\tilde{\mathcal{A}}$ denote an infinite set of positive integers, and define

$$\mathcal{A}(x) = \{a \in \tilde{\mathcal{A}} : a \leq x\} \quad (12)$$

(indeed, $\tilde{\mathcal{A}}$ may even be a multiset, as long as $\mathcal{A}(x)$ is finite for every x). Furthermore, let $\tilde{\mathcal{P}}$ denote an infinite set of primes, and define $\mathcal{P}(z) = \{p \in \tilde{\mathcal{P}} : p \leq z\}$. Using Theorem 2.7, we can prove the following result, which (as we subsequently show) implies Theorems 1.3 and 1.6.

Theorem 2.8. *Let $Q(T_1, \dots, T_\ell)$ be a polynomial of degree δ with nonnegative real coefficients. Let $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{P}}$ be infinite sets as described above, and assume that there exists a positive constant η such that*

$$\mu^{\mathcal{P}(t)}(1) = \sum_{p \leq t} \frac{h(p)}{p} \sim \eta \log \log t.$$

Let x and $z = z(x)$ be parameters tending to infinity, with z chosen so that for all integers $k \geq 0$,

$$\mu^{\mathcal{P}(z)}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P}(z))} |E_d| \ll_k X \mathfrak{K}^{k/2-1}, \quad (13)$$

where \mathfrak{K} is as in Definition 2.3 with \mathcal{P} replaced by $\mathcal{P}(z)$. Let g_1, g_2, \dots, g_ℓ be nonnegative strongly additive functions such that for each $1 \leq j \leq \ell$:

- (a) $g_j(p) \ll 1$ uniformly for all primes $p \in \tilde{\mathcal{P}}$;
- (b) the series $\sum_{p \in \tilde{\mathcal{P}}} g_j^2(p)/p$ diverges;
- (c) $\sigma^{\mathcal{P}(z)}(g_j)^2 \gg \mathfrak{M}$, with \mathfrak{M} as in Definition 2.3 with \mathcal{P} replaced by $\mathcal{P}(z)$.

Further, assume that

$$\frac{\log x}{\log z} = o(\mathfrak{M}^{1/2}). \quad (14)$$

Then the values $Q(g_1(a), \dots, g_\ell(a))$, as a runs through the elements of \mathcal{A} , satisfy an Erdős–Kac law with mean $A_Q(\mathcal{P}(x))$ from equation (9) and variance $B_Q(\mathcal{P}(x))^2$ from equation (10); in other words, for every $u \in \mathbb{R}$,

$$\lim_{x \rightarrow \infty} \frac{\#\{a \in \mathcal{A}(x) : Q(g_1(a), \dots, g_\ell(a)) < A_Q(\mathcal{P}(x)) + uB_Q(\mathcal{P}(x))\}}{\#\mathcal{A}(x)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt. \quad (15)$$

Before we show that Theorem 2.8 is a consequence of Theorem 2.7, we make the following convention regarding constants implicit in O , \ll , and \asymp notation: these constants may always depend upon the polynomial Q and its number of arguments (usually ℓ), the pointwise bound G for the additive functions g_j , the integer m dictating which moment we are looking at, and the finite sets \mathcal{A} and \mathcal{P} where appropriate. (We remind the reader that we write $f \asymp g$ if there exist absolute constants $C > c > 0$ such that $cf(x) \leq g(x) \leq Cf(x)$ throughout the domain of f and g .) On the other hand, these implicit constants will never depend upon the parameters x and z , and therefore will be independent of all quantities that depend upon these parameters, such as M_m , $\mu^{\mathcal{P}}(g_j)$ and \mathfrak{M} , and $\kappa^{\mathcal{P}}(g_i, g_j)$ and \mathfrak{K} .

Deduction of Theorem 2.8 from Theorem 2.7. For this proof, we will set $\mathcal{A} = \mathcal{A}(x)$ and $\mathcal{P} = \mathcal{P}(z)$ (and will write out $\mathcal{P}(x)$ explicitly when necessary). We begin by noting some consequences of the assumptions of Theorem 2.8. First, thanks to assumption (a), there exists a constant $G > 0$ such that $g_j(p) \leq G$ for each j and for all primes $p \in \tilde{\mathcal{P}}$. In particular, all of the hypotheses of Theorem 2.7 are satisfied for any positive integer m .

Next, for each $1 \leq j \leq \ell$ we have $\mathfrak{M} \geq \mu^{\mathcal{P}}(g_j) \geq \frac{1}{G}\sigma^{\mathcal{P}}(g_j)^2 \gg \mathfrak{M}$ by assumption (c), so that each $\mu^{\mathcal{P}}(g_j) \asymp \mathfrak{M}$ and hence $\mathfrak{K} \asymp \mathfrak{M}$; furthermore, $\mathfrak{M} \rightarrow \infty$ as $x \rightarrow \infty$ thanks to assumption (b). In summary, for every $1 \leq j \leq \ell$,

$$\sigma^{\mathcal{P}}(g_j)^2 \asymp \mu^{\mathcal{P}}(g_j) \asymp \mathfrak{K} \asymp \mathfrak{M} \rightarrow \infty. \quad (16)$$

Furthermore, by Cauchy–Schwarz, we also have $0 \leq \kappa^{\mathcal{P}}(g_i, g_j) \ll \mathfrak{M}$. Note also that when $m = 0$, the assumption (13) becomes simply $|E_1| \ll \frac{X}{\mathfrak{K}}$; this implies that $\#\mathcal{A} = \#\mathcal{A}_1 = \frac{h(1)}{1}X + E_1 = X(1 + O(\frac{1}{\mathfrak{K}})) \sim X$ as $x \rightarrow \infty$.

Since Q is of degree δ , there exists $1 \leq j_0 \leq \ell$ such that $\frac{\partial Q}{\partial T_{j_0}}(T_1, \dots, T_\ell)^2$ is of degree $2\delta - 2$. It follows from the nonnegativity of the covariances and of the coefficients of Q that

$$\begin{aligned} B_Q(\mathcal{P})^2 &= \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \frac{\partial Q}{\partial T_i}(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell)) \frac{\partial Q}{\partial T_j}(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell)) \kappa^{\mathcal{P}}(g_i, g_j) \\ &\geq \frac{\partial Q}{\partial T_{j_0}}(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell))^2 \kappa^{\mathcal{P}}(g_{j_0}, g_{j_0}) \gg \mathfrak{M}^{2\delta-2} \sigma^{\mathcal{P}}(g_{j_0})^2 \gg \mathfrak{M}^{2\delta-1} \end{aligned} \quad (17)$$

by equation (16). If m is even, then Theorem 2.7 (with $\mathcal{A} = \mathcal{A}(x)$ and $\mathcal{P} = \mathcal{P}(z)$) implies

$$\frac{M_m}{XB_Q(\mathcal{P})^m} = C_m + O(\mathfrak{M}^{-1/2}) = C_m + o(1);$$

on the other hand, if m is odd, then Theorem 2.7 implies

$$\frac{M_m}{XB_Q(\mathcal{P})^m} \ll B_Q(\mathcal{P})^{-1} + \mathfrak{M}^{-1/2} \ll \mathfrak{M}^{-2\delta+1} + \mathfrak{M}^{-1/2} = o(1).$$

These estimates establish the limits

$$\lim_{x \rightarrow \infty} \frac{M_m}{\#\mathcal{A} \cdot B_Q(\mathcal{P})^m} = \lim_{x \rightarrow \infty} \frac{M_m}{XB_Q(\mathcal{P})^m} = \begin{cases} C_m, & \text{if } m \text{ is even,} \\ 0, & \text{if } m \text{ is odd,} \end{cases} \quad (18)$$

where M_m is as in Definition 2.5; the first equality is due to the fact that $\#\mathcal{A} \sim X$ as $x \rightarrow \infty$. Note that $M_0 = \#\mathcal{A}$, and so the limit (18) is trivially true when $m = 0$ as well.

It follows from the limits (18), by the method of moments (in a way that is standard in these applications to Erdős–Kac theorems; see [7, Section 7] for more details on this type of deduction), that

$$\lim_{x \rightarrow \infty} \frac{\#\{a \in \mathcal{A}(x) : Q(g_1^{\mathcal{P}}(a), \dots, g_\ell^{\mathcal{P}}(a)) < A_Q(\mathcal{P}) + uB_Q(\mathcal{P})\}}{\#\mathcal{A}(x)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt, \quad (19)$$

Note that this deduction is not exactly the same as the conclusion of the theorem we are proving: we would rather that the polynomial Q were being evaluated at the original additive functions $g_j(a)$ rather than their truncations $g_j^{\mathcal{P}}(a)$, and that both occurrences of $\mathcal{P} = \mathcal{P}(z)$ on the right-hand side of the inequality were instead $\mathcal{P}(x)$. However, it is easy to

see (essentially by the continuity of the integral as a function of u) that we can make these adjustments provided that:

- (i) $A_Q(\mathcal{P}(x)) - A_Q(\mathcal{P}(z)) = o(B_Q(\mathcal{P}(z)))$;
- (ii) $B_Q(\mathcal{P}(x)) - B_Q(\mathcal{P}(z)) = o(B_Q(\mathcal{P}(z)))$;
- (iii) $Q(g_1(a), \dots, g_\ell(a)) - Q(g_1^{\mathcal{P}(z)}(a), \dots, g_\ell^{\mathcal{P}(z)}(a)) = o(B_Q(\mathcal{P}(z)))$ for every $a \in \mathcal{A}(x)$.

First, note that $A_Q(\mathcal{P}(z)) \asymp \mathfrak{M}^\delta$ by equations (9) and (16). Furthermore, equation (17) gives the lower bound $B_Q(\mathcal{P}(z)) \gg \mathfrak{M}^{\delta-1/2}$, and the corresponding upper bound follows from equation (16) and the estimate $\kappa^{\mathcal{P}}(g_i, g_j) \ll \mathfrak{M}$. Also note that for each $1 \leq j \leq \ell$,

$$0 \leq \mu^{\mathcal{P}(x)}(g_j) - \mu^{\mathcal{P}(z)}(g_j) = \sum_{z < p \leq x} g_j(p) \frac{h(p)}{p} \leq G \sum_{z < p \leq x} \frac{h(p)}{p} = G\eta \log \frac{\log x}{\log z} + O(1) \ll \log \mathfrak{M}$$

by the assumption (14); by a similar calculation, $\kappa^{\mathcal{P}}(g_i, g_j) \ll \log \mathfrak{M}$ for all $1 \leq i, j \leq \ell$. The difference

$$A_Q(\mathcal{P}(x)) - A_Q(\mathcal{P}(z)) = Q(\mu^{\mathcal{P}(x)}(g_1), \dots, \mu^{\mathcal{P}(x)}(g_\ell)) - Q(\mu^{\mathcal{P}(z)}(g_1), \dots, \mu^{\mathcal{P}(z)}(g_\ell))$$

can therefore be bounded, using the multivariable mean value theorem, by $\mathfrak{M}^{\delta-1} \log \mathfrak{M} = o(\mathfrak{M}^{\delta-1/2}) = o(B_Q(\mathcal{P}(z)))$, establishing (i).

An analogous argument shows that the difference $B_Q(\mathcal{P}(x))^2 - B_Q(\mathcal{P}(z))^2$ can be bounded by $\mathfrak{M}^{2\delta-2} \log \mathfrak{M}$; since

$$B_Q(\mathcal{P}(x)) - B_Q(\mathcal{P}(z)) = \frac{B_Q(\mathcal{P}(x))^2 - B_Q(\mathcal{P}(z))^2}{B_Q(\mathcal{P}(x)) + B_Q(\mathcal{P}(z))} \asymp \frac{B_Q(\mathcal{P}(x))^2 - B_Q(\mathcal{P}(z))^2}{\mathfrak{M}^{\delta-1/2}},$$

we see that $B_Q(\mathcal{P}(x)) - B_Q(\mathcal{P}(z)) \ll \mathfrak{M}^{\delta-3/2} \log \mathfrak{M} = o(B_Q(\mathcal{P}(z)))$, establishing (ii).

Finally, for any $a \in \mathcal{A}(x)$ and any $1 \leq j \leq \ell$,

$$g(a) - g^{\mathcal{P}(z)}(a) = \sum_{\substack{z < p \leq x \\ p|a}} g(p)$$

since $a \leq x$; the number of summands in this sum is at most $\frac{\log x}{\log z}$, whence $g^{\mathcal{P}(x)}(a) - g^{\mathcal{P}(z)}(a) \leq G \frac{\log x}{\log z} = o(\mathfrak{M}^{1/2})$ by the assumption (14). Again the multivariable mean value theorem gives the estimate $Q(g_1(a), \dots, g_\ell(a)) - Q(g_1^{\mathcal{P}(z)}(a), \dots, g_\ell^{\mathcal{P}(z)}(a)) \ll \mathfrak{M}^{\delta-1} \cdot o(\mathfrak{M}^{1/2}) = o(B_Q(\mathcal{P}(z)))$, establishing (iii).

These estimates confirm that we can deduce equation (15) from equation (19), which concludes the proof of Theorem 2.8. \square

Theorem 1.3 follows quickly from Theorem 2.8.

Deduction of Theorem 1.3 from Theorem 2.8. Let $\tilde{\mathcal{P}}$ be the set of all primes and $\tilde{\mathcal{A}}$ the set of all positive integers; then $\#\mathcal{A}_d = \frac{x}{d} + O(1)$, and so setting $X = x$ and $h(d) = 1$ for all $d \geq 1$ yields $|E_d| \ll 1$, as well as $\mu^{\mathcal{P}(t)}(1) = \sum_{p \leq t} \frac{1}{p} \sim \log \log t$. Let the polynomial $Q(T_1, \dots, T_\ell)$ and the additive functions g_1, \dots, g_ℓ satisfy the hypotheses of Theorem 1.3; note that one of those hypotheses, namely $\kappa(g_i, g_i) \gg \mu(g_j)$ (with the implicit parameter x replaced by z), implies that $\sigma^{\mathcal{P}(z)}(g_j)^2 \gg \mathfrak{M}$.

Let $z = z(x)$ be the smallest positive real number satisfying $z^{\max\{1, \mathfrak{M}^{1/3}\}} \geq x$, with \mathfrak{M} as in Definition 2.3 with \mathcal{P} replaced by $\mathcal{P}(z)$. Since $\mathfrak{M}^{1/3}$ is an increasing, right-continuous

function of z , this minimum is well-defined and tends to infinity with x . Also, $z \ll_\varepsilon x^\varepsilon$ for any $\varepsilon > 0$ since \mathfrak{M} tends to infinity as well (again because $\sum_{p \leq z} \frac{g_j(p)}{p} \geq \frac{1}{G} \sum_{p \leq z} \frac{g_j(p)^2}{p}$, with the latter series diverging by assumption as $z \rightarrow \infty$). We verify quickly that the hypothesis (13) is satisfied for all $k \geq 0$ via the calculation

$$\begin{aligned} \mu^{\mathcal{P}(z)}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P}(z))} |E_d| &\ll (\log \log x)^k \#\mathcal{D}_k(\mathcal{P}(z)) \\ &\leq (\log \log x)^k z^k \ll_\varepsilon x^\varepsilon \ll_\varepsilon \frac{X}{\log \log x} \ll X \mathfrak{R}^{k/2-1}, \end{aligned}$$

since each $\kappa^{\mathcal{P}(z)}(g_i, g_j) \leq \kappa^{\mathcal{P}(x)}(g_i, g_j) \leq \sum_{p \leq x} \frac{G^2}{p} (1 - \frac{1}{p}) \ll \log \log x$. We also confirm that $\frac{\log x}{\log z} = \max\{1, \mathfrak{M}^{1/3}\} = o(\mathfrak{M}^{1/2})$, which completes the verification of the hypotheses of Theorem 2.8.

Therefore the conclusion of Theorem 2.8 holds, and it is easy to see that this is identical to the conclusion of Theorem 1.3 since $\mu^{\mathcal{P}(x)}(g_j) = \mu(g_j)$ and $\kappa^{\mathcal{P}(x)}(g_i, g_j) = \kappa(g_i, g_j)$. \square

Theorem 1.6 follows in almost exactly the same way, other than requiring a more powerful tool in the Bombieri–Vinogradov theorem to control the accumulation of the sieve error terms. See for example [6, Theorem 17.1] for the statement of the Bombieri–Vinogradov theorem for the function $\psi(x; q, a)$, from which it is simple to derive the analogous version for $\pi(x; q, a)$ (an example of such a derivation is the proof of [1, Corollary 1.4]):

Proposition 2.9. *For any positive real number A , there exists a positive real number $B = B(A)$ such that for all $x \geq 2$,*

$$\sum_{2 \leq q \leq x^{1/2} (\log x)^{-B}} \max_{(a, q)=1} \left| \pi(x; q, a) - \frac{\text{li}(x)}{\varphi(q)} \right| \ll_A \frac{x}{(\log x)^A}, \quad (20)$$

where $\text{li}(x) = \int_2^x \frac{dt}{\log t}$ is the usual logarithmic integral.

Deduction of Theorem 1.6 from Theorem 2.8. Let $\tilde{\mathcal{P}}$ be the set of all prime numbers, and set $\tilde{\mathcal{A}} = \{p - a : p \text{ prime}, p > a\}$. Define a (strongly) multiplicative function $h(d)$ by setting $h(p) = \frac{p}{p-1}$ if $p \nmid a$ and $h(p) = 0$ if $p \mid a$. Then, with $X = \text{li}(x) \sim \pi(x)$, we have

$$\#\mathcal{A}_d = \pi(x; d, a) = \frac{h(d)}{d} \text{li}(x) + E_d(x)$$

where, if $(a, d) = 1$, then $E_d(x) = \pi(x; d, a) - \frac{\text{li}(x)}{\varphi(d)}$ is the error term in the prime number theorem for arithmetic progressions, while if $(a, d) > 1$, then $|E_d(x)| \leq 1$. Therefore, for any parameter $z = x^{o(1)}$,

$$\mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d(x)| \leq \left(\sum_{p \leq z} \frac{1}{p-1} \right)^k \sum_{d \leq z^k} |E_d(x)| \ll (\log \log x)^k \frac{x}{(\log x)^2} \ll X \mathfrak{R}^{k/2-1}$$

by the Bombieri–Vinogradov bound (20). The rest of the proof is the same as the deduction of Theorem 1.3 from Theorem 2.8 above. \square

From now on, we concern ourselves entirely with the proof of Theorem 2.7; in particular, we assume for the rest of this paper that the finite sets \mathcal{A} and \mathcal{P} and the additive functions g_1, \dots, g_ℓ satisfy all the hypotheses of Theorem 2.7.

3. POLYNOMIAL ACCOUNTING

We attack the m th moment M_m by expanding the m th power in each summand. In this section we present the system we use for writing down this expansion, starting by defining a few helpful objects and pieces of notation. For any given $a \in \mathcal{A}$, define $f_r(a)$ to be the completely multiplicative function of r (not of a) that satisfies

$$f_p(a) = \begin{cases} 1 - h(p)/p, & \text{if } p \mid a, \\ -h(p)/p, & \text{if } p \nmid a. \end{cases} \quad (21)$$

(We note that this function was first introduced by Granville and Soundararajan in [4], where it served largely the same role as it will for us.) For any strongly additive function g , define the notation

$$F_g^{\mathcal{P}}(a) = \sum_{p \in \mathcal{P}} g(p) f_p(a), \quad (22)$$

and note that equation (8) implies that $g^{\mathcal{P}}(a) = \mu^{\mathcal{P}}(g) + F_g^{\mathcal{P}}(a)$. Therefore, from Definition 2.5 and equation (9),

$$\begin{aligned} M_m &= \sum_{a \in \mathcal{A}} (Q(g_1^{\mathcal{P}}(a), \dots, g_\ell^{\mathcal{P}}(a)) - A_Q(\mathcal{P}))^m \\ &= \sum_{a \in \mathcal{A}} (Q(F_{g_1}^{\mathcal{P}}(a) + \mu^{\mathcal{P}}(g_1), \dots, F_{g_\ell}^{\mathcal{P}}(a) + \mu^{\mathcal{P}}(g_\ell)) - Q(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell)))^m. \end{aligned} \quad (23)$$

We are now ready to expand the m th power. In addition to setting out the necessary notation for writing down this expansion, this section establishes a key result, Proposition 3.3, which allows us to simplify the main term of the moments M_m into the form of Theorem 2.7. We note that this part of our proof is extremely similar to [7, Section 4]; indeed, we need only quote two relevant definitions and two relevant results from that section, beginning with [7, Definition 4.1]:

Definition 3.1. For any positive even integer k , define T_k to be the set of all 2-to-1 functions from $\{1, \dots, k\}$ to $\{1, \dots, k/2\}$. A typical element of T_k will be denoted by τ . For $\tau \in T_k$ and $j \in \{1, \dots, k/2\}$, define $\Upsilon_1(j)$ and $\Upsilon_2(j)$ to be the two preimages in $\{1, \dots, k\}$; we will never be in a situation where we need to distinguish them from each other.

We also quote [7, Definition 4.5]:

Definition 3.2. Given a polynomial $Q(y_1, \dots, y_\ell) \in \mathbb{R}[y_1, \dots, y_\ell]$ of degree δ , and a positive integer m , define a polynomial in the 2ℓ variables $x_1, \dots, x_\ell, y_1, \dots, y_\ell$ by

$$R_m(x_1, \dots, x_\ell, y_1, \dots, y_\ell) = (Q(x_1 + y_1, \dots, x_\ell + y_\ell) - Q(y_1, \dots, y_\ell))^m.$$

To expand this out in gruesome detail, R_m can be written as the sum of B_m monomials, the β th of which will have total x -degree equal to $k_{m\beta}$ and total y -degree equal to $\tilde{k}_{m\beta}$:

$$R_m(x_1, \dots, x_\ell, y_1, \dots, y_\ell) = \sum_{\beta=1}^{B_m} r_{m\beta} \prod_{i=1}^{k_{m\beta}} x_{v(m,\beta,i)} \prod_{j=1}^{\tilde{k}_{m\beta}} y_{w(m,\beta,j)}. \quad (24)$$

Here each $v(m, \beta, i)$ and $w(m, \beta, j)$ is an integer in $\{1, \dots, \ell\}$. Note that the total x -degree of the β th monomial in the sum is $k_{m\beta}$, while its total y -degree is $\tilde{k}_{m\beta}$. We note that, in

particular, each $k_{m\beta}$ is at least m . Note also that each $k_{m\beta} + \tilde{k}_{m\beta}$ is at most δm , and that there exists at least one β for which $k_{m\beta} = \tilde{k}_{m\beta} = m$.

Remark. Comparing Definition 3.2 with the expression (23), in our calculation we will take $x_i = F_{g_i}^{\mathcal{P}}(a)$ and $y_i = \mu^{\mathcal{P}}(g_i)$, for each $1 \leq i \leq \ell$. Thus the right-hand side of equation (24) becomes a sum of products of $\mu^{\mathcal{P}}(g_i)$ and $F_{g_j}^{\mathcal{P}}(a)$ for various choices of $1 \leq i, j \leq \ell$.

Note that the monomial in R_m of smallest x -degree has x -degree equal to m . It turns out that these terms contribute to the main term of the moments M_m . The following proposition will allow us to prove this fact; it appears as [7, Proposition 4.7], and its proof can be found there. We use $Q_j(y_1, \dots, y_\ell)$ to denote the partial derivative of $Q(y_1, \dots, y_\ell)$ with respect to the j th variable.

Proposition 3.3. *Let m be a positive even integer, and let z_{ij} ($1 \leq i, j \leq \ell$) be real numbers. In the notation of Definitions 3.1 and 3.2,*

$$\begin{aligned} \frac{1}{(m/2)!} \sum_{\substack{\beta \leq B_m \\ k_{m\beta} = m}} r_{m\beta} \prod_{j=1}^{\tilde{k}_{m\beta}} y_{w(m,\beta,j)} \sum_{\tau \in T_m} \prod_{i=1}^{m/2} z_{v(m,\beta,\Upsilon_1(i))v(m,\beta,\Upsilon_2(i))} \\ = C_m \left(\sum_{i=1}^{\ell} \sum_{j=1}^{\ell} Q_i(y_1, \dots, y_\ell) Q_j(y_1, \dots, y_\ell) z_{ij} \right)^{m/2}, \end{aligned} \quad (25)$$

where C_m is as in Definition 2.6.

4. PRELIMINARY CALCULATIONS

As we have seen, expanding the m th power in M_m will result in a sum of terms, each of which takes the shape of a product of means $\mu^{\mathcal{P}}(g_i)$ and functions $F_{g_j}^{\mathcal{P}}(a)$ for various choices $1 \leq i, j \leq \ell$. In this section, we obtain formulas for these products. When m is even, those terms which contribute to the main term of the m th moment will involve a product of $m/2$ of the functions $F_{g_i}^{\mathcal{P}}(a)$. Proposition 4.8 allows us to write such products in terms of covariances $\kappa(g_i, g_j)$, thereby (after some rearranging) recovering, via Proposition 3.3, the main term stated in Theorem 2.7. We also obtain upper bounds for those terms not contributing to this main term. Our method here is deeply inspired by the work of Granville and Soundararajan in [4].

We begin by defining two new functions which serve important technical roles (these functions were called $G(n)$ and $E(r, s)$ in [4], but we have renamed them to avoid a clash of notation herein).

Definition 4.1. Let $H(n)$ be the multiplicative function defined by

$$H(n) = \prod_{p^\alpha \parallel n} \left(\frac{h(p)}{p} \left(1 - \frac{h(p)}{p} \right)^\alpha + \left(-\frac{h(p)}{p} \right)^\alpha \left(1 - \frac{h(p)}{p} \right) \right).$$

For a given natural number s , let $J(r, s)$ be the multiplicative function of r defined by

$$J(r, s) = \prod_{\substack{p^\alpha \parallel r \\ p \mid s}} \left(\left(1 - \frac{h(p)}{p} \right)^\alpha - \left(-\frac{h(p)}{p} \right)^\alpha \right) \prod_{\substack{p^\alpha \parallel r \\ p \nmid s}} \left(-\frac{h(p)}{p} \right)^\alpha.$$

The following two lemmas list some properties of these two functions that will be useful in later proofs.

Lemma 4.2. *For all primes p and all positive integers m, s , and α :*

- (a) $H(p) = 0$, and $H(n) = 0$ unless n is squarefull.
- (b) $|H(p^\alpha)| \leq H(p^2)$.
- (c) $|J(p^\alpha, s)| \leq 1$; furthermore, if $p \nmid s$ then $|J(p^\alpha, s)| \leq h(p)/p$.

Proof. Part (a) is obvious. As for part (b), the triangle inequality yields

$$|H(p^\alpha)| \leq \frac{h(p)}{p} \left(1 - \frac{h(p)}{p}\right)^\alpha + \left(\frac{h(p)}{p}\right)^\alpha \left(1 - \frac{h(p)}{p}\right).$$

Since $0 \leq h(p)/p \leq 1$, both terms on the right-hand side are decreasing functions of α , and so

$$|H(p^\alpha)| \leq \frac{h(p)}{p} \left(1 - \frac{h(p)}{p}\right)^2 + \left(\frac{h(p)}{p}\right)^2 \left(1 - \frac{h(p)}{p}\right) = \frac{h(p)}{p} \left(1 - \frac{h(p)}{p}\right) = H(p^2). \quad (26)$$

Finally, we consider

$$J(p^\alpha, s) = \begin{cases} (1 - h(p)/p)^\alpha - (-h(p)/p)^\alpha, & \text{if } p \mid s, \\ (-h(p)/p)^\alpha, & \text{if } p \nmid s. \end{cases}$$

As $h(p) \leq p$, the second assertion of part (c) is immediate; and, by the triangle inequality, if $p \mid s$ then

$$|J(p^\alpha, s)| \leq \left|1 - \frac{h(p)}{p}\right|^\alpha + \left|\frac{h(p)}{p}\right|^\alpha.$$

The right-hand side is a decreasing function of $\alpha \geq 1$ which takes the value 1 if $\alpha = 1$, which establishes the last remaining assertion. \square

Lemma 4.3. *Let r be a positive integer and let R be the largest squarefree divisor of r . Then for any divisor s of R ,*

$$\sum_{d \mid R} f_r(d) \frac{h(d)}{d} \prod_{p \mid R/d} \left(1 - \frac{h(p)}{p}\right) = H(r) \quad \text{and} \quad \sum_{de=s} f_r(d) \mu(e) = J(r, s).$$

Proof. For each prime factor p of r , let α_p denote the exponent of p in the factorization of r , so that $p^{\alpha_p} \parallel r$. We start with the first claimed equation. By Definition 4.1,

$$H(r) = \prod_{p \mid R} \left(\frac{h(p)}{p} \left(1 - \frac{h(p)}{p}\right)^{\alpha_p} + \left(-\frac{h(p)}{p}\right)^{\alpha_p} \left(1 - \frac{h(p)}{p}\right) \right).$$

Expanding this product results in a sum where each summand is a product over all the prime divisors of R . To each summand we assign a squarefree divisor $d \mid R$ by setting d equal to the product of those primes contributing a factor of the form $(h(p)/p)(1 - h(p)/p)^{\alpha_p}$ to the

summand; this accounts for all such divisors d of R . Therefore,

$$\begin{aligned} H(r) &= \sum_{d|R} \left(\prod_{p|d} \frac{h(p)}{p} \left(1 - \frac{h(p)}{p}\right)^{\alpha_p} \right) \left(\prod_{p|R/d} \left(-\frac{h(p)}{p}\right)^{\alpha_p} \left(1 - \frac{h(p)}{p}\right) \right) \\ &= \sum_{d|R} \frac{h(d)}{d} \left(\prod_{p|d} \left(1 - \frac{h(p)}{p}\right)^{\alpha_p} \prod_{p|R/d} \left(-\frac{h(p)}{p}\right)^{\alpha_p} \right) \prod_{p|R/d} \left(1 - \frac{h(p)}{p}\right) \\ &= \sum_{d|R} \frac{h(d)}{d} f_r(d) \prod_{p|R/d} \left(1 - \frac{h(p)}{p}\right) \end{aligned}$$

as claimed, where the last equality uses the definition (21) of $f_r(d)$.

Now, suppose that s is a divisor of R (hence itself squarefree). From Definition 4.1,

$$J(r, s) = \prod_{p|R/s} \left(-\frac{h(p)}{p}\right)^{\alpha_p} \prod_{p|s} \left(\left(1 - \frac{h(p)}{p}\right)^{\alpha_p} - \left(-\frac{h(p)}{p}\right)^{\alpha_p} \right).$$

Expanding the second product using the same method as above, we obtain

$$\begin{aligned} J(r, s) &= \prod_{p|R/s} \left(-\frac{h(p)}{p}\right)^{\alpha_p} \sum_{d|s} \left(\prod_{p|d} \left(1 - \frac{h(p)}{p}\right)^{\alpha_p} \right) \left(\prod_{p|s/d} -\left(-\frac{h(p)}{p}\right)^{\alpha_p} \right) \\ &= \sum_{d|s} \left(\prod_{p|d} \left(1 - \frac{h(p)}{p}\right)^{\alpha_p} \right) \left(\prod_{p|R/d} \left(-\frac{h(p)}{p}\right)^{\alpha_p} \right) \left(\prod_{p|s/d} -1 \right) \\ &= \sum_{d|s} f_r(d) \mu(s/d) \end{aligned}$$

(using equation (21) again), which is equivalent to the second assertion of the lemma. \square

We now present several lemmas which will aid in the proof of Proposition 4.8, which is the main result of this section. Proposition 4.8 handles expressions of the form

$$\sum_{a \in \mathcal{A}} \prod_{j=1}^k F_{g_j}^{\mathcal{P}}(a),$$

providing an asymptotic formula when k is even and an upper bound when k is odd. Lemmas 4.4 through 4.6 serve to streamline the proof of Proposition 4.8.

In the following formulas and proofs, all implied constants may depend on the positive integer k , which in practice will satisfy $k \leq \delta m$ (recall that δ is the degree of the polynomial Q), in addition to the finite sets \mathcal{A} and \mathcal{P} and the constant G .

Remark. We note that the lemmas below are nearly identical to those in [7, Section 5], with slightly different notation; for example, the expression $\text{cov}(g_1, g_2)$ in [7] is replaced by the slightly more general $\kappa(g_1, g_2)$ here, which is defined in nearly the same way, save for the multiplicative factor $h(p)$ measuring the local densities in our sieve setup. Similarly, the function $H(n)$ in this manuscript is nearly the same as $H(n)$ in [7], save for the presence of this same $h(p)$. However, we give the proofs of these formulas here, as the situation in [7] (involving one additive function with much larger mean than the others) required somewhat more complicated handling.

Lemma 4.4. *Let k be a positive even integer and let g_1, \dots, g_k be any strongly additive functions. Then*

$$\begin{aligned} & \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} = k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \\ &= \frac{1}{(k/2)!} \sum_{\tau \in T_k} \prod_{j=1}^{k/2} \kappa^{\mathcal{P}}(g_{\Upsilon_1(j)}, g_{\Upsilon_2(j)}) + O(\mathfrak{K}^{k/2-1}), \end{aligned} \quad (27)$$

where T_k , Υ_1 , and Υ_2 are as in Definition 3.1 and $\kappa^{\mathcal{P}}(g_i, g_j)$ is the covariance of g_i and g_j as defined in Section 2.

Proof. To each k -tuple (p_1, \dots, p_k) counted by the sum on the left-hand side, we can uniquely associate a $(k/2)$ -tuple $(q_1, \dots, q_{k/2})$ of primes satisfying $q_1 < \dots < q_{k/2}$ and each q_j equals exactly two p_i (at least two because $p_1 \cdots p_k$ is squarefull, and then exactly two because $\#\{p_1, \dots, p_k\} = k/2$). This defines a unique $\tau \in T_k$, namely $\tau(i)$ equals the integer j such that $p_i = q_j$. Therefore

$$\begin{aligned} & \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} = k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \\ &= \sum_{\tau \in T_k} \sum_{\substack{q_1 < \dots < q_{k/2} \\ q_j \in \mathcal{P}}} H(q_1^2 \cdots q_{k/2}^2) g_1(q_{\tau(1)}) \cdots g_k(q_{\tau(k)}) \\ &= \frac{1}{(k/2)!} \sum_{\tau \in T_k} \sum_{\substack{q_1, \dots, q_{k/2} \in \mathcal{P} \\ q_1, \dots, q_{k/2} \text{ distinct}}} H(q_1^2 \cdots q_{k/2}^2) g_1(q_{\tau(1)}) \cdots g_k(q_{\tau(k)}). \end{aligned} \quad (28)$$

By the multiplicativity of H and its values (26) on squares of primes, we see that

$$\begin{aligned} & \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} = k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \\ &= \frac{1}{(k/2)!} \sum_{\tau \in T_k} \sum_{\substack{q_1, \dots, q_{k/2} \in \mathcal{P} \\ q_1, \dots, q_{k/2} \text{ distinct}}} \left(\prod_{j=1}^{k/2} \frac{h(q_j)}{q_j} \left(1 - \frac{h(q_j)}{q_j} \right) \right) g_1(q_{\tau(1)}) \cdots g_k(q_{\tau(k)}) \\ &= \frac{1}{(k/2)!} \sum_{\tau \in T_k} \sum_{\substack{q_1, \dots, q_{k/2} \in \mathcal{P} \\ q_1, \dots, q_{k/2} \text{ distinct}}} \prod_{j=1}^{k/2} g_{\Upsilon_1(j)}(q_j) g_{\Upsilon_2(j)}(q_j) \frac{h(q_j)}{q_j} \left(1 - \frac{h(q_j)}{q_j} \right). \end{aligned} \quad (29)$$

If we fix τ and $q_1, \dots, q_{k/2-1}$, the innermost sum over $q_{k/2}$ is

$$\begin{aligned} \sum_{\substack{q_{k/2} \in \mathcal{P} \\ q_{k/2} \notin \{q_1, \dots, q_{k/2-1}\}}} g_{\Upsilon_1(k/2)}(q_{k/2}) g_{\Upsilon_2(k/2)}(q_{k/2}) \frac{h(q_{k/2})}{q_{k/2}} \left(1 - \frac{h(q_{k/2})}{q_{k/2}}\right) \\ = \kappa^{\mathcal{P}}(g_{\Upsilon_1(k/2)}, g_{\Upsilon_2(k/2)}) + O(1). \end{aligned}$$

Summing in turn over $q_{k/2-1}, \dots, q_1$ in the same way, equation (29) becomes

$$\begin{aligned} \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} = k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \\ = \frac{1}{(k/2)!} \sum_{\tau \in T_k} \prod_{j=1}^{k/2} (\kappa^{\mathcal{P}}(g_{\Upsilon_1(j)}, g_{\Upsilon_2(j)}) + O(1)) \\ = \frac{1}{(k/2)!} \sum_{\tau \in T_k} \left(\prod_{j=1}^{k/2} \kappa^{\mathcal{P}}(g_{\Upsilon_1(j)}, g_{\Upsilon_2(j)}) + O(\mathfrak{K}^{k/2-1}) \right) \end{aligned} \quad (30)$$

as desired, since each $\kappa^{\mathcal{P}}(g_{\Upsilon_1(j)}, g_{\Upsilon_2(j)}) \ll \mathfrak{K}$ by Definition 2.3. \square

Lemma 4.5. *Let k be an integer, and let g_1, \dots, g_k be strongly additive functions satisfying $0 \leq g_j(p) \leq G$ for all primes p and all $1 \leq j \leq k$. When k is even,*

$$\sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} < k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \ll \mathfrak{K}^{k/2-1},$$

while when k is odd,

$$\sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} < k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \ll \mathfrak{K}^{(k-1)/2}.$$

Proof. To each k -tuple (p_1, \dots, p_k) counted by the sum, we associate the positive integer $s = \#\{p_1, \dots, p_k\}$, the primes $q_1 < \dots < q_s$ in \mathcal{P} such that $\{q_1, \dots, q_s\} = \{p_1, \dots, p_k\}$, and the integers $\alpha_1, \dots, \alpha_s \geq 2$ such that q_j equals exactly α_j of the p_i . Let T_α denote the set of functions from $\{1, \dots, k\}$ to $\{1, \dots, s\}$ such that for each $1 \leq j \leq s$, exactly α_j elements of $\{1, \dots, k\}$ are mapped to j . Define $\Upsilon_1(j)$ and $\Upsilon_2(j)$ to be two of the preimages in $\{1, \dots, k\}$; we will never put ourselves in a situation where we need to know which two. We then have

$$\begin{aligned} \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} < k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \\ = \sum_{1 \leq s < k/2} \sum_{\substack{\alpha_1, \dots, \alpha_s \geq 2 \\ \alpha_1 + \dots + \alpha_s = k}} \sum_{\tau \in T_\alpha} \sum_{\substack{q_1 < \dots < q_s \\ q_j \in \mathcal{P}}} H(q_1^{\alpha_1} \cdots q_s^{\alpha_s}) g_1(q_{\tau(1)}) \cdots g_k(q_{\tau(k)}). \end{aligned}$$

By Lemma 4.2(b), $|H(q_1^{\alpha_1} \cdots q_s^{\alpha_s})| \leq H(q_1^2 \cdots q_s^2)$; note also that

$$\sum_{\tau \in T_\alpha} g_1(q_{\tau(1)}) \cdots g_k(q_{\tau(k)}) \leq G^{k-2s} \prod_{j=1}^s g_{\Upsilon_1(j)}(q_j) g_{\Upsilon_2(j)}(q_j).$$

Therefore

$$\begin{aligned} & \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} < k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \\ & \ll \sum_{1 \leq s < k/2} \sum_{\substack{\alpha_1, \dots, \alpha_s \geq 2 \\ \alpha_1 + \cdots + \alpha_s = k}} \sum_{\tau \in T_\alpha} \sum_{\substack{q_1 < \cdots < q_{k/2} \\ q_j \in \mathcal{P}}} H(q_1^2 \cdots q_s^2) g_{\Upsilon_1(j)}(q_j) g_{\Upsilon_2(j)}(q_j). \end{aligned}$$

The innermost double sum can be evaluated just as in equations (28) and (30), yielding

$$\begin{aligned} & \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} < k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \\ & \ll \sum_{1 \leq s < k/2} \sum_{\substack{\alpha_1, \dots, \alpha_s \geq 2 \\ \alpha_1 + \cdots + \alpha_s = k}} \frac{1}{s!} \sum_{\tau \in T_\alpha} \prod_{j=1}^s (\kappa^{\mathcal{P}}(g_{\Upsilon_1(j)}, g_{\Upsilon_2(j)}) + O(1)) \\ & \ll \sum_{1 \leq s < k/2} \sum_{\substack{\alpha_1, \dots, \alpha_s \geq 2 \\ \alpha_1 + \cdots + \alpha_s = k}} \frac{1}{s!} \sum_{\tau \in T_\alpha} \mathfrak{K}^s \ll \mathfrak{K}^{\max\{s \in \mathbb{N} : s < k/2\}} \end{aligned}$$

as desired, since the implicit constant is allowed to depend upon k . □

Lemma 4.6. *For any positive integer r , we have*

$$\sum_{a \in \mathcal{A}} f_r(a) = H(r)X + \sum_{s|r} \mu^2(s) J(r, s) E_s,$$

where E_s is defined by equation (7).

Proof. This is [4, equation (13)]; for completeness, we give the full argument here. Let R denote the largest squarefree divisor of r , and note from equation (21) that $f_r(a) = f_r(d)$ whenever $(a, R) = d$. Then

$$\begin{aligned} \sum_{a \in \mathcal{A}} f_r(a) &= \sum_{d|R} f_r(d) \sum_{\substack{a \in \mathcal{A} \\ (a, R) = d}} 1 = \sum_{d|R} f_r(d) \sum_{\substack{a \in \mathcal{A} \\ d|a}} \sum_{e|(a/d, R/d)} \mu(e) \\ &= \sum_{d|R} f_r(d) \sum_{e|R/d} \mu(e) \sum_{\substack{a \in \mathcal{A} \\ de|a}} 1 \\ &= \sum_{d|R} f_r(d) \sum_{e|R/d} \mu(e) \left(\frac{h(de)}{de} X + E_{de} \right) \end{aligned}$$

by equation (7). Since R is squarefree, we have $h(de) = h(d)h(e)$, and therefore

$$\begin{aligned} \sum_{a \in \mathcal{A}} f_r(a) &= X \sum_{d|R} f_r(d) \frac{h(d)}{d} \sum_{e|R/d} \frac{\mu(e)h(e)}{e} + \sum_{d|R} f_r(d) \sum_{e|R/d} \mu(e) E_{de} \\ &= X \sum_{d|R} f_r(d) \frac{h(d)}{d} \prod_{p|R/d} \left(1 - \frac{h(p)}{p}\right) + \sum_{s|R} E_s \sum_{de=s} f_r(d) \mu(e) \\ &= H(r)X + \sum_{s|R} E_s J(r, s) \end{aligned}$$

by Lemma 4.3, which is equivalent to the assertion of the lemma. \square

Lemma 4.7. *Let k be a positive integer, and let g_1, \dots, g_k be strongly additive functions satisfying $0 \leq g_j(p) \leq G$ for all primes p and all $1 \leq j \leq k$. Then*

$$\sum_{p_1, \dots, p_k \in \mathcal{P}} g_1(p_1) \cdots g_k(p_k) \sum_{s|p_1 \cdots p_k} \mu^2(s) J(p_1 \cdots p_k, s) E_s \ll \mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d|,$$

where E_s and E_d are defined in equation (7).

Proof. By Lemma 4.2(c),

$$|J(p_1 \cdots p_k, s)| \leq \prod_{\substack{1 \leq i \leq k \\ p_i \nmid s}} \frac{h(p_i)}{p_i}.$$

Since $g_j(p_j) \ll_G 1$ for all j ,

$$\begin{aligned} \sum_{p_1, \dots, p_k \in \mathcal{P}} g_1(p_1) \cdots g_k(p_k) \sum_{s|p_1 \cdots p_k} J(p_1 \cdots p_k, s) E_s \\ \ll \sum_{\ell=0}^k \sum_{\substack{s=q_1 \cdots q_\ell \geq 1 \\ q_1 < \cdots < q_\ell \in \mathcal{P}}} |E_s| \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ s|p_1 \cdots p_k}} \prod_{\substack{1 \leq i \leq k \\ p_i \nmid s}} \frac{h(p_i)}{p_i}. \end{aligned}$$

At the cost of a constant depending only on k , we can arrange the primes p_1, \dots, p_k so that $p_1 = q_1, \dots, p_\ell = q_\ell$. Summing the factor $h(p_i)/p_i$ over the remaining primes p_i for $\ell + 1 \leq i \leq k$ yields $\mu^{\mathcal{P}}(1)^{k-\ell} \ll \mu^{\mathcal{P}}(1)^k$. Each term $|E_s|$ is equal to $|E_d|$ for some integer $d \in \mathcal{D}_k(\mathcal{P})$, and therefore the total error term is $\ll \mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d|$, as desired. \square

With these preliminary calculations out of the way, we may now establish the main proposition of this section, which will feature prominently in the proof of Theorem 2.7 in the next section.

Proposition 4.8. *Let g_1, \dots, g_k be strongly additive functions satisfying $0 \leq g_j(p) \leq G$ for all primes p and all $1 \leq j \leq k$. Let $F_g^{\mathcal{P}}$ be as in equation (22), and let k be a positive integer. If k is even, then*

$$\begin{aligned} \sum_{a \in \mathcal{A}} \prod_{j=1}^k F_{g_j}^{\mathcal{P}}(a) &= \frac{X}{(k/2)!} \sum_{\tau \in T_k} \prod_{j=1}^{k/2} \kappa^{\mathcal{P}}(g_{\tau_1(j)}, g_{\tau_2(j)}) \\ &\quad + O\left(X \mathfrak{K}^{k/2-1} + \mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d|\right). \end{aligned}$$

If k is odd, then

$$\sum_{a \in \mathcal{A}} \prod_{j=1}^k F_{g_j}^{\mathcal{P}}(a) \ll X \mathfrak{K}^{(k-1)/2} + \mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d|.$$

Proof. Expanding out the definition,

$$\sum_{a \in \mathcal{A}} \prod_{j=1}^k F_{g_j}^{\mathcal{P}}(a) = \sum_{a \in \mathcal{A}} \prod_{j=1}^k \sum_{p \in \mathcal{P}} g(p) f_p(a) = \sum_{p_1, \dots, p_k \in \mathcal{P}} g_1(p_1) \cdots g_k(p_k) \sum_{a \in \mathcal{A}} f_{p_1 \cdots p_k}(a).$$

By Lemma 4.6,

$$\begin{aligned} \sum_{a \in \mathcal{A}} \prod_{j=1}^k F_{g_j}^{\mathcal{P}}(a) &= \sum_{p_1, \dots, p_k \in \mathcal{P}} g_1(p_1) \cdots g_k(p_k) \left(H(p_1 \cdots p_k) X + \sum_{s | p_1 \cdots p_k} \mu^2(s) J(p_1 \cdots p_k, s) E_s \right) \\ &= X \sum_{p_1, \dots, p_k \in \mathcal{P}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) + O \left(\mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d| \right), \end{aligned}$$

using Lemma 4.7 to obtain the error term. Since $H(p_1 \cdots p_k)$ vanishes unless $p_1 \cdots p_k$ is squarefull by Lemma 4.2(a), there are at most $k/2$ distinct primes among p_1, \dots, p_k , and so we can write

$$\begin{aligned} \sum_{p_1, \dots, p_k \in \mathcal{P}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) &= \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} = k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k) \\ &\quad + \sum_{\substack{p_1, \dots, p_k \in \mathcal{P} \\ p_1 \cdots p_k \text{ squarefull} \\ \#\{p_1, \dots, p_k\} < k/2}} H(p_1 \cdots p_k) g_1(p_1) \cdots g_k(p_k). \end{aligned}$$

The proposition now follows upon appeals to Lemmas 4.4 and 4.5 (and the observation that the first sum on the right-hand side is empty if k is odd). \square

Corollary 4.9. *Under the hypotheses of Theorem 2.7, for any positive integer k ,*

$$\sum_{a \in \mathcal{A}} \prod_{j=1}^k |F_{g_j}^{\mathcal{P}}(a)| \ll \begin{cases} X \mathfrak{K}^{k/2}, & \text{if } k \text{ is even,} \\ X \mathfrak{K}^{(k-1)/2}, & \text{if } k \text{ is odd.} \end{cases}$$

Proof. First, note that if the estimate (11) holds, then it holds with δm replaced throughout by any integer $k \leq \delta m$. Indeed, under the hypotheses of Theorem 2.7, we have $\mu^{\mathcal{P}}(1) \geq \frac{1}{G} \mu^{\mathcal{P}}(g_j) \asymp \mathfrak{K}$ for any $1 \leq j \leq k$ (as in the deduction of Theorem 2.8). Since $\mathcal{D}_k(\mathcal{P}) \subseteq \mathcal{D}_{\delta m}(\mathcal{P})$, we see that

$$\begin{aligned} \mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d| &\leq \frac{1}{\mu^{\mathcal{P}}(1)^{\delta m - k}} \cdot \mu^{\mathcal{P}}(1)^{\delta m} \sum_{d \in \mathcal{D}_{\delta m}(\mathcal{P})} |E_d| \\ &\ll \frac{1}{\mu^{\mathcal{P}}(1)^{\delta m - k}} \cdot X \mathfrak{K}^{\delta m / 2 - 1} \\ &\ll X \mathfrak{K}^{k - \delta m / 2 - 1} \leq X \mathfrak{K}^{k/2 - 1}. \end{aligned}$$

If k is even, then by Proposition 4.8 and the triangle inequality,

$$\begin{aligned} \sum_{a \in \mathcal{A}} \prod_{j=1}^k |F_{g_j}^{\mathcal{P}}(a)| &\leq \frac{X}{(k/2)!} \sum_{\tau \in T_k} \prod_{j=1}^{k/2} \mathfrak{K} + O\left(X \mathfrak{K}^{k/2-1} + \mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d|\right) \\ &\ll X \mathfrak{K}^{k/2} + O\left(X \mathfrak{K}^{k/2-1} + \mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d|\right) \ll X \mathfrak{K}^{k/2} \end{aligned}$$

by equation (11). On the other hand, if k is odd, then again by Proposition 4.8 and equation (11),

$$\sum_{a \in \mathcal{A}} \prod_{j=1}^k |F_{g_j}^{\mathcal{P}}(a)| \ll X \mathfrak{K}^{(k-1)/2} + \mu^{\mathcal{P}}(1)^k \sum_{d \in \mathcal{D}_k(\mathcal{P})} |E_d| \ll X \mathfrak{K}^{(k-1)/2},$$

as desired. \square

5. CALCULATING THE MOMENTS

We are finally ready to compute the moments, using the technology built up in prior sections. The reader will observe that the computation ends up being rather brief, thanks to this earlier work. By equation (23) and Definition 3.2,

$$\begin{aligned} M_m &= \sum_{a \in \mathcal{A}} (Q(F_{g_1}^{\mathcal{P}}(a) + \mu^{\mathcal{P}}(g_1), \dots, F_{g_\ell}^{\mathcal{P}}(a) + \mu^{\mathcal{P}}(g_\ell)) - Q(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell)))^m \\ &= \sum_{a \in \mathcal{A}} R_m(F_{g_1}^{\mathcal{P}}(a), \dots, F_{g_\ell}^{\mathcal{P}}(a), \mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_\ell)) \\ &= \sum_{a \in \mathcal{A}} \sum_{\beta=1}^{B_m} r_{m\beta} \prod_{i=1}^{k_{m\beta}} F_{g_{v(m,\beta,i)}}^{\mathcal{P}}(a) \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \\ &= \sum_{\beta=1}^{B_m} r_{m\beta} \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{a \in \mathcal{A}} \prod_{i=1}^{k_{m\beta}} F_{g_{v(m,\beta,i)}}^{\mathcal{P}}(a), \end{aligned} \tag{31}$$

where we used equation (24) in the middle equality.

We analyze the expression (31) to obtain Theorem 2.7, splitting our work into two cases depending on whether m is even or odd. We remind the reader that all implied constants may depend on the polynomial Q , the constant G from the hypotheses of Theorem 2.7, the positive integer m , and the finite set \mathcal{A} .

Proof of Theorem 2.7 when m is even. Our goal is to show that

$$M_m = C_m X B_Q(\mathcal{P})^m + O(X \mathfrak{M}^{\delta m - (m+1)/2}).$$

We first isolate the terms in the expression (31) with $k_{h\beta} = m$, the minimum possible (see the discussion following Definition 3.2), with the goal of showing that these terms contribute

the main term of M_m . By Proposition 4.8, we have for these terms

$$\begin{aligned}
& \prod_{i=1}^{k_{m,\beta}} F^{\mathcal{P}}_{g_{v(m,\beta,i)}}(a) \\
&= \frac{X}{(m/2)!} \sum_{\tau \in T_m} \prod_{i=1}^{m/2} \kappa^{\mathcal{P}}(g_{v(m,\beta,\Upsilon_1(i))}, g_{v(m,\beta,\Upsilon_2(i))}) + O\left(X\mathfrak{K}^{m/2-1} + \mu^{\mathcal{P}}(1)^m \sum_{d \in \mathcal{D}_m(\mathcal{P})} |E_d|\right) \\
&= \frac{X}{(m/2)!} \sum_{\tau \in T_m} \prod_{i=1}^{m/2} \kappa^{\mathcal{P}}(g_{v(m,\beta,\Upsilon_1(i))}, g_{v(m,\beta,\Upsilon_2(i))}) + O(X\mathfrak{K}^{m/2-1})
\end{aligned}$$

since we assume the error bound (11) (which, as discussed in the proof of Corollary 4.9, holds for the smaller integer m since it is assumed to hold for δm). Therefore

$$\begin{aligned}
& \sum_{\substack{\beta \leq B_m \\ k_{m,\beta} = m}} r_{m,\beta} \prod_{j=1}^{\tilde{k}_{m,\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{a \in \mathcal{A}} \prod_{i=1}^{k_{m,\beta}} F^{\mathcal{P}}_{g_{v(m,\beta,i)}}(a) \\
&= \sum_{\substack{\beta \leq B_m \\ k_{h,\beta} = m}} r_{m,\beta} \prod_{j=1}^{\tilde{k}_{m,\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \\
&\quad \times \left(\frac{X}{(m/2)!} \sum_{\tau \in T_m} \prod_{i=1}^{m/2} \kappa^{\mathcal{P}}(g_{v(m,\beta,\Upsilon_1(i))}, g_{v(m,\beta,\Upsilon_2(i))}) + O(X\mathfrak{K}^{m/2-1}) \right) \\
&= \frac{X}{(m/2)!} \sum_{\substack{\beta \leq B_m \\ k_{m,\beta} = m}} r_{m,\beta} \prod_{j=1}^{\tilde{k}_{m,\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{\tau \in T_m} \prod_{i=1}^{m/2} \kappa^{\mathcal{P}}(g_{v(m,\beta,\Upsilon_1(i))}, g_{v(m,\beta,\Upsilon_2(i))}) \\
&\quad + O(X\mathfrak{M}^{(\delta-1)m} \mathfrak{K}^{m/2-1}), \tag{32}
\end{aligned}$$

since $\tilde{k}_{m,\beta} \leq \delta m - k_{m,\beta} = (\delta - 1)m$. By Proposition 3.3, where y_j has been replaced by $\mu^{\mathcal{P}}(g_j)$ and z_{ij} by $\kappa^{\mathcal{P}}(g_i, g_j)$, this main term equals

$$\begin{aligned}
& C_m X \left(\sum_{i=1}^{\ell} \sum_{j=1}^{\ell} Q_i(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_{\ell})) Q_j(\mu^{\mathcal{P}}(g_1), \dots, \mu^{\mathcal{P}}(g_{\ell})) \kappa^{\mathcal{P}}(g_i, g_j) \right)^{m/2} \\
&= C_m X (B_Q(\mathcal{P})^2)^{m/2} \tag{33}
\end{aligned}$$

by equation (10). As for the error term, since we assume $\mathfrak{K} \leq G\mathfrak{M}$,

$$X\mathfrak{M}^{(\delta-1)m} \mathfrak{K}^{m/2-1} \ll X\mathfrak{M}^{\delta m - m/2 - 1}. \tag{34}$$

The remaining terms in the expression (31) have $m < k_{m\beta} \leq \delta m$. For these terms,

$$\begin{aligned}
& \sum_{\substack{\beta \leq B_m \\ k_{m\beta} > m}} r_{m\beta} \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{a \in \mathcal{A}} \prod_{i=1}^{k_{m\beta}} F_{g_{v(m,\beta,i)}}^{\mathcal{P}}(a) \\
& \ll \sum_{\substack{\beta \leq B_m \\ k_{m\beta} > m}} |r_{m\beta}| \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{a \in \mathcal{A}} \prod_{i=1}^{k_{m\beta}} |F_{g_{v(m,\beta,i)}}^{\mathcal{P}}(a)| \\
& \ll \sum_{\substack{\beta \leq B_m \\ k_{m\beta} > m}} |r_{m\beta}| \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) X \mathfrak{K}^{k_{m\beta}/2}
\end{aligned}$$

by Corollary 4.9 (recalling again that equation (11) holds for the integers $k_{m\beta}$ under consideration), giving

$$\sum_{\substack{\beta \leq B_m \\ k_{m\beta} > m}} r_{m\beta} \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{a \in \mathcal{A}} \prod_{i=1}^{k_{m\beta}} F_{g_{v(m,\beta,i)}}^{\mathcal{P}}(a) \ll X \sum_{\substack{\beta \leq B_m \\ k_{m\beta} > m}} |r_{m\beta}| \mathfrak{M}^{\tilde{k}_{m\beta}} \mathfrak{K}^{k_{m\beta}/2}.$$

Since we assume $\mathfrak{K} \leq G\mathfrak{M}$, this estimate becomes

$$\sum_{\substack{\beta \leq B_m \\ k_{m\beta} > m}} r_{m\beta} \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{a \in \mathcal{A}} \prod_{i=1}^{k_{m\beta}} F_{g_{v(m,\beta,i)}}^{\mathcal{P}}(a) \ll X \mathfrak{M}^{\delta m - (m+1)/2} \quad (35)$$

since $\tilde{k}_{m\beta} + \frac{k_{m\beta}}{2} = \tilde{k}_{m\beta} + k_{m\beta} - \frac{k_{m\beta}}{2} \leq \tilde{k}_{m\beta} + k_{m\beta} - \frac{m+1}{2} \leq \delta m - \frac{m+1}{2}$ for all terms in the sum.

Gathering the results of the calculations (32)–(35) into the expression (31) yields

$$M_m = C_m X B_Q(\mathcal{P})^m + O(X \mathfrak{M}^{\delta m - (m+1)/2})$$

as desired. \square

A slight modification of the above proof suffices for odd integers m .

Proof of Theorem 2.7 when m is odd. Now, our goal is to show that $M_m \ll X \mathfrak{M}^{\delta m - (m+1)/2}$. The estimate (35) still holds for those terms in expression (31) with $k_{m\beta} > m$. For the remaining terms with $k_{m\beta} = m$, we use Corollary 4.9 to write

$$\begin{aligned}
& \sum_{\substack{\beta \leq B_m \\ k_{m\beta} = m}} r_{m\beta} \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{a \in \mathcal{A}} \prod_{i=1}^{k_{m\beta}} F_{g_{v(m,\beta,i)}}^{\mathcal{P}}(a) \\
& \ll X \mathfrak{K}^{(m-1)/2} \sum_{\substack{\beta \leq B_m \\ k_{m\beta} = m}} |r_{m\beta}| \prod_{j=1}^{\tilde{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \\
& \ll X \mathfrak{K}^{(m-1)/2} \mathfrak{M}^{\tilde{k}_{m\beta}} \leq X \mathfrak{K}^{(m-1)/2} \mathfrak{M}^{(\delta-1)m}
\end{aligned}$$

as before. Since $\mathfrak{K} \leq G\mathfrak{M}$, we obtain

$$\sum_{\substack{\beta \leq B_m \\ k_{m\beta} = m}} r_{m\beta} \prod_{j=1}^{\bar{k}_{m\beta}} \mu^{\mathcal{P}}(g_{w(m,\beta,j)}) \sum_{a \in \mathcal{A}} \prod_{i=1}^{k_{m\beta}} F_{g_{v(m,\beta,i)}}^{\mathcal{P}}(a) \ll X\mathfrak{M}^{\delta m - (m+1)/2}$$

as desired. \square

6. EXAMPLES

We conclude by providing examples of Theorem 1.3 applied to certain choices of the polynomial Q which may be of independent interest. Example 6.1 (which establishes Corollary 1.4) pertains to powers of strongly additive functions; Example 6.2 (which establishes Corollary 1.5) pertains to products of strongly additive functions; and Example 6.3 pertains to linear combinations of strongly additive functions.

Example 6.1. Let δ be a positive integer and set $Q(T) = T^\delta$, so that $Q'(T) = \delta T^{\delta-1}$. Then, for any strongly additive function $g(n)$ satisfying the hypotheses of Theorem 1.3, the function $g(n)^\delta$ satisfies an Erdős–Kac law with mean $A_Q(x) = \mu(g)^\delta$ and variance

$$B_Q(x)^2 = (Q'(\mu(g)))^2 \kappa(g, g) = \delta^2 \mu(g)^{2\delta-2} \sigma^2(g).$$

For a concrete example, if $g(n) = \omega(n)$, the number-of-distinct-prime-divisors function, then $\omega(n)^\delta$ satisfies an Erdős–Kac law with mean $(\log \log n)^\delta$ and variance $\delta^2 (\log \log n)^{2\delta-2}$; and more generally, the same holds for $Q(g(n))$ for any polynomial $Q(T)$ with nonnegative coefficients and leading term T^δ .

Example 6.2. Let $g_1(n)$ and $g_2(n)$ be strongly additive functions satisfying the hypotheses of Theorem 1.3, and set $Q(T) = T_1 T_2$, so that $\frac{\partial Q}{\partial T_1} = T_2$ and $\frac{\partial Q}{\partial T_2} = T_1$. Then the product $g_1(n)g_2(n)$ satisfies an Erdős–Kac law with mean $A_Q(x) = \mu(g_1)\mu(g_2)$ and variance

$$\begin{aligned} B_Q(x)^2 &= \sum_{i=1}^2 \sum_{j=1}^2 \frac{\partial Q}{\partial T_i}(\mu(g_1), \mu(g_2)) \frac{\partial Q}{\partial T_j}(\mu(g_1), \mu(g_2)) \kappa(g_i, g_j) \\ &= \mu(g_1)^2 \sigma^2(g_2) + 2\mu(g_1)\mu(g_2)\kappa(g_1, g_2) + \mu(g_2)^2 \sigma^2(g_1). \end{aligned}$$

For example, suppose P_1 and P_2 are sets of primes of positive relative densities α and β , such that $P_1 \cap P_2$ has relative density γ in the primes; and set $\omega_j(n) = \#\{p \mid n: p \in P_j\}$ for $j = 1, 2$. Using the definitions (3), it is easy to calculate that

$$\mu(g_1) \sim \sigma^2(g_1) \sim \alpha \log \log x, \quad \mu(g_2) \sim \sigma^2(g_2) \sim \beta \log \log x, \quad \kappa(g_1, g_2) \sim \gamma \log \log x.$$

Therefore $\omega_1(n)\omega_2(n)$ satisfies an Erdős–Kac law with mean $\alpha\beta(\log \log x)^2$ and variance $\alpha\beta(\alpha + 2\gamma + \beta)(\log \log x)^3$. In particular, note that the variance depends not just on the two additive functions g_1 and g_2 individually but also upon the correlations in their values.

More generally, let $g_1(n), \dots, g_\ell(n)$ be strongly additive functions satisfying the hypotheses of Theorem 1.3, and set $Q(T) = T_1 \cdots T_\ell$, so that $\frac{\partial Q}{\partial T_j} = (T_1 \cdots T_\ell)/T_j$ for all $1 \leq j \leq \ell$. Then the function $g_1(n) \cdots g_\ell(n)$ satisfies an Erdős–Kac law with mean $A_Q(x) = \mu(g_1) \cdots \mu(g_\ell)$ and

variance

$$\begin{aligned} B_Q(x)^2 &= \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \frac{\partial Q}{\partial T_i}(\mu(g_1), \dots, \mu(g_{\ell})) \frac{\partial Q}{\partial T_j}(\mu(g_1), \dots, \mu(g_{\ell})) \kappa(g_i, g_j) \\ &= (\mu(g_1) \cdots \mu(g_{\ell}))^2 \sum_{1 \leq i, j \leq \ell} \frac{\kappa(g_i, g_j)}{\mu(g_i) \mu(g_j)}. \end{aligned}$$

Example 6.3. Again let $g_1(n)$ and $g_2(n)$ be strongly additive functions satisfying the hypotheses of Theorem 1.3; this time, set $Q(T_1, T_2) = vT_1^{\delta} + wT_2^{\delta}$ for a positive integer δ and nonnegative real numbers v and w , so that $\frac{\partial Q}{\partial T_1} = \delta v T_1^{\delta-1}$ and $\frac{\partial Q}{\partial T_2} = \delta w T_2^{\delta-1}$. Then the function $vg_1(n)^{\delta} + wg_2(n)^{\delta}$ satisfies an Erdős–Kac law with mean $v\mu(g_1)^{\delta} + w\mu(g_2)^{\delta}$ and variance

$$\begin{aligned} B_Q(x)^2 &= \sum_{i=1}^2 \sum_{j=1}^2 \frac{\partial Q}{\partial T_i}(\mu(g_1), \mu(g_2)) \frac{\partial Q}{\partial T_j}(\mu(g_1), \mu(g_2)) \kappa(g_i, g_j) \\ &= v^2 \delta^2 \mu(g_1)^{2\delta-2} \sigma^2(g_1) + 2vw\delta^2 (\mu(g_1)\mu(g_2))^{\delta-1} \kappa(g_1, g_2) + w^2 \delta^2 \mu(g_2)^{2\delta-2} \sigma^2(g_2). \end{aligned}$$

In particular, when $\delta = 1$, the function $g(n) = vg_1(n) + wg_2(n)$ satisfies an Erdős–Kac law with mean $v\mu(g_1) + w\mu(g_2)$ and variance $v^2\sigma^2(g_1) + 2vw\kappa(g_1, g_2) + w^2\sigma^2(g_2)$. But notice that g itself is an additive function. Using the definitions (3), it is trivial to check that $\mu(g) = v\mu(g_1) + w\mu(g_2)$, and still easy to check that $\sigma^2(g) = v^2\sigma^2(g_1) + 2vw\kappa(g_1, g_2) + w^2\sigma^2(g_2)$. Therefore our theorem is self-consistent in this case.

REFERENCES

1. A. Akbary and K. Hambrook, *A variant of the Bombieri–Vinogradov theorem with explicit constants and applications*, Math. Comp. **84** (2015), no. 294, 1901–1932.
2. K. Alladi, *An Erdős–Kac theorem for integers without large prime factors*, Acta Arith. **49** (1987), no. 1, 81–105. MR 913766
3. P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, American Journal of Mathematics **62** (1940), no. 1/4, 343–352.
4. A. Granville and K. Soundararajan, *Sieving and the Erdős–Kac theorem*, Equidistribution in number theory, an introduction, NATO Sci. Ser. II Math. Phys. Chem., vol. 237, Springer, Dordrecht, 2007, pp. 15–27.
5. H. Halberstam, *On the distribution of additive number-theoretic functions*, J. London Math. Soc. **30** (1955), 43–53.
6. H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.
7. G. Martin and L. Troupe, *The distribution of the number of subgroups of the multiplicative group*, J. Aust. Math. Soc. (to appear).