HIGHER CORRELATIONS OF DIVISOR SUMS RELATED TO PRIMES I: TRIPLE CORRELATIONS

D. A. Goldston[1]
Department of Mathematics and Computer Science, San Jose State University, San Jose, CA 95192
goldston@mathcs.sjsu.edu

C. Y. Yıldırım
Department of Mathematics, Boğaziçi University, Istanbul 34342, Turkey
yalciny@boun.edu.tr

Received: 4/29/02, Revised: 4/18/03, Accepted: 4/18/03, Published: 4/22/03

Abstract

We obtain the triple correlations for a truncated divisor sum related to primes. We also obtain the mixed correlations for this divisor sum when it is summed over the primes, and give some applications to primes in short intervals.

1. Introduction

This is the first in a series of papers concerned with the calculation of higher correlations of short divisor sums that are approximations for the von Mangoldt function $\Lambda(n)$, where $\Lambda(n)$ is defined to be $\log p$ if $n = p^m$, p a prime, m a positive integer, and to be zero otherwise. These higher correlations have applications to the theory of primes which is our motivation for their study. In this first paper we will calculate the pair and triple correlations for

$$\Lambda_R(n) = \sum_{\substack{d|n \\ d\le R}} \mu(d)\log(R/d), \qquad \text{for } n \ge 1, \tag{1.1}$$

and $\Lambda_R(n) = 0$ if $n \le 0$. In later papers in this series we will examine quadruple and higher correlations, and also examine the more delicate divisor sum

$$\lambda_R(n) = \sum_{r\le R} \frac{\mu^2(r)}{\phi(r)} \sum_{\substack{d|r \\ d|n}} d\mu(d), \qquad \text{for } n \ge 1, \tag{1.2}$$

[1]The first author was supported in part by an NSF Grant

and $\lambda_R(n) = 0$ if $n \leq 0$. The correlations we are interested in evaluating are

$$\mathcal{S}_k(N, \boldsymbol{j}, \boldsymbol{a}) = \sum_{n=1}^{N} \Lambda_R(n+j_1)^{a_1} \Lambda_R(n+j_2)^{a_2} \cdots \Lambda_R(n+j_r)^{a_r} \tag{1.3}$$

and

$$\tilde{\mathcal{S}}_k(N, \boldsymbol{j}, \boldsymbol{a}) = \sum_{n=1}^{N} \Lambda_R(n+j_1)^{a_1} \Lambda_R(n+j_2)^{a_2} \cdots \Lambda_R(n+j_{r-1})^{a_{r-1}} \Lambda(n+j_r) \tag{1.4}$$

where $\boldsymbol{j} = (j_1, j_2, \ldots, j_r)$ and $\boldsymbol{a} = (a_1, a_2, \ldots a_r)$, the j_i's are distinct integers, $a_i \geq 1$ and $\sum_{i=1}^{r} a_i = k$. In (1.4) we assume that $r \geq 2$ and take $a_r = 1$. For later convenience we define

$$\tilde{\mathcal{S}}_1(N, \boldsymbol{j}, \boldsymbol{a}) = \sum_{n=1}^{N} \Lambda(n+j_1) \sim N \tag{1.5}$$

with $|j_1| \leq N$ by the prime number theorem. For $k = 1$ and $k = 2$ these correlations have been evaluated before [8] (and for $\lambda_Q(n)$ they have been evaluated in [9]); the results show that Λ_R and λ_R mimic the behavior of Λ, and this is also the case in arithmetic progressions, see [17], [18], [11].

When $k \geq 3$ the procedure for evaluating these correlations is complicated, and it is easy to make mistakes in the calculations. Therefore we have chosen to first treat the triple correlations in detail. The main terms in the theorems can often be obtained in an easier way by evaluating the multiple sums in a different order or with a different decomposition of the initial summands; the method used here was chosen to control the error terms and generalize to higher correlations. Recently we have found a somewhat different method which is preferable for higher values of k. This method will be used in the third paper in this series. We can not compute correlations which contain a factor $\Lambda(n)\Lambda(n+k)$, $k \neq 0$, without knowledge about prime twins. This limits our applications, and further the mixed correlations (1.4) can only be calculated for shorter divisor sums than the pure correlations (1.3) of $\Lambda_R(n)$, which degrades to some extent the results we obtain. When we assume the Elliott-Halberstam conjecture we can eliminate this latter problem and obtain stronger results.

One motivation for the study of the correlations of $\Lambda_R(n)$ or $\lambda_R(n)$ is to provide further information on the moments

$$M_k(N, h, \psi) = \sum_{n=1}^{N} (\psi(n+h) - \psi(n))^k \tag{1.6}$$

where $\psi(x) = \sum_{n \leq x} \Lambda(n)$. We always take $N \to \infty$, and let

$$h \sim \lambda \log N, \tag{1.7}$$

where we will usually be considering the case $\lambda \ll 1$. When h is larger we need to subtract the expected value h in the moments above, which leads to more delicate questions which we will not consider in this paper (see [23]). Gallagher [6] proved that the moments in (1.6) may be computed from the Hardy-Littlewood prime r-tuple conjecture [14]. This conjecture states that for $\boldsymbol{j} = (j_1, j_2, \dots, j_r)$ with the j_i's distinct integers,

$$\psi_{\boldsymbol{j}}(N) = \sum_{n=1}^{N} \Lambda(n+j_1)\Lambda(n+j_2)\cdots\Lambda(n+j_r) \sim \mathfrak{S}(\boldsymbol{j})N \tag{1.8}$$

when $\mathfrak{S}(\boldsymbol{j}) \neq 0$, where

$$\mathfrak{S}(\boldsymbol{j}) = \prod_{p} \left(1 - \frac{1}{p}\right)^{-r} \left(1 - \frac{\nu_p(\boldsymbol{j})}{p}\right) \tag{1.9}$$

and $\nu_p(\boldsymbol{j})$ is the number of distinct residue classes modulo p that the j_i's occupy. If $r = 1$ we see $\mathfrak{S}(\boldsymbol{j}) = 1$, and for $|j_1| \leq N$ equation (1.8) reduces to (1.5), which is the only case where (1.8) has been proved. To compute the moments in (1.6) we have

$$\begin{aligned} M_k(N,h,\psi) &= \sum_{n=1}^{N} \left(\sum_{1\leq m \leq h} \Lambda(n+m) \right)^k \\ &= \sum_{\substack{1\leq m_i \leq h \\ 1\leq i \leq k}} \sum_{n=1}^{N} \Lambda(n+m_1)\Lambda(n+m_2)\cdots\Lambda(n+m_k). \end{aligned}$$

Now suppose that the k numbers $m_1, m_2, \dots, m_k$ take on r distinct values $j_1, j_2, \dots, j_r$ with j_i having multiplicity a_i, so that $\sum_{1\leq i \leq r} a_i = k$. Grouping the terms above, we have that

$$M_k(N,h,\psi) = \sum_{r=1}^{k} \sum_{\substack{a_1,a_2,\dots,a_r \\ a_i \geq 1, \sum a_i = k}} \binom{k}{a_1, a_2, \dots, a_r} \sum_{1\leq j_1 < j_2 < \cdots < j_r \leq h} \psi_k(N, \boldsymbol{j}, \boldsymbol{a}), \tag{1.10}$$

where

$$\psi_k(N,\boldsymbol{j},\boldsymbol{a}) = \sum_{n=1}^{N} \Lambda(n+j_1)^{a_1}\Lambda(n+j_2)^{a_2}\cdots\Lambda(n+j_r)^{a_r} \tag{1.11}$$

and the multinomial coefficient counts the number of different innermost sums that occur. If $n + j_i$ is a prime then $\Lambda(n+j_i)^{a_i} = \Lambda(n+j_i)(\log(n+j_i))^{a_i-1}$, and we easily see that

$$\begin{aligned} \psi_k(N,\boldsymbol{j},\boldsymbol{a}) &= (1+o(1))(\log N)^{k-r} \sum_{n=1}^{N} \Lambda(n+j_1)\Lambda(n+j_2)\cdots\Lambda(n+j_r) + O(N^{\frac{1}{2}+\epsilon}) \\ &= (1+o(1))(\log N)^{k-r} \psi_{\boldsymbol{j}}(N) + O(N^{\frac{1}{2}+\epsilon}). \end{aligned} \tag{1.12}$$

Hence we may apply the conjecture (1.8) assuming it is valid uniformly for $\max_i |j_i| \le h$ and obtain

$$M_k(N,h,\psi) \sim N \sum_{r=1}^{k} (\log N)^{k-r} \sum_{\substack{a_1,a_2,\cdots,a_r \\ a_i \ge 1, \sum a_i = k}} \binom{k}{a_1,a_2,\ldots,a_r} \sum_{1\le j_1<j_2<\cdots<j_r\le h} \mathfrak{S}(\boldsymbol{j}).$$

Gallagher [6] proved that, as $h \to \infty$,

$$\sum_{\substack{1\le j_1,j_2,\cdots,j_r\le h \\ \text{distinct}}} \mathfrak{S}(\boldsymbol{j}) \sim h^r, \tag{1.13}$$

and since this sum includes $r!$ permutations of the specified vector $\boldsymbol{j}$ when the components are ordered, we have

$$M_k(N,h,\psi) \sim N(\log N)^k \sum_{r=1}^{k} \frac{1}{r!} \left(\frac{h}{\log N}\right)^r \sum_{\substack{a_1,a_2,\cdots,a_r \\ a_i \ge 1, \sum a_i = k}} \binom{k}{a_1,a_2,\ldots,a_r}.$$

Letting $\left\{ \begin{matrix} k \\ r \end{matrix} \right\}$ denote the Stirling numbers of the second type, then it may be easily verified (see [12]) that

$$\sum_{\substack{a_1,a_2,\cdots,a_r \\ a_i \ge 1, \sum a_i = k}} \binom{k}{a_1,a_2,\ldots,a_r} = r! \left\{ \begin{matrix} k \\ r \end{matrix} \right\}. \tag{1.14}$$

We conclude that for $h \sim \lambda \log N$,

$$M_k(N,h,\psi) \sim N(\log N)^k \sum_{r=1}^{k} \left\{ \begin{matrix} k \\ r \end{matrix} \right\} \lambda^r, \tag{1.15}$$

which are the moments of a Poisson distribution with mean λ. The first 4 moments are, for $\lambda \ll 1$,

$$\begin{aligned} M_1(N,h,\psi) &\sim \lambda N \log N, \\ M_2(N,h,\psi) &\sim (\lambda+\lambda^2) N \log^2 N, \\ M_3(N,h,\psi) &\sim (\lambda+3\lambda^2+\lambda^3) N \log^3 N, \\ M_4(N,h,\psi) &\sim (\lambda+7\lambda^2+6\lambda^3+\lambda^4) N \log^4 N. \end{aligned}$$

The asymptotic formula for the first moment is known to be true as a simple consequence of the prime number theorem. The other moment formulas have never been proved. It is known that the asymptotic formula for the second moment follows from the assumption of the Riemann Hypothesis and the pair correlation conjecture for zeros of the Riemann zeta-function [10].

Turning to our approximation $\Lambda_R(n)$, we define

$$\psi_R(x) = \sum_{n\le x} \Lambda_R(n) \tag{1.16}$$

and first wish to examine the moments $M(N,h,\psi_R)$ defined as in (1.6). The same computation used for ψ to obtain (1.10) clearly applies and therefore we obtain

$$M_k(N,h,\psi_R) = \sum_{r=1}^{k} \sum_{\substack{a_1,a_2,\dots,a_r\\ a_i\ge 1, \sum a_i = k}} \binom{k}{a_1,a_2,\dots,a_r} \sum_{1\le j_1<j_2<\cdots<j_r\le h} \mathcal{S}_k(N,\boldsymbol{j},\boldsymbol{a}), \tag{1.17}$$

where $\mathcal{S}_k(N,\boldsymbol{j},\boldsymbol{a})$ is the correlation given in (1.3). Since $\Lambda_R(n)$ is not supported on the primes and prime powers as $\Lambda(n)$ is, we can not use (1.12) to reduce the problem to correlations without powers, and as we will see these powers sometimes effect the correlations for $\Lambda_R(n)$. Our main result on these correlations is contained in the following theorem.

Theorem 1.1 *Given $1\le k\le 3$, let $\boldsymbol{j} = (j_1,j_2,\dots,j_r)$ and $\boldsymbol{a} = (a_1,a_2,\dots a_r)$, where the j_i's are distinct integers, and $a_i \ge 1$ with $\sum_{i=1}^r a_i = k$. Assume $\max_i |j_i| \ll R^\epsilon$ and $R \gg N^\epsilon$. Then we have*

$$\mathcal{S}_k(N,\boldsymbol{j},\boldsymbol{a}) = \big(\mathcal{C}_k(\boldsymbol{a})\mathfrak{S}(\boldsymbol{j}) + o(1)\big)N(\log R)^{k-r} + O(R^k), \tag{1.18}$$

where $\mathcal{C}_k(\boldsymbol{a})$ has the values

$$\begin{aligned}
\mathcal{C}_1(1) &= 1,\\
\mathcal{C}_2(2) &= 1, \quad \mathcal{C}_2(1,1) = 1,\\
\mathcal{C}_3(3) &= \frac{3}{4}, \quad \mathcal{C}_3(2,1) = 1, \quad \mathcal{C}_3(1,1,1) = 1.
\end{aligned}$$

Here we have used the notational convention of dropping extra parentheses, so for example $\mathcal{C}_2((1,1)) = \mathcal{C}_2(1,1)$. The method of proof used in this paper is not limited to $k\le 3$, but it does become extremely complicated even for $k=4$. Using the new method mentioned before we will prove Theorem 1.1 holds for all k in the third paper in this series. The computation of the constants $\mathcal{C}_k(\boldsymbol{a})$ as k gets larger becomes increasingly difficult. We also believe the error term $O(R^k)$ can be improved. This has been done in the case $k=2$ (unpublished) where the error term $O(R^2)$ may be replaced by $O(R^{2-\delta})$ for a small constant δ. In the special case of $S_2(N,(0),(2))$ Graham [13] has removed the error term $O(R^2)$ entirely.

In proving Theorem 1.1 we will assume $j_1 = 0$. This may be done without loss of generality since we may shift the sum over n in $\mathcal{S}_k$ to $m = n + j_1$ and then return to the original summation range with an error $O(N^\epsilon)$ since $|j_1| \ll R^\epsilon$ and $\Lambda_R(n) \ll n^\epsilon$. Further $\mathfrak{S}(\boldsymbol{j}) = \mathfrak{S}(\boldsymbol{j} - \boldsymbol{j_1})$ where $\boldsymbol{j_1}$ is a vector with j_1 in every component, and so the singular series are unaffected by this shift.

We now apply Theorem 1.1 in (1.17), and obtain immediately using (1.13) that

$$M_k(N,h,\psi_R) = (1+o(1))P_k(\lambda,R)N(\log R)^k + O(h^kR^k) \tag{1.19}$$

where

$$P_k(\lambda,R) = \sum_{r=1}^{k} \frac{1}{r!}\left(\frac{h}{\log R}\right)^r \sum_{\substack{a_1,a_2,\dots,a_r\\ a_i\ge 1,\sum a_i=k}} \binom{k}{a_1,a_2,\dots,a_r} \mathcal{C}_k(\boldsymbol{a}). \tag{1.20}$$

Using the values of the constants $\mathcal{C}(\boldsymbol{a})$ in Theorem 1.1 we obtain the following result on moments of ψ_R.

Corollary 1.2 *For $h \sim \lambda \log N$, $\lambda \ll R^\epsilon$, and $R = N^{\theta_k}$, where θ_k is fixed and $0 < \theta_k < \frac{1}{k}$ for $1 \le k \le 3$, we have*

$$\begin{aligned} M_1(N,h,\psi_R) &\sim \lambda N \log N,\\ M_2(N,h,\psi_R) &\sim (\theta_2\lambda + \lambda^2)N\log^2 N,\\ M_3(N,h,\psi_R) &\sim (\frac{3}{4}{\theta_3}^2\lambda + 3\theta_3\lambda^2 + \lambda^3)N\log^3 N. \end{aligned}$$

We next consider the mixed moments

$$\tilde{M}_k(N,h,\psi_R) = \sum_{n=1}^{N}(\psi_R(n+h)-\psi_R(n))^{k-1}(\psi(n+h)-\psi(n)) \tag{1.21}$$

for $k \ge 2$, while if $k = 1$ we take

$$\tilde{M}_1(N,h,\psi_R) = M_1(N,h,\psi) \sim \lambda N \log N \tag{1.22}$$

for $1 \le R \le N$ by the prime number theorem. Now assume $k \ge 2$. On multiplying out and grouping as before for the terms involving Λ_R we have

$$\tilde{M}_k(N,h,\psi_R) = \sum_{r=2}^{k} \sum_{\substack{a_1,a_2,\dots,a_{r-1}\\ a_i\ge 1,\sum a_i=k-1}} \frac{1}{(r-1)!}\binom{k-1}{a_1,a_2,\dots,a_{r-1}} \sum_{\substack{1\le j_1,j_2,\cdots,j_{r-1}\le h\\ \text{distinct}}} V_{r-1}(N,\boldsymbol{j},\boldsymbol{a}), \tag{1.23}$$

where

$$V_{r-1}(N,\boldsymbol{j},\boldsymbol{a}) = \sum_{1\le m\le h}\sum_{n=1}^{N}\Lambda_R(n+j_1)^{a_1}\Lambda_R(n+j_2)^{a_2}\cdots\Lambda_R(n+j_{r-1})^{a_{r-1}}\Lambda(n+m)$$

Since, provided $n+j > R$ and $n+j \ne p^m$ for some $m \ge 2$ and $p < R$, we have

$$\Lambda_R(n+j)^a\Lambda(n+j) = (\log R)^a\Lambda(n+j),$$

we see

$$\begin{aligned}
&V_{r-1}(N,\boldsymbol{j},\boldsymbol{a})\\
&\quad=\sum_{i=1}^{r-1}(\log R)^{a_i}\sum_{n=1}^{N}\Big(\prod_{\substack{1\le s\le r-1\\ s\neq i}}\Lambda_R(n+j_s)^{a_s}\Big)\Lambda(n+j_i)\\
&\quad+\sum_{\substack{1\le j_r\le h\\ j_r\neq j_i\\ 1\le i\le r-1}}\sum_{n=1}^{N}\Lambda_R(n+j_1)^{a_1}\Lambda_R(n+j_2)^{a_2}\cdots\Lambda_R(n+j_{r-1})^{a_{r-1}}\Lambda(n+j_r)+O(RN^{\epsilon})\\
&\quad=\sum_{i=1}^{r-1}(\log R)^{a_i}\tilde{\mathcal{S}}_{k-a_i}(N,\boldsymbol{j_i},\boldsymbol{a_i})+\sum_{\substack{1\le j_l\le h\\ j_l\neq j_i\\ 1\le i\le r-1}}\tilde{\mathcal{S}}_k(N,\boldsymbol{j},\boldsymbol{a})+O(RN^{\epsilon})
\end{aligned}$$

where

$$\boldsymbol{j_i}=(j_1,j_2,\dots,j_{i-1},j_{i+1},\dots,j_{r-1},j_i),\quad \boldsymbol{a_i}=(a_1,a_2,\dots,a_{i-1},a_{i+1},\dots,a_{r-1},1). \tag{1.24}$$

We conclude for $k\ge 2$

$$\tilde{M}_k(N,h,\psi_R)=\sum_{r=2}^{k}\sum_{\substack{a_1,a_2,\cdots,a_{r-1}\\ a_i\ge1,\sum a_i=k-1}}\frac{1}{(r-1)!}\binom{k-1}{a_1,a_2,\dots,a_{r-1}}W_r(N,\boldsymbol{j},\boldsymbol{a})+O(RN^{\epsilon}), \tag{1.25}$$

where

$$W_r(N,\boldsymbol{j},\boldsymbol{a})=\sum_{i=1}^{r-1}(\log R)^{a_i}\sum_{\substack{1\le j_1,j_2,\cdots,j_{r-1}\le h\\ \text{distinct}}}\tilde{\mathcal{S}}_{k-a_i}(N,\boldsymbol{j_i},\boldsymbol{a_i})+\sum_{\substack{1\le j_1,j_2,\cdots,j_r\le h\\ \text{distinct}}}\tilde{\mathcal{S}}_k(N,\boldsymbol{j},\boldsymbol{a}). \tag{1.26}$$

We have reduced the calculation of the mixed moments to mixed correlations. Our method for evaluating the mixed correlations will prove as a by-product that the mixed correlations are asymptotically equal to the corresponding pure correlations in a certain range of R. Our results depend on the uniform distribution of primes in arithmetic progressions. We let

$$\psi(x;q,a)=\sum_{\substack{n\le x\\ n\equiv a(q)}}\Lambda(n), \tag{1.27}$$

and

$$E_{a,b}=\begin{cases}1, & \text{if } (a,b)=1,\\ 0, & \text{if } (a,b)>1.\end{cases} \tag{1.28}$$

On taking

$$E(x;q,a) = \psi(x;q,a) - E_{a,q}\frac{x}{\phi(q)}, \tag{1.29}$$

the estimate we need is, for some fixed $0 < \vartheta \leq 1$,

$$\sum_{1\leq q\leq x^{\vartheta-\epsilon}} \max_{\substack{a \\ (a,q)=1}} |E(x;q,a)| \ll \frac{x}{\log^{\mathcal{A}} x}, \tag{1.30}$$

for any $\epsilon > 0$, any $\mathcal{A} = \mathcal{A}(\epsilon) > 0$, and x sufficiently large. This is a weakened form of the Bombieri-Vinogradov theorem if $\vartheta = \frac{1}{2}$, and therefore (1.30) holds unconditionally if $\vartheta \leq \frac{1}{2}$. Elliott and Halberstam conjectured (1.30) is true with $\vartheta = 1$. The range of R where our results on mixed correlations hold depends on ϑ in (1.30). We first prove the following general result.

Theorem 1.3 *Given $k \geq 2$, let $\boldsymbol{j} = (j_1, j_2, \ldots, j_r)$ and $\boldsymbol{a} = (a_1, a_2, \ldots a_r)$, where the j_i's are distinct integers, $a_i \geq 1$, $a_r = 1$, and $\sum_{i=1}^r a_i = k$. Assume $\max_i |j_i| \ll R^{\frac{1}{k}}$ and that $N^\epsilon \ll R \ll N^{\min(\frac{\vartheta}{k-1},\frac{1}{k})-\epsilon}$. Then we have, with $\mathcal{A}$ from (1.30),*

$$\tilde{\mathcal{S}}_k(N,\boldsymbol{j},\boldsymbol{a}) = \mathcal{S}_k(N,\boldsymbol{j},\boldsymbol{a}) + O(R^k) + O\Big(\frac{N}{(\log N)^{\frac{\mathcal{A}}{2}-4^{k-3/2}+2^{k-1}-1}}\Big). \tag{1.31}$$

The proof of Theorem 1.3 involves proving that both $\tilde{\mathcal{S}}_k(N,\boldsymbol{j},\boldsymbol{a})$ and $\mathcal{S}_k(N,\boldsymbol{j},\boldsymbol{a})$ are asymptotic to the same main term and therefore they are asymptotic to each other in the range where both asymptotic formulas hold. Using Theorems 1.1 and 1.3 we can now immediately evaluate the mixed moments. There is, however, an inefficiency in the use of Theorem 1.3 which imposes the condition that $R \ll N^{\min(\frac{\vartheta}{k-1},\frac{1}{k})-\epsilon}$. The restriction $R \ll N^{\frac{1}{k}-\epsilon}$ in this condition arises from applying Theorem 1.1, but by directly evaluating the main term that arises in the proof of Theorem 1.3 we can remove this condition and prove the following result.

Theorem 1.4 *Given $2 \leq k \leq 3$, let $\boldsymbol{j} = (j_1, j_2, \ldots, j_r)$ and $\boldsymbol{a} = (a_1, a_2, \ldots a_r)$, where $r \geq 2$, $a_r = 1$, and where the j_i's are distinct integers, and $a_i \geq 1$ with $\sum_{i=1}^r a_i = k$. Assume $\max_i |j_i| \ll R^\epsilon$. Then we have, for $N^\epsilon \ll R \ll N^{\frac{\vartheta}{k-1}-\epsilon}$ where (1.30) holds with ϑ,*

$$\tilde{\mathcal{S}}_k(N,\boldsymbol{j},\boldsymbol{a}) = \big(\mathfrak{S}(\boldsymbol{j}) + o(1)\big)N(\log R)^{k-r}. \tag{1.32}$$

For larger k the constants $\mathcal{C}(\boldsymbol{a})$ will appear in this theorem as in Theorem 1.1, but for $k \leq 3$ all these constants for the mixed correlations are equal to 1.

Next, using (1.25) we are able to evaluate the first three mixed moments.

Corollary 1.5 *For $h \sim \lambda \log N$, $\lambda \ll R^{\epsilon}$, and $R = N^{\theta_k}$, where θ_k is fixed, $0 < \theta_1 \leq 1$, and $0 < \theta_k < \frac{\vartheta}{k-1}$ for $2 \leq k \leq 3$, we have,*

$$\begin{aligned} \tilde{M}_1(N,h,\psi_R) &\sim \lambda N \log N, \\ \tilde{M}_2(N,h,\psi_R) &\sim (\theta_2 \lambda + \lambda^2) N \log^2 N, \\ \tilde{M}_3(N,h,\psi_R) &\sim ({\theta_3}^2 \lambda + 3\theta_3 \lambda^2 + \lambda^3) N \log^3 N. \end{aligned}$$

The starting point of Bombieri and Davenport's [1] work on small gaps between primes is essentially equivalent to the inequality

$$\sum_{n=N+1}^{2N} \Big(\big(\psi(n+h) - \psi(n)\big) - \big(\psi_R(n+h) - \psi_R(n)\big) \Big)^2 \geq 0. \tag{1.33}$$

Letting

$$M_k'(N,h,\psi) = M_k(2N,h,\psi) - M_k(N,h,\psi), \tag{1.34}$$

with the corresponding definition for $M_k'(N,h,\psi_R)$ and $\tilde{M}_k'(N,h,\psi_R)$, we see that Corollary 1.2 holds with $M_k(N,h,\psi_R)$ replaced by $M_k'(N,h,\psi_R)$ and Corollary 1.5 holds with $\tilde{M}_k(N,h,\psi_R)$ replaced with $\tilde{M}_k'(N,h,\psi_R)$. On expanding (1.33) we have

$$M_2'(N,h,\psi) \geq 2\tilde{M}_2'(N,h,\psi_R) - M_2'(N,h,\psi_R)$$

which implies on taking $\theta_2 = 1/2 - \epsilon$ in Corollaries 1.2 and 1.5 that

$$M_2'(N,h,\psi) \geq (\frac{1}{2}\lambda + \lambda^2 - \epsilon) N \log^2 N. \tag{1.35}$$

Let p_j denote the j-th prime. If it is the case that $p_{j+1} - p_j > h = \lambda \log N$ for all $\frac{N}{2} < p_j \leq 2N$, then each of the intervals $(n, n+h]$ for $N < n \leq 2N$ contains at most one prime. Hence, since the prime powers may be removed with a negligible error, we have that $M_2'(N,h,\psi) \sim (\log N) M_1'(N,h,\psi) \sim \lambda N \log^2 N$ so that (1.35) implies

$$\lambda \geq \frac{1}{2}\lambda + \lambda^2 - \epsilon$$

which is false if $\lambda > \frac{1}{2}$. We conclude that

$$\liminf_{n\to\infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) \leq \frac{1}{2}. \tag{1.36}$$

More generally, we define for r any positive integer

$$\Xi_r = \liminf_{n\to\infty} \left(\frac{p_{n+r} - p_n}{\log p_n} \right) \tag{1.37}$$

and see that if $p_{n+r} - p_n > h = \lambda \log N$ for $N < p_n \leq 2N$ then

$$M_2'(N,h,\psi) \leq (r+\epsilon)(\log N)M_1'(N,h,\psi) \leq (r+\epsilon)\lambda N \log^2 N$$

which then implies that

$$r\lambda \geq \frac{1}{2}\lambda + \lambda^2 - \epsilon$$

and hence

$$\Xi_r \leq r - \frac{1}{2}. \tag{1.38}$$

Bombieri and Davenport were also able to improve (1.36) by incorporating an earlier method of Erdös into their argument. This method depends on the sieve upper bound for primes differing by an even number k given by

$$\sum_{n\leq N} \Lambda(n)\Lambda(n+k) \leq (\mathcal{B}+\epsilon)\mathfrak{S}(k)N \tag{1.39}$$

where $\mathfrak{S}(k) = \mathfrak{S}(\boldsymbol{j})$ with $\boldsymbol{j} = (0,k)$, and $\mathcal{B}$ is a constant. In [1] Bombieri and Davenport proved that (1.39) holds with $\mathcal{B} = 4$, and using this value they improved (1.36) and obtained

$$\Xi_1 \leq \frac{2+\sqrt{3}}{8} = 0.46650\ldots . \tag{1.40}$$

While (1.35) has never been improved, the refinements based on the Erdös method together with the choice of certain weights in a more general form of (1.35) has led to further improvements. Huxley [20] [21] proved that, letting θ_r be the smallest positive solution of

$$\theta_r + \sin\theta_r = \frac{\pi}{\mathcal{B}r}, \quad \sin\theta_r < (\pi+\theta_r)\cos\theta_r\ , \tag{1.41}$$

then

$$\Xi_r \leq \frac{2r-1}{4\mathcal{B}r}\left\{\mathcal{B}r + (\mathcal{B}r-1)\frac{\theta_r}{\sin\theta_r}\right\}. \tag{1.42}$$

With the value $\mathcal{B} = 4$ this gives

$$\Xi_1 \leq 0.44254\ldots, \quad \Xi_2 \leq 1.41051\ldots, \quad \Xi_3 \leq 2.39912\ldots, \quad \Xi_4 \leq 3.39326\ldots .$$

We note that the expression on the right-hand side of (1.42) is equal to

$$r - \frac{1+\frac{1}{\mathcal{B}}}{2} + O(\frac{1}{r}),$$

and thus for large r this bound approaches $r - \frac{5}{8}$ with $\mathcal{B} = 4$.

The best result known for $\mathcal{B}$ which holds uniformly for all k is $\mathcal{B} = 3.9171\ldots$ due to Chen [4]. However, in the application to obtain (1.42) one only needs (1.39) to hold uniformly for a restricted range of k; the condition $0 < |k| \le \log^2 N$ is more than sufficient. In this case there have been a string of improvements. For ease of comparison with the value $\mathcal{B} = 4$ used above, the value $\mathcal{B} = 3.5$ obtained by Bombieri, Friedlander, and Iwaniec [2] gives the values

$$\Xi_1 \le 0.43493\ldots, \quad \Xi_2 \le 1.39833\ldots, \quad \Xi_3 \le 2.38519\ldots, \quad \Xi_4 \le 3.37842\ldots.$$

All of these results above actually hold for a positive percentage of gaps. Maier [22] introduced a new method to prove that

$$\liminf_{n\to\infty} \left(\frac{p_{n+1}-p_n}{\log p_n}\right) \le e^{-\gamma} = 0.56145\ldots\ . \tag{1.43}$$

This method, which applies to special sets of sparse intervals, may be combined with the earlier methods to include this factor of $e^{-\gamma}$ times the earlier results. The argument was carried out with $\mathcal{B} = 4$ in [22], which then gives in particular

$$\Xi_1 \le 0.24846\ldots, \quad \Xi_2 \le 0.79194\ldots, \quad \Xi_3 \le 1.34700\ldots, \quad \Xi_4 \le 1.90518\ldots.$$

Our approach for examining gaps between primes is to consider the mixed third moment

$$\tilde{M}_3'(N,h,\psi_R,C) = \sum_{n=N+1}^{2N} \big(\psi(n+h)-\psi(n)\big)\big(\psi_R(n+h)-\psi_R(n)-C\log N\big)^2, \tag{1.44}$$

Here C may be chosen as a function of h and R to optimize the argument. The idea behind the use of $\tilde{M}_3'(N,h,\psi_R,C)$ is that it will approximate and thus provide some of the same information as the third moment $M_3'(N,h,\psi)$. If $p_{n+r}-p_n > h = \lambda\log N$ for all $N < p_n \le 2N$ then, removing prime powers as before, we have

$$\tilde{M}_3'(N,h,\psi_R,C) \le (r+\epsilon)\log N \sum_{n=N+1}^{2N} \big(\psi_R(n+h)-\psi_R(n)-C\log N\big)^2. \tag{1.45}$$

Corollaries 1.2 and 1.5 allow us to evaluate both sides of (1.45), and on choosing C appropriately we are able to prove the following result.

Theorem 1.6 *For $r \ge 1$, we have*

$$\Xi_r \le r - \frac{1}{2}\sqrt{r}. \tag{1.46}$$

Further, assuming that for $h \ll \log N$ and $R \le N^{\frac{1}{4}}$

$$M_4(N,h,\psi_R) \ll N\log^4 N, \tag{1.47}$$

then we have that, for $h = \lambda \log N$ and $\lambda > r - \frac{1}{2}\sqrt{r}$,

$$\sum_{\substack{N+1 \le p_n \le 2N \\ p_{n+r} - p_n < h}} 1 \gg_\lambda \frac{N}{\log N}. \tag{1.48}$$

Thus we have

$$\Xi_1 \le \frac{1}{2}, \quad \Xi_2 \le 1.29289\ldots, \quad \Xi_3 \le 2.13397\ldots, \quad \Xi_4 \le 3.$$

We see that our result improves on the results of Huxley when $r \ge 2$, although Maier's results are still better. Our theorem corresponds to (1.38) in that it does not use the Erdös method. It is possible to incorporate the Erdös method into our method too, but this requires we first obtain an asymptotic formula for $M_4(N, h, \psi_R)$. One should also be able to incorporate Maier's method as well, which would then give better results than are currently known for $r \ge 2$.

The result in (1.48) shows that the small gaps produced in the theorem form a positive proportion of all the gaps assuming that (1.47) holds. We will prove (1.47) in a later paper in this series and thus show that (1.48) holds unconditionally.

We will actually prove

$$\Xi_r \le r - \sqrt{\frac{\vartheta r}{2}}, \tag{1.49}$$

where ϑ is the number in (1.30). Therefore, assuming the Elliott-Halberstam Conjecture in the form that one may take $\vartheta = 1$ in (1.30), we have

$$\Xi_r \le r - \sqrt{\frac{r}{2}}$$

which in particular gives

$$\Xi_1 \le 0.29289\ldots, \quad \Xi_2 \le 1, \quad \Xi_3 \le 1.77525\ldots, \quad \Xi_4 \le 2.58578\ldots.$$

These results are in contrast to the method of Bombieri and Davenport where the Elliott-Halberstam conjecture does not improve their results directly. (The Elliott-Halberstam conjecture does allow one to take $\mathcal{B} = 2$ in (1.39), and therefore leads to small improvements in Huxley's results, which for $r \ge 2$ are weaker than the result in Theorem 1.6.) We cannot extend these last results obtained assuming an Elliott-Halberstam conjecture to a positive proportion of gaps because we can not prove (1.47) for $\theta > \frac{1}{4}$. Our proof gives that the number of gaps we produce in this case is $\gg N \log^{-B} N$ for some positive constant $B > 1$.

Our method can also be used to examine larger than average gaps between primes. In this case much more is known than for small gaps; the latest result being that [24]

$$\max_{p_n\le N}\frac{p_{n+1}-p_n}{\log p_n}\ge (2e^\gamma-o(1))\frac{\log\log N\log\log\log\log N}{(\log\log\log N)^2}. \tag{1.50}$$

If one were to ask however for a positive proportion of gaps larger than the average, then it is a remarkable fact that nothing non-trivial is known. [2] What can be proved is that a positive proportion of the interval $(N,2N]$ is contained between consecutive primes whose difference is a little larger than average. To formalize this, we let Θ_r be the supremum over all λ for which

$$\sum_{\substack{N<p_n\le 2N\\ p_{n+r}-p_n\ge\lambda\log N}}(p_{n+r}-p_n)\gg_\lambda N \tag{1.51}$$

for all sufficiently large N. Then using the Erdös method one finds that [3]

$$\Theta_1\ge 1+\frac{1}{2\mathcal{B}} \tag{1.52}$$

where $\mathcal{B}$ is the number in (1.39).

To apply (1.44) to this problem, we assume that $p_{j+r}-p_j<h=\lambda\log N$ for all $N<p_j\le 2N$ in which case the interval $(n,n+h]$ always contains at least r primes, and therefore (1.45) holds with the inequality reversed. On optimizing C we obtain the following result.

Theorem 1.7 *Assume that* (1.47) *holds. For* $r\ge 1$ *we have that*

$$\Theta_r\ge r+\frac{1}{2}\sqrt{r}. \tag{1.53}$$

As mentioned above, we will prove (1.47) in a later paper in this series, which will show that Theorem 1.7 holds unconditionally.

The proof of Theorems 1.6 and 1.7 only require the asymptotic formula for the third mixed moment in Corollary 1.5 and the second moment for ψ_R in Corollary 1.2. Thus the results in sections 6–10 which are concerned with triple correlations of ψ_R may be skipped by the reader who is only interested in our applications to primes.

Notation. In this paper N will always be a large integer, p denotes a prime number, and sums will start at 1 if a lower limit is unspecified. When a sum is denoted with a dash as $\sum'$ this always indicates we will sum over all variables expressed by inequalities in the conditions of summation and these variables will all be pairwise relatively prime

[2]The first-named author of this paper learned of this from Carl Pomerance after a talk in which the author had claimed to have such a result.

with each other. We will always take the value of a void sum to be zero and the value of a void product to be 1. The letter ϵ will denote a small positive number which may change in each equation. We will also use the Iverson notation [12] that putting brackets around a true-false statement will replace the statement by 1 if it is true, and 0 if it is false:

$$[P(x)] = \begin{cases} 1, & \text{if } P(x) \text{ is true,} \\ 0, & \text{if } P(x) \text{ is false.} \end{cases} \tag{1.54}$$

As usual, (a, b) denotes the greatest common divisor of a and b and $[a_1, a_2, \cdots, a_n]$ denotes the least common multiple of $a_1, a_2, \ldots, a_n$.

Acknowledgements

A preliminary version of this paper was presented at the First Workshop on L-functions and Random Matrices at the American Institute of Mathematics in May, 2001. Thanks to a suggestion of Peter Sarnak during the talk, and encouragement of John Friedlander after the talk, the authors found an alternative method for proving Theorem 1.1 which generalizes to the case of k-correlations. This new method has its own complications which make it very easy for mistakes to creep into the calculations. Following the conference, the first-named author visited AIM where these problems were solved jointly with Brian Conrey and David Farmer. Farmer has written a program in Mathematica to compute the constants $\mathcal{C}_k(\boldsymbol{a})$ and has found that $\mathcal{C}_4(4) = \frac{3}{4}$, $\mathcal{C}_5(5) = \frac{11065}{2^{14}} = .67535\ldots$, and $\mathcal{C}_6(6) = \frac{11460578803}{2^{34}} = .66709\ldots$. The authors would like to thank these individuals and the American Institute of Mathematics.

2. Lemmas

For j a non-negative integer, we define the arithmetic function $\phi_j(n)$ on the primes by

$$\phi_j(p) = p - j, \tag{2.1}$$

$\phi_j(1) = 1$, and extend the definition to squarefree integers by multiplicativity. Thus for n squarefree $\phi_0(n) = n$, and $\phi_1(n) = \phi(n)$. We will not need to extend the definition beyond the squarefree integers here. Letting

$$p(j) = \begin{cases} j, & \text{if } j \text{ is a prime,} \\ 1, & \text{otherwise,} \end{cases} \tag{2.2}$$

we next define

$$H_j(n) = \prod_{\substack{p|n \\ p\neq j-1,\ p\neq j}} \left(1 + \frac{1}{p-j}\right) = \prod_{\substack{p|n \\ p\neq j-1,\ p\neq j}} \left(1 + \frac{1}{\phi_j(p)}\right) = \sum_{\substack{d|n \\ (d,p(j-1)p(j))=1}} \frac{\mu^2(d)}{\phi_j(d)}. \quad (2.3)$$

We see that for n squarefree $H_0(n) = \sigma(n)/n$, $H_1(n) = n/\phi(n)$, and in general for $j \geq 1$

$$H_j(n) = \prod_{\substack{p|n \\ p\neq j-1,\ p\neq j}} \left(\frac{p-j+1}{p-j}\right) = \frac{\phi_{j-1}(\frac{n}{(n,p(j-1)p(j))})}{\phi_j(\frac{n}{(n,p(j-1)p(j))})}, \qquad (\mu^2(n) \neq 0). \quad (2.4)$$

Next, we define the singular series for $j \geq 1$ and $n \neq 0$ by

$$\mathfrak{S}_j(n) = \begin{cases} C_j G_j(n) H_j(n), & \text{if } p(j)|n, \\ 0, & \text{otherwise.} \end{cases} \quad (2.5)$$

where

$$G_j(n) = \prod_{\substack{p|n \\ p=j-1 \text{ or } p=j}} \left(\frac{p}{p-1}\right), \quad (2.6)$$

and

$$C_j = \prod_{\substack{p \\ p\neq j-1,\ p\neq j}} \left(1 - \frac{j-1}{(p-1)(p-j+1)}\right). \quad (2.7)$$

The case of $j = 3$ is special because it is the only case where $p(j-1)$ and $p(j)$ are both greater than one. We see that for $j = 1$ and $n \neq 0$

$$\mathfrak{S}_1(n) = \prod_{p|n} \left(\frac{p}{p-1}\right) = \frac{n}{\phi(n)} \quad (2.8)$$

and for $j = 2$ we have the familiar singular series for the Goldbach and prime twins conjectures

$$\mathfrak{S}_2(n) = \begin{cases} 2C_2 \prod\limits_{\substack{p|n \\ p>2}} \left(\frac{p-1}{p-2}\right), & \text{if } n \text{ is even, } n \neq 0; \\ 0, & \text{if } n \text{ is odd;} \end{cases} \quad (2.9)$$

where

$$C_2 = \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right). \quad (2.10)$$

Lemma 2.1 *For $R \geq 1$, $j \geq 0$, $p(j)|k$, and $0 \leq \log|k| \ll \log R$, we have*

$$\sum_{\substack{d \leq R \\ (d,k)=1}} \frac{\mu(d)}{\phi_j(d)} \log \frac{R}{d} = \mathfrak{S}_{j+1}(k) + r_j(R,k), \tag{2.11}$$

where

$$r_j(R,k) \ll_j e^{-c_1\sqrt{\log R}}, \tag{2.12}$$

and c_1 is an absolute positive constant. Also,

$$\sum_{\substack{d \leq R \\ (d,k)=1}} \frac{\mu(d)}{\phi_j(d)} \ll_j e^{-c_1\sqrt{\log R}}. \tag{2.13}$$

Special cases of Lemma 2.1 have been proved before. When $j = 0$ this was used by Selberg [25], and also Graham [13], but we have made the error term stronger with regard to k by an argument suggested in [5]. It is easy to make the j dependence explicit in the error term, but in this paper we will assume j is fixed (actually we only use $j \leq 2$.) We will sometimes use Lemma 2.1 in the weaker form

$$\sum_{\substack{d \leq R \\ (d,k)=1}} \frac{\mu(d)}{\phi_j(d)} \log \frac{R}{d} = \mathfrak{S}_{j+1}(k) + O_j(\frac{1}{(\log 2R)^A})$$

for A any positive number, and assuming the same conditions as in Lemma 2.1. Further, in handling error terms, we will need to remove the restriction $0 \leq \log|k| \ll \log R$ in Lemma 2.1, in which case we have the error estimate

$$r_j(R,k) \ll_j m(k)e^{-c_1\sqrt{\log R}}, \tag{2.14}$$

which holds uniformly for $k \geq 1$ and $R \geq 1$, where $m(k)$ is defined below in equation (2.18). This estimate also holds in (2.13).

The next lemma is a generalization of a result of Hildebrand [15].

Lemma 2.2 *For $R \geq 1$, $j \geq 1$ and $p(j)|k$, we have*

$$\sum_{\substack{d \leq R \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} = \begin{cases} \frac{1}{\mathfrak{S}_j(k)}(\log R + D_j + h_j(k)) + O(\frac{m(k)}{\sqrt{R}}) & \text{if } p(j-1)|k, \\ O(\frac{m(k)}{\sqrt{R}}), & \text{if } p(j-1) \not| k, \end{cases} \tag{2.15}$$

where

$$D_j = \gamma + \sum_{p \neq j-1} \frac{(2-j)\log p}{(p-j+1)(p-1)}, \tag{2.16}$$

$$\begin{aligned} h_j(k) &= \sum_{p|k} \frac{\log p}{p-1} - \sum_{\substack{p|k \\ p\neq j-1}} \frac{(2-j)\log p}{(p-j+1)(p-1)} \\ &= \sum_{\substack{p|k \\ p\neq j-1}} \frac{\log p}{(p-j+1)} + [(p(j-1),k)>1]\frac{\log(j-1)}{j-2}, \end{aligned} \tag{2.17}$$

and

$$m(k) = \sum_{d|k} \frac{\mu^2(d)}{\sqrt{d}} = \prod_{p|k}\left(1+\frac{1}{\sqrt{p}}\right). \tag{2.18}$$

The case $j=1$ of this lemma is Hilfssatz 2 of [15]. The proof of this generalization only requires minor modifications in Hildebrand's proof which we sketch. In applying this lemma we will sometimes use the simple estimates (see [11])

$$h_j(k) \ll_j \log\log 3k, \qquad m(k) \ll \exp\left(\frac{c\sqrt{\log k}}{\log\log 3k}\right). \tag{2.19}$$

We will frequently use the estimate, for $p(j)|k$ and $\log|k| \ll \log R$,

$$\sum_{\substack{d\leq R \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} \ll \log 2R \tag{2.20}$$

which follows immediately from Lemma 2.2 or may be seen directly. We also need the following result that is obtained by partial summation in Lemma 2.2. For $j\geq 1$ and $p(j)|k$, we have

$$\begin{aligned} &\sum_{\substack{d\leq R \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} \log\frac{R}{d} \\ &\qquad = \begin{cases} \frac{1}{\mathfrak{S}_j(k)}\left(\frac{1}{2}\log^2 R + (D_j+h_j(k))\log R + E_j(k)\right) + O\left(\frac{m(k)}{\sqrt{R}}\right) & \text{if } p(j-1)|k, \\ E_j(k) + O\left(\frac{m(k)}{\sqrt{R}}\right), & \text{if } p(j-1)\not|\,k, \end{cases} \end{aligned} \tag{2.21}$$

where $E_j(k)$ is given by

$$E_j(k) = \int_1^\infty \Bigg(\sum_{\substack{d\leq u \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} - \frac{1}{\mathfrak{S}_j(k)}\left(\log u + D_j + h_j(k)\right)\Bigg)\frac{du}{u}. \tag{2.22}$$

Lemma 2.3 *For $R \geq 1$, $j \geq 1$, $p(j)|k$, and $0 \leq \log|k| \ll \log R$, we have*

$$\sum_{\substack{d \leq R \\ (d,k)=1}} \frac{\mu(d)}{\phi_j(d)} \mathfrak{S}_{j+1}(dk) \log \frac{R}{d}$$
$$= \mu(p(j+1))\mu((k,p(j+1)))\mathfrak{S}_{j+1}(kp(j+1))\Big(\mathfrak{S}_{j+2}(kp(j+1)) \tag{2.23}$$
$$+ r_{j+1}(\frac{R(k,p(j+1))}{p(j+1)}, kp(j+1))\Big),$$

where $r_j(R,k)$ is the error term in Lemma 2.1.

Lemma 2.4 *For $R \geq 1$, $j \geq 1$ and $p(j)|k$, we have*

$$\begin{aligned}\sum_{\substack{d \leq R \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} \mathfrak{S}_{j+1}(dk) &= \log\left(\frac{R(k,p(j+1))}{p(j+1)}\right) + D_{j+1} + h_{j+1}(kp(j+1)) \\ &\quad + O\left(\frac{\mathfrak{S}_{j+1}(kp(j+1))m(kp(j+1))}{\sqrt{\frac{R(k,p(j+1))}{p(j+1)}}}\right)\end{aligned} \tag{2.24}$$

and

$$\sum_{\substack{d | r \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} \mathfrak{S}_{j+1}(dk) = \mathfrak{S}_{j+1}(rk). \tag{2.25}$$

Our final lemma relates the singular series given in (1.9) for r equal to two and three to the singular series in (2.5).

Lemma 2.5 *For $\boldsymbol{k} = (0,k)$, with $k \neq 0$, we have*

$$\mathfrak{S}(\boldsymbol{k}) = \mathfrak{S}_2(k), \tag{2.26}$$

and for $\boldsymbol{k} = (0,k_1,k_2)$, $k_1 \neq k_2 \neq 0$, $\kappa = (k_1,k_2)$, and $\Delta = k_1k_2(k_2-k_1)$, we have

$$\mathfrak{S}(\boldsymbol{k}) = \mathfrak{S}_2(\kappa)\mathfrak{S}_3(\Delta). \tag{2.27}$$

3. Proof of the Lemmas

Proof of Lemma 2.1. We assume $p(j)|k$. Let $s=\sigma+it$, $\sigma>0$, and define

$$\begin{aligned} F(s) &= \sum_{\substack{n=1\\(n,k)=1}}^{\infty} \frac{\mu(n)}{\phi_j(n)n^s} = \prod_{p\nmid k}\left(1-\frac{1}{(p-j)p^s}\right) \\ &= \frac{1}{\zeta(s+1)}\prod_{p|k}\left(1-\frac{1}{p^{s+1}}\right)^{-1}\prod_{p\nmid k}\left(1-\frac{1}{(p-j)p^s}\right)\left(1-\frac{1}{p^{s+1}}\right)^{-1} \\ &= \frac{1}{\zeta(s+1)}\prod_{p|k}\left(1-\frac{1}{p^{s+1}}\right)^{-1}\prod_{p\nmid k}\left(1-\frac{j}{(p-j)(p^{s+1}-1)}\right) \\ &= \frac{1}{\zeta(s+1)}g_k(s)h_k(s). \end{aligned} \tag{3.1}$$

We see the product for $h_k(s)$ converges absolutely for $\mathrm{Re}(s)>-1$, and therefore $F(s)$ is an analytic function in this half-plane except possibly for poles at the zeros of $\zeta(s+1)$.

We now apply the formula, for $m\geq 2$ and $b>0$,

$$\frac{(m-1)!}{2\pi i}\int_{b-i\infty}^{b+i\infty}\frac{x^s}{s^m}\,ds = \begin{cases} 0, & \text{if } 0<x\leq 1,\\ (\log x)^{m-1}, & \text{if } x\geq 1,\end{cases}$$

which, in the case $m=2$, gives

$$\sum_{\substack{d\leq R\\(d,k)=1}}\frac{\mu(d)}{\phi_j(d)}\log\frac{R}{d} = \frac{1}{2\pi i}\int_{b-i\infty}^{b+i\infty}F(s)\frac{R^s}{s^2}\,ds. \tag{3.2}$$

By Theorem 3.8 and (3.11.8) of [26] there exists a small positive constant c such that $\zeta(\sigma+it)\neq 0$ in the region $\sigma\geq 1-\frac{c}{\log(|t|+2)}$ and all t, and further

$$\frac{1}{\zeta(\sigma+it)} \ll \log(|t|+2) \tag{3.3}$$

in this region. (There are stronger results but this suffices for our needs.) We now move the contour to the left to the path $\mathcal{L}$ given by $s=-\frac{c}{\log(|t|+2)}+it$. When $p(j+1)\nmid k$, $h_k(s)$ has a simple zero at $s=0$ and hence $F(s)/s^2$ is analytic at $s=0$ and no contribution occurs, but when $p(j+1)|k$ (including $p(j+1)=1$), $F(s)/s^2$ has a simple pole at $s=0$

which gives a contribution from the residue of

$$\begin{aligned} g_k(0)h_k(0) &= \prod_{p|k}\left(1-\frac{1}{p}\right)^{-1}\prod_{p\not| k}\left(1-\frac{j}{(p-1)(p-j)}\right) \\ &= \prod_{\substack{p|k \\ p=j \text{ or } p=j+1}}\left(\frac{p}{p-1}\right)\prod_{\substack{p|k \\ p\neq j,\ p\neq j+1}}\left(1-\frac{1}{p}\right)^{-1}\left(1-\frac{j}{(p-1)(p-j)}\right)^{-1} \\ &\qquad \prod_{\substack{p \\ p\neq j,\ p\neq j+1}}\left(1-\frac{j}{(p-1)(p-j)}\right) \\ &= G_{j+1}(k)\prod_{\substack{p \\ p\neq j, p\neq j+1}}\left(1-\frac{j}{(p-1)(p-j)}\right)\prod_{\substack{p|k \\ p\neq j,\ p\neq j+1}}\left(\frac{p-j}{p-j-1}\right) \\ &= C_{j+1}G_{j+1}(k)H_{j+1}(k). \end{aligned}$$

Hence we have

$$\sum_{\substack{d\leq R \\ (d,k)=1}}\frac{\mu(d)}{\phi_j(d)}\log\frac{R}{d} = \mathfrak{S}_{j+1}(k) + \frac{1}{2\pi i}\int_{\mathcal{L}}F(s)\frac{R^s}{s^2}\,ds = \mathfrak{S}_{j+1}(k) + r_j(R,k). \tag{3.4}$$

It remains to estimate the integral in (3.4). On $\mathcal{L}$ we have $-\frac{1}{4}\leq\sigma<0$, and therefore

$$|h_k(s)| \ll \prod_p\left(1+O_j(\frac{1}{p^{2+\sigma}})\right) \ll_j 1.$$

For $g_k(s)$, we see that $g_1(s)=1$ since the product defining $g_1(s)$ is void, and for $k>1$

$$\begin{aligned} \log|g_k(s)| &\leq -\sum_{p|k}\log(1-\frac{1}{p^{\sigma+1}}) \\ &\ll \sum_{p|k}\frac{1}{p^{\sigma+1}} \\ &\ll \sum_{p<2\log 2k}\frac{1}{p^{\sigma+1}} \ll (\log k)^{-\sigma} \end{aligned}$$

and hence

$$|g_k(s)| \ll e^{\sqrt[4]{\log k}}, \quad \text{for } k\geq 1. \tag{3.5}$$

Thus the integral in (3.4) is

$$\ll_j e^{\sqrt[4]{\log k}}\int_{-\infty}^{\infty}R^{-\frac{c}{\log(|t|+2)}}\log(|t|+2)\frac{1}{\left(\frac{c}{\log(|t|+2))}\right)^2+t^2}\,dt$$

This last integral is, for any $w \geq 2$,

$$\ll \int_0^w R^{-\frac{c}{\log(|t|+2)}}\,dt + \int_w^\infty \frac{\log t}{t^2}\,dt \ll we^{\frac{-c\log R}{\log w}} + \frac{\log w}{w}$$

and hence, on choosing $\log w = \frac{1}{2}\sqrt{c\log R}$ we have, since $\log|k| \ll \log R$, that the error term is

$$\ll_j e^{\sqrt[4]{\log k}}\left(e^{-\sqrt{c\log R}} + \sqrt{c\log R}e^{-\frac{1}{2}\sqrt{c\log R}}\right) \ll_j e^{-c_1\sqrt{\log R}},$$

which proves the first part of Lemma 1.

The bound in equation (2.14) follows from the previous argument when we replace the estimate for $g_k(s)$ used above by the bound, for $-\frac{1}{4} < \sigma < 0$,

$$|g_k(s)| \ll \prod_{p|k}\left(1 + \frac{1}{p^{1+\sigma}}\right) \ll m(k),$$

which follows from

$$\begin{aligned}\log|g_k(s)| &\leq -\sum_{p|k}\log(1 - \frac{1}{p^{\sigma+1}})\\ &= \sum_{p|k}\left(\log(1 + \frac{1}{p^{\sigma+1}}) + O(\frac{1}{p^{2(\sigma+1)}})\right)\\ &= \sum_{p|k}\log(1 + \frac{1}{p^{\sigma+1}}) + O(\sum_p \frac{1}{p^{\frac{3}{2}}})\\ &= \sum_{p|k}\log(1 + \frac{1}{p^{\sigma+1}}) + O(1).\end{aligned}$$

To prove (2.13), we apply Perron's formula (see [26], Chapter 3) in the usual way to obtain, with $b = 1/\log R$,

$$\sum_{\substack{d\leq R\\(d,k)=1}} \frac{\mu(d)}{\phi_j(d)} = \frac{1}{2\pi i}\int_{b-iT}^{b+iT} F(s)\frac{R^s}{s}\,ds + O_j(\frac{\log^2 R}{T}).$$

Moving the contour to the left to $\mathcal{L}$ for $-T \leq t \leq T$ we have no residue, and we may estimate the integral along $\mathcal{L}$ and the upper and lower horizontal paths by

$$\ll_j e^{\sqrt[4]{\log k}}\Big(\log^2 Te^{-c\frac{\log R}{\log T}} + \frac{\log T}{T}\Big).$$

Now $\log|k| \ll \log R$, and taking $T = e^{\sqrt{\log R}}$ shows the above error is $\ll e^{-c_1\sqrt{\log R}}$ for a sufficiently small constant c_1. This completes the proof of Lemma 2.1.

Proof of Lemma 2.2. We follow Hildebrand's proof of the case $j = 1$, indicating only the main steps. We assume $p(j)|k$, and $j \geq 1$. Letting

$$f_k(n) = \begin{cases} \frac{\mu^2(n)n}{\phi_j(n)}, & \text{if } (n,k) = 1, \\ 0, & \text{otherwise,} \end{cases} \tag{3.6}$$

and defining $g_k(n)$ by

$$f_k(n) = \sum_{d|n} g_k(d)$$

so that

$$g_k(p^m) = f_k(p^m) - f_k(p^{m-1}) = \begin{cases} \frac{j}{p-j}, & m = 1, p \not| k, \\ -1, & m = 1, p|k, \\ \frac{-p}{p-j} & m = 2, p \not| k, \\ 0, & m = 2 \text{ and } p|k, \text{ or } m > 2, \end{cases} \tag{3.7}$$

then we have

$$\begin{aligned} \sum_{\substack{d \leq R \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} &= \sum_{n \leq R} \frac{f_k(n)}{n} \\ &= \sum_{n \leq R} \frac{1}{n} \sum_{d|n} g_k(d) \\ &= \sum_{d \leq R} \frac{g_k(d)}{d} \sum_{m \leq R/d} \frac{1}{m} \\ &= \sum_{d \leq R} \frac{g_k(d)}{d} \left(\log \frac{R}{d} + \gamma + O(\frac{d}{R}) \right) \\ &= \sum_{d \leq R} \frac{g_k(d)}{d} \log \frac{R}{d} + \gamma \sum_{d \leq R} \frac{g_k(d)}{d} + O(\frac{1}{R} \sum_{d \leq R} |g_k(d)|) \\ &= S_1 + \gamma S_2 + O(\frac{1}{R} S_3). \end{aligned} \tag{3.8}$$

Using (3.7) and the multiplicativity of $g_k(n)$ we easily verify that

$$M_j(k) = \sum_{d=1}^{\infty} \frac{g_k(d)}{d} = \begin{cases} \frac{1}{\mathfrak{S}_j(k)}, & \text{if } p(j-1)|k, \\ 0, & \text{if } p(j-1) \not| k. \end{cases}$$

As in [15] we find

$$S_3 \ll_j \sqrt{R} m(k)$$

which by partial summation implies

$$\sum_{d>R} \frac{g_k(d)}{d} \ll \frac{m(k)}{\sqrt{R}}.$$

These results now imply

$$S_2 = M_j(k) + O(\frac{m(k)}{\sqrt{R}})$$

and

$$S_1 = M_j(k)\log R - \sum_{d=1}^{\infty} \frac{g_k(d)\log d}{d} + O(\frac{m(k)}{\sqrt{R}}).$$

Finally, letting

$$G_k(s) = \sum_{n=1}^{\infty} \frac{g_k(n)}{n^{s+1}},$$

we have

$$-\sum_{d=1}^{\infty} \frac{g_k(d)\log d}{d} = {G_k}'(0)$$

which on using (2.16) and the Euler product for $G_k(s)$ gives ${G_k}'(0) = M_j(k)(D_j + h_j(k) - \gamma)$ where $h_j(k)$ is given by (2.17).

Proof of Lemma 2.3. We have, assuming $p(j)|k$,

$$\begin{aligned}
&\sum_{\substack{d\le R\\ (d,k)=1}} \frac{\mu(d)}{\phi_j(d)} \mathfrak{S}_{j+1}(dk)\log\frac{R}{d}\\
&= C_{j+1}G_{j+1}(k)H_{j+1}(k) \sum_{\substack{d\le R\\ (d,k)=1\\ p(j+1)|dk}} \frac{\mu(d)}{\phi_j(d)} G_{j+1}(d)H_{j+1}(d)\log\frac{R}{d}\\
&= C_{j+1}G_{j+1}(k)H_{j+1}(k) \sum_{\substack{d\le R\\ (d,k)=1\\ \frac{p(j+1)}{(k,p(j+1))}|d}} \frac{\mu(d)}{\phi_j(d)} G_{j+1}(d)H_{j+1}(d)\log\frac{R}{d}\\
&= \mu(p(j+1))\mu((k,p(j+1)))C_{j+1}G_{j+1}(kp(j+1))H_{j+1}(k)\\
&\qquad \sum_{\substack{m\le \frac{R(k,p(j+1))}{p(j+1)}\\ (m,kp(j+1))=1}} \frac{\mu(m)}{\phi_{j+1}(m)} \log\Big(\frac{R(k,p(j+1))}{mp(j+1)}\Big)\\
&= \mu(p(j+1))\mu((k,p(j+1)))\mathfrak{S}_{j+1}(kp(j+1))\\
&\qquad \Big(\mathfrak{S}_{j+2}(kp(j+1)) + r_{j+1}\big(\frac{R(k,p(j+1))}{p(j+1)}, kp(j+1)\big)\Big),
\end{aligned}$$

by Lemma 2.1.

Proof of Lemma 2.4. The same argument used above to prove Lemma 2.3 shows that the sum in (2.24) is equal to

$$\mathfrak{S}_{j+1}(kp(j+1)) \sum_{\substack{m\le \frac{R(k,p(j+1))}{p(j+1)} \\ (m,kp(j+1))=1}} \frac{\mu^2(m)}{\phi_{j+1}(m)}.$$

Since $p(j)|k$, equation (2.24) now follows from Lemma 2.2.

To prove (2.25), we proceed as before in the proof of Lemma 2.3 to obtain

$$\sum_{\substack{d|r \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)}\mathfrak{S}_{j+1}(dk) = C_{j+1}G_{j+1}(k)H_{j+1}(k) \sum_{\substack{d|r \\ (d,k)=1 \\ p(j+1)|dk}} \frac{\mu^2(d)}{\phi_j(d)}G_{j+1}(d)H_{j+1}(d). \tag{3.9}$$

Suppose first that $p(j+1)|k$. Then since $p(j)|k$ also we have that $G_{j+1}(d) = 1$ in the sum on the right, and hence our expression becomes

$$\begin{aligned} &= C_{j+1}G_{j+1}(k)H_{j+1}(k) \sum_{\substack{d|r \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)}H_{j+1}(d) \\ &= C_{j+1}G_{j+1}(k)H_{j+1}(k) \sum_{\substack{d|r \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_{j+1}(d)} \\ &= C_{j+1}G_{j+1}(k)H_{j+1}(k) \prod_{\substack{p|r \\ p\nmid k}} \left(1+\frac{1}{p-j-1}\right) \\ &= C_{j+1}G_{j+1}(k)H_{j+1}(rk) = \mathfrak{S}_{j+1}(rk) \end{aligned}$$

since here $G_{j+1}(k) = G_{j+1}(p(j)p(j+1)) = G_{j+1}(rk)$.

Now assume the alternative case that $p(j+1) \nmid k$. Then the right-hand side of (3.9) is

$$= C_{j+1}G_{j+1}(k)H_{j+1}(k) \sum_{\substack{p(j+1)e|r \\ (e,kp(j+1))=1}} \frac{\mu^2(e)}{\phi_j(e)}G_{j+1}(ep(j+1))H_{j+1}(e).$$

If $p(j+1) \nmid r$ this sum has no terms and is zero which proves (2.25) in this case. If

$p(j+1)|r$ our expression becomes

$$\begin{aligned} &= C_{j+1}G_{j+1}(p(j)p(j+1))H_{j+1}(k) \sum_{\substack{e|r \\ (e,kp(j+1))=1}} \frac{\mu^2(e)}{\phi_{j+1}(e)} \\ &= C_{j+1}G_{j+1}(p(j)p(j+1))H_{j+1}(k) \prod_{\substack{p|r \\ p \nmid kp(j+1)}} \left(1+\frac{1}{p-j-1}\right) \\ &= \mathfrak{S}_{j+1}(rk), \end{aligned}$$

which completes the proof of Lemma 2.4.

Proof of Lemma 2.5. For the case $\boldsymbol{k}=(0,k)$ we have by (1.9) that

$$\mathfrak{S}(\boldsymbol{k}) = 4\left(1-\frac{\nu_2(\boldsymbol{k})}{2}\right)\prod_{p>2}\left(1-\frac{1}{p}\right)^{-2}\left(1-\frac{\nu_p(\boldsymbol{k})}{p}\right).$$

Now $\nu_p(\boldsymbol{k})=1$ if $p|k$ and $\nu_p(\boldsymbol{k})=2$ if $p \nmid k$, and hence

$$\begin{aligned} \mathfrak{S}(\boldsymbol{k}) &= 2[2|k]\prod_{\substack{p|k \\ p>2}}\left(1-\frac{1}{p}\right)^{-1}\prod_{\substack{p\nmid k \\ p>2}}\left(1-\frac{1}{p}\right)^{-2}\left(1-\frac{2}{p}\right) \\ &= 2[2|k]\prod_{\substack{p|k \\ p>2}}\left(1-\frac{1}{p}\right)\left(1-\frac{2}{p}\right)^{-1}\prod_{p>2}\left(1-\frac{1}{p}\right)^{-2}\left(1-\frac{2}{p}\right) \\ &= 2[2|k]\prod_{\substack{p|k \\ p>2}}\left(\frac{p-1}{p-2}\right)\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right) \\ &= \mathfrak{S}_2(k). \end{aligned}$$

Next, in general,

$$\begin{aligned} \mathfrak{S}(\boldsymbol{k}) &= \prod_{p}\left(1-\frac{1}{p}\right)^{-r}\left(1-\frac{\nu_p(\boldsymbol{k})}{p}\right) \\ &= \left(1-\frac{1}{2}\right)^{-r}\left(1-\frac{1}{3}\right)^{-r}\left(1-\frac{\nu_2(\boldsymbol{k})}{2}\right)\left(1-\frac{\nu_3(\boldsymbol{k})}{3}\right)\prod_{p>3}\left(\frac{p}{p-1}\right)^{r-1}\left(\frac{p-\nu_p(\boldsymbol{k})}{p-1}\right). \end{aligned}$$

With $r=3$, $\boldsymbol{k}=(0,k_1,k_2)$, $\kappa=(k_1,k_2)$, and $\Delta=k_1k_2(k_2-k_1)$ we have

$$\nu_2(\boldsymbol{k}) = \begin{cases} 1, & \text{if } 2|\kappa\ , \\ 2, & \text{if } 2 \nmid \kappa, \end{cases}$$

and for $p \geq 3$

$$\nu_p(\boldsymbol{k}) = \begin{cases} 1, & \text{if } p|\kappa, \\ 2, & \text{if } p|\Delta,\ p \nmid \kappa, \\ 3, & \text{if } p \nmid \Delta. \end{cases}$$

We now see that $\mathfrak{S}(\boldsymbol{k}) = 0$ if $2 \not| \kappa$ or $3 \not| \Delta$, which by (2.5) proves the lemma in these cases. Thus we now assume that $2|\kappa$ and $3|\Delta$, and have

$$\begin{aligned}
\mathfrak{S}(\boldsymbol{k}) &= \frac{9}{\nu_3(\boldsymbol{k})}\prod_{p>3}\left(\frac{p}{p-1}\right)^2\left(\frac{p-\nu_p(\boldsymbol{k})}{p-1}\right) \\
&= \frac{9}{\nu_3(\boldsymbol{k})}\prod_{p>3}\left(\frac{p}{p-1}\right)^2\left(\frac{p-3}{p-1}\right)\prod_{\substack{p|\kappa\\ p>3}}\left(\frac{p-1}{p-3}\right)\prod_{\substack{p|\Delta\\ p\not|\kappa\\ p>3}}\left(\frac{p-2}{p-3}\right) \\
&= \frac{9}{\nu_3(\boldsymbol{k})}\prod_{p>3}\left(1-\frac{1}{(p-1)^2}\right)\left(1-\frac{2}{(p-1)(p-2)}\right)\prod_{\substack{p|\kappa\\ p>3}}\left(\frac{p-1}{p-2}\right)\prod_{\substack{p|\Delta\\ p>3}}\left(\frac{p-2}{p-3}\right) \\
&= 6C_2C_3H_2(\kappa)H_3(\Delta) \\
&= \mathfrak{S}_2(\kappa)\mathfrak{S}_3(\Delta),
\end{aligned}$$

which proves Lemma 2.5.

4. Proof of Theorem 1.3

We now prove that the mixed correlations can be reduced to the pure correlations through an application of the Bombieri-Vinogradov theorem. For $k \ge 2$ we consider the general sum, for $R \le P \le N$,

$$\mathcal{S}_P(N, \boldsymbol{j}) = \sum_{n=1}^{N} \Lambda_R(n+j_1)\Lambda_R(n+j_2)\cdots\Lambda_R(n+j_{k-1})\Lambda_P(n) \tag{4.1}$$

where the j_i's are not necessarily distinct, but satisfy

$$j_i \neq 0, \quad 1 \le i \le k-1. \tag{4.2}$$

Note that for $1 < n \le N$, $\Lambda_N(n) = \Lambda(n)$. We prove Theorem 1.3 by proving the following theorem which shows that both $\mathcal{S}_R(N, \boldsymbol{j})$ and $\mathcal{S}_N(N, \boldsymbol{j})$ in certain ranges are asymptotic to the same main term. Let

$$W_R(\boldsymbol{j}) = \sum_{\substack{d_1, d_2, \dots, d_{k-1} \le R \\ (d_i, j_i)=1,\, 1\le i\le k-1 \\ (d_r, d_s)|j_s - j_r,\, 1\le r<s\le k-1}} \frac{1}{\phi(D_{k-1})}\prod_{i=1}^{k-1}\mu(d_i)\log\frac{R}{d_i}, \tag{4.3}$$

where

$$D_{k-1} = [d_1, d_2, \dots, d_{k-1}]. \tag{4.4}$$

Theorem 4.1 *We have, for $k \geq 2$, $N^{\epsilon} \leq R \leq N^{\frac{1}{k}}$, and* $\max_i |j_i| \ll R^{\frac{1}{k}}$,

$$\mathcal{S}_R(N, \boldsymbol{j}) = NW_R(\boldsymbol{j}) + O(R^k) + O_k\Big(N(\log^{4^{k-1}+k-2} N)e^{-c_1\sqrt{\log \frac{R}{\max |j_i|^{k-1}}}}\Big), \tag{4.5}$$

and, for $N^{\epsilon} \leq R \leq N^{\frac{\vartheta}{k-1}-\epsilon}$,

$$\mathcal{S}_N(N, \boldsymbol{j}) = NW_R(\boldsymbol{j}) + O_k\big(\frac{N}{(\log N)^{\frac{A}{2}-4^{k-3/2}+2^{k-1}-1}}\big). \tag{4.6}$$

We have

$$\mathcal{S}_P(N, \boldsymbol{j}) = \sum_{d_1, d_2, \ldots, d_{k-1} \leq R} \Big(\prod_{i=1}^{k-1} \mu(d_i) \log \frac{R}{d_i}\Big)\Big(\sum_{\substack{n=1 \\ d_i | n+j_i,\ 1 \leq i \leq k-1}}^{N} \Lambda_P(n)\Big). \tag{4.7}$$

Let

$$T_P(N, \boldsymbol{j}) = \sum_{\substack{n=1 \\ d_i | n+j_i,\ 1 \leq i \leq k-1}}^{N} \Lambda_P(n). \tag{4.8}$$

The $k-1$ congruence relations $n \equiv -j_i (\operatorname{mod} d_i)$ will have no solutions unless $(d_r, d_s) | j_s - j_r$ for all $1 \leq r < s \leq k-1$. If these divisibility conditions hold, then by the Chinese remainder theorem there exists a unique solution to these congruences $n \equiv a (\operatorname{mod} [d_1, d_2, \ldots, d_{k-1}])$ for some $a = a(\boldsymbol{d}, \boldsymbol{j})$. Here a satisfies the original congruences $a \equiv -j_i (\operatorname{mod} d_i)$ for $1 \leq i \leq k-1$. Thus we have

$$\begin{aligned} T_P(N, \boldsymbol{j}) &= [(d_r, d_s) | j_s - j_r,\ 1 \leq r < s \leq k-1]\Big(\sum_{\substack{1 \leq n \leq N \\ n \equiv a (\operatorname{mod} D_{k-1})}} \Lambda_P(n)\Big) \\ &= [(d_r, d_s) | j_s - j_r,\ 1 \leq r < s \leq k-1] \psi_P(N; D_{k-1}, a). \end{aligned} \tag{4.9}$$

We next have that

$$\psi_P(N; D_{k-1}, a) = \sum_{d \leq P} \mu(d) \log \frac{P}{d} \sum_{\substack{1 \leq n \leq N \\ n \equiv a (\operatorname{mod} D_{k-1}) \\ n \equiv 0 (\operatorname{mod} d)}} 1.$$

The two congruences in this sum are solvable provided $(d, D_{k-1}) | a$, in which case we have that n runs through a residue class modulo $[D_{k-1}, d]$. Hence

$$\psi_P(N; D_{k-1}, a) = N \sum_{\substack{d \leq P \\ (d, D_{k-1}) | a}} \frac{\mu(d)}{[D_{k-1}, d]} \log \frac{P}{d} + O(P).$$

We now write $d = gd'$, where $g = (d, D_{k-1})$. Hence $[d, D_{k-1}] = D_{k-1}d'$, and by Lemma 2.1

$$\begin{aligned}\psi_P(N; D_{k-1}, a) &= \frac{N}{D_{k-1}} \sum_{\substack{g|D_{k-1},\ g|a \\ g\le P}} \mu(g) \sum_{\substack{d'\le \frac{P}{g} \\ (d', D_{k-1})=1}} \frac{\mu(d')}{d'} \log\frac{P}{gd'} + O(P) \\ &= \frac{N}{\phi(D_{k-1})} \sum_{\substack{g|D_{k-1},\ g|a \\ g\le P}} \mu(g) \\ &\qquad + O\Big(\frac{N}{D_{k-1}} \sum_{\substack{g|D_{k-1},\ g|a \\ g\le P}} \mu^2(g) e^{-c_1\sqrt{\log(P/g)}}\Big) + O(P).\end{aligned}$$

Since $a \equiv -j_i (\bmod d_i)$ for $1 \le i \le k-1$ and $g|a$ we see that $(g, d_i)|j_i$ for $1 \le i \le k-1$, and conversely since $g|D_{k-1}$ these divisibility conditions imply that $g|a$. Hence

$$\begin{aligned}\psi_P(N; D_{k-1}, a) &= \frac{N}{\phi(D_{k-1})} \sum_{\substack{g|D_{k-1} \\ g\le P \\ (g,d_i)|j_i,\, 1\le i\le k-1}} \mu(g) \\ &\qquad + O\Big(\frac{N}{D_{k-1}} \sum_{\substack{g|D_{k-1} \\ g\le P \\ (g,d_i)|j_i,\, 1\le i\le k-1}} \mu^2(g) e^{-c_1\sqrt{\log(P/g)}}\Big) + O(P). \qquad (4.10)\end{aligned}$$

The truncated Möbius function sum complicates the calculations of our pure correlations when one or more of the $j_i = 0$, but when all the $j_i \ne 0$ the truncation problem disappears. Thus, we see in this sum that $g \le \prod_{i=1}^{k-1} |j_i| \le \max |j_i|^{k-1}$, and hence provided

$$\max_{1\le i\le k-1} |j_i| \le P^{\frac{1}{k-1}} \qquad (4.11)$$

we have

$$\sum_{\substack{g|D_{k-1} \\ g\le P \\ (g,d_i)|j_i,\, 1\le i\le k-1}} \mu(g) = \sum_{\substack{g|D_{k-1} \\ (g,d_i)|j_i,\, 1\le i\le k-1}} \mu(g) = \begin{cases} 1, & \text{if } (d_i, j_i) = 1,\ 1 \le i \le k-1, \\ 0, & \text{otherwise.}\end{cases}$$

We conclude that subject to (4.11),

$$\begin{aligned}\psi_P(N; D_{k-1}, a) &= [(d_i, j_i) = 1,\ 1 \le i \le k-1] \frac{N}{\phi(D_{k-1})} \\ &\quad + O\Big(\frac{N}{D_{k-1}} d(D_{k-1}) e^{-c_1\sqrt{\log(P/\max|j_i|^{k-1})}}\Big) + O(P).\end{aligned} \qquad (4.12)$$

Hence by (4.3), (4.7), (4.8), (4.9), and (4.12), we obtain

$$\begin{aligned}\mathcal{S}_P(N, \boldsymbol{j}) =& NW_R(\boldsymbol{j}) + O(PR^{k-1}) \\ &+ O\Big(N \log^{k-1} N e^{-c_1\sqrt{\log\frac{P}{\max|j_i|^{k-1}}}} \sum_{d_1, d_2, \ldots, d_{k-1}\le R} \frac{d(D_{k-1})}{D_{k-1}}\Big).\end{aligned} \qquad (4.13)$$

We now estimate the sum in the second error term sufficiently well for our needs. Letting

$$D_R(n) = \sum_{\substack{e|n\\ e\le R}} d(e),$$

then

$$\begin{aligned}\sum_{n\le N} D_R(n)^m &= \sum_{n\le N} \sum_{\substack{e_1,e_2,\dots,e_m\le R\\ e_i|n,\,1\le i\le m}} d(e_1)d(e_2)\dots d(e_m)\\ &= \sum_{e_1,e_2,\dots,e_m\le R} d(e_1)d(e_2)\dots d(e_m) \sum_{\substack{n\le N\\ e_i|n,1\le i\le m}} 1\\ &= N \sum_{e_1,e_2,\dots,e_m\le R} \frac{d(e_1)d(e_2)\dots d(e_m)}{[e_1,e_2,\dots,e_m]} + O\Big(\big(\sum_{e\le R} d(e)\big)^m\Big).\end{aligned}$$

The last error term is $O(R^m \log^m R)$, and $D_R(n) \le \sum_{e|n} d(n) = d^2(n)$. Hence, using the estimate [19]

$$\sum_{m\le N} d(m)^k \ll_k N \log^{2^k-1} N, \tag{4.14}$$

we have

$$\sum_{n\le N} D_R(n)^m \le \sum_{n\le N} d(n)^{2m} \ll_m N \log^{4^m-1} N.$$

Thus

$$\begin{aligned}\sum_{d_1,d_2,\dots,d_m\le R} \frac{d(D_m)}{D_m} &\le \sum_{e_1,e_2,\dots,e_m\le R} \frac{d(e_1)d(e_2)\dots d(e_m)}{[e_1,e_2,\dots,e_m]}\\ &\ll_m \log^{4^m-1} N + \frac{R^m\log^m R}{N}.\end{aligned}$$

Hence we conclude

$$\begin{aligned}\mathcal{S}_P(N,\boldsymbol{j}) =& NW_R(\boldsymbol{j}) + O(PR^{k-1})\\ &+ O_k\bigg(\Big(N\log^{4^{k-1}+k-2} N + R^{k-1}\log^{2k-2} N\Big)e^{-c_1\sqrt{\log \frac{P}{\max|j_i|^{k-1}}}}\bigg).\end{aligned} \tag{4.15}$$

Taking $P = R$ proves the first part of Theorem 4.1. Equation (4.15) may also be useful when P is not too large but larger than R.

We next turn to the case $P = N$. In this case $\psi_P(N; q, a) = \psi(N; q, a) + O(\log N)$, the error term coming from $\Lambda_N(1)$. We apply (1.29) and have

$$\psi(N; D_{k-1}, a) = [(D_{k-1}, a) = 1]\frac{N}{\phi(D_{k-1})} + E(N; D_{k-1}, a).$$

The condition that $(D_{k-1}, a) = 1$ is equivalent to having $(d_i, a) = 1$ for $1 \le i \le k-1$, and since $a \equiv -j_i (\bmod d_i)$ these conditions are equivalent to $(d_i, j_i) = 1$ for $1 \le i \le k-1$. We conclude that

$$\psi(N; D_{k-1}, a) = [(d_i, j_i) = 1,\, 1 \le i \le k-1] \frac{N}{\phi(D_{k-1})} + E(N; D_{k-1}, a). \tag{4.16}$$

By (4.9) we thus obtain in place of (4.13)

$$\begin{aligned} \mathcal{S}_N(N, \boldsymbol{j}) = N W_R(\boldsymbol{j}) + O\Big(\log^{k-1} N \sum_{d_1, d_2, \dots, d_{k-1} \le R} \mu^2(D_{k-1}) |E(N; D_{k-1}, a)| \Big) \\ + O(\log N \log^{k-1} R \prod_{i=1}^{k-1} d(1 + j_i)), \end{aligned} \tag{4.17}$$

the last error term coming from the $n = 1$ term. This last error will be negligible since it is $\ll (\log N (\max_i |j_i|)^\epsilon)^k \ll_k N^\epsilon$ since $\max_i |j_i| \ll R^{\frac{1}{k}}$. For the sum in the error term, we have

$$\begin{aligned} &\sum_{d_1, d_2, \dots, d_{k-1} \le R} \mu^2(D_{k-1}) |E(N; D_{k-1}, a)| \\ &\qquad = \sum_{m \le R^{k-1}} \mu^2(m) \max_{a (\bmod m)} |E(N; m, a)| \sum_{\substack{m = D_{k-1} \\ d_1, d_2, \dots, d_{k-1} \le R}} 1 \\ &\qquad \le \sum_{m \le R^{k-1}} \mu^2(m) \max_{a (\bmod m)} |E(N; m, a)| d_l(m), \end{aligned}$$

where $l = 2^{k-1} - 1$. The factor of $d_l(m)$ arises since, given m, the number of solutions of $m = D_r$ is bounded by $d_{2^r - 1}(m)$, since the least common multiple of r squarefree numbers can always be expressed uniquely as the product of up to $2^r - 1$ numbers which are pairwise relatively prime, determined by exactly which of the original numbers $d_1, d_2, \cdots, d_r$ each factor divides. (We will use this decomposition in later sections.) Applying Cauchy's inequality we see the previous expression is

$$\ll \sqrt{\sum_{m \le R^{k-1}} \frac{d_l(m)^2}{m}} \sqrt{\sum_{m \le R^{k-1}} m \max_{a (\bmod m)} |E(N; m, a)|^2}.$$

We now use the generalization of (4.14)

$$\sum_{m \le N} d_r(m)^k \ll_k N \log^{r^k - 1} N, \tag{4.18}$$

and the trivial estimate $|E(N; m, a)| \ll \frac{N \log N}{m}$ to see the error term above is

$$\ll_l \sqrt{(\log R)^{l^2} N \log N \sum_{m \le R^{k-1}} \max_{a (\bmod m)} |E(N; m, a)|}.$$

We now apply (1.30) to conclude this error is, for $R^{k-1} \le N^{\vartheta-\epsilon}$,

$$\ll_k \frac{N}{(\log N)^{\frac{A}{2}-4^{k-3/2}+2^{k-1}-1}},$$

which finishes the proof of Theorem 4.1.

5. Pair Correlation of $\Lambda_R(n)$

We first prove the case $k = 1$ of Theorem 1.1. We have

$$\begin{aligned}
\mathcal{S}_1(N, j, (1)) &= \sum_{n\le N} \Lambda_R(n+j) = \sum_{d\le R} \mu(d) \log \frac{R}{d} \sum_{\substack{1\le n\le N \\ d|n+j}} 1 \\
&= N \sum_{d\le R} \frac{\mu(d)}{d} \log \frac{R}{d} + O(R) \\
&= N + O(\frac{N}{(\log R)^A}) + O(R) \qquad (5.1)
\end{aligned}$$

by Lemma 2.1. This proves Theorem 1.1 in this case.

We now examine the case $k = 2$ of Theorem 1.1. These results have been proved before in [8]. The proof we give here models the procedure we will use for higher correlations without any of the truncation complications which will arise there. In view of the comment following Theorem 1.1, we need to consider

$$\mathcal{S}_2(k) = \sum_{n=1}^{N} \Lambda_R(n)\Lambda_R(n+k), \qquad (5.2)$$

In our earlier notation, we have $\mathcal{S}_2(k) = \mathcal{S}_2(N, (0, k), (1, 1))$ if $k \ne 0$, and $\mathcal{S}_2(0) = \mathcal{S}_2(N, (0), (2))$.

Theorem 5.1 *We have*

$$\mathcal{S}_2(0) = N \log R + O(N) + O(R^2), \qquad (5.3)$$

and for $0 < |k| \le R$ *and any* $A > 0$*, we have*

$$\mathcal{S}_2(k) = \mathfrak{S}_2(k)N + O(\frac{k}{\phi(k)} \frac{N}{(\log 2R/k)^A}) + O(R^2). \qquad (5.4)$$

Graham [13] has proved (5.3) for $1 \le R \le N$ with the error term $O(R^2)$ removed. By Theorem 5.1 we see that Theorem 1.1 is true for $k = 2$, with $\mathcal{C}(2) = 1$ and $\mathcal{C}(1, 1) = 1$.

Proof of Theorem 5.1. Applying the definition of $\Lambda_R(n)$, we have

$$\mathcal{S}_2(k) = \sum_{d,e\le R} \mu(d)\log(R/d)\mu(e)\log(R/e) \sum_{\substack{n\le N\\ d|n,e|n+k}} 1.$$

In the inner sum the two divisibility conditions imply that n will run through a residue class modulo $[d,e]$ provided $(d,e)|k$, and there will be no solution for n otherwise. Therefore we have

$$\begin{aligned}\mathcal{S}_2(k) &= N\sum_{\substack{d,e\le R\\ (d,e)|k}} \frac{\mu(d)\log(R/d)\mu(e)\log(R/e)}{[d,e]} + O\Big(\sum_{d,e\le R}\log(R/d)\log(R/e)\Big)\\ &= NT_2(k) + O(R^2).\end{aligned} \tag{5.5}$$

To evaluate $T_2(k)$ we break the sum into relatively prime summands in order to handle $[d,e]$. We let $d=a_1b_{12}$ and $e=a_2b_{12}$ where $b_{12}=(d,e)$ so that a_1, a_2, and b_{12} are pairwise relatively prime. For higher correlations this decomposition notation will be used as well. Hence we have

$$\begin{aligned} T_2(k) &= \sideset{}{'}\sum_{\substack{a_1b_{12}\le R\\ a_2b_{12}\le R\\ b_{12}|k}} \frac{\mu^2(b_{12})}{b_{12}}\frac{\mu(a_1)}{a_1}\frac{\mu(a_2)}{a_2}\log\frac{R}{a_1b_{12}}\log\frac{R}{a_2b_{12}}\\ &= \sideset{}{'}\sum_{\substack{a_2b_{12}\le R\\ b_{12}|k}} \frac{\mu^2(b_{12})}{b_{12}}\frac{\mu(a_2)}{a_2}\log\frac{R}{a_2b_{12}} \sum_{\substack{a_1\le R/b_{12}\\ (a_1,a_2b_{12})=1}} \frac{\mu(a_1)}{a_1}\log\frac{R/b_{12}}{a_1},\end{aligned}$$

where the prime on the summation indicates that all the summands are relatively prime to each other. We now apply Lemma 2.1 with $j=0$ to obtain for any $B>0$

$$T_2(k) = \sideset{}{'}\sum_{\substack{a_2b_{12}\le R\\ b_{12}|k}} \frac{\mu^2(b_{12})}{\phi(b_{12})}\frac{\mu(a_2)}{\phi(a_2)}\log\frac{R}{a_2b_{12}} + O\left(\sideset{}{'}\sum_{\substack{a_2b_{12}\le R\\ b_{12}|k}} \frac{\mu^2(b_{12})\mu^2(a_2)\log\frac{R/b_{12}}{a_2}}{b_{12}a_2\log^B(2R/b_{12})}\right).$$

For the main term above we sum over a_2 and apply Lemma 2.1 again with $j=1$ to see this term is equal to

$$\sum_{\substack{b_{12}\le R\\ b_{12}|k}} \frac{\mu^2(b_{12})}{\phi(b_{12})}\mathfrak{S}_2(b_{12}) + O\Big(\sum_{\substack{b_{12}\le R\\ b_{12}|k}} \frac{\mu^2(b_{12})}{\phi(b_{12})\log^B(2R/b_{12})}\Big).$$

Summing over a_2 in the error term in the formula for $T_2(k)$ above, we conclude

$$T_2(k) = \sum_{\substack{b_{12}\le R\\ b_{12}|k}} \frac{\mu^2(b_{12})}{\phi(b_{12})}\mathfrak{S}_2(b_{12}) + O\Big(\sum_{\substack{b_{12}\le R\\ b_{12}|k}} \frac{\mu^2(b_{12})}{\phi(b_{12})\log^{B-2}(2R/b_{12})}\Big). \tag{5.6}$$

We now consider two cases. If $k = 0$ the main term is

$$\sum_{b_{12}\le R} \frac{\mu^2(b_{12})}{\phi(b_{12})}\mathfrak{S}_2(b_{12}) = \log R + O(1)$$

by Lemma 2.4 with $j = 1$, and the error term in (5.6) is also $O(1)$. This proves (5.3). If $0 < |k| \le R$ the main term is

$$\sum_{b_{12}|k} \frac{\mu^2(b_{12})}{\phi(b_{12})}\mathfrak{S}_2(b_{12}) = \mathfrak{S}_2(k)$$

by the second part of Lemma 2.4, and the error term is

$$\ll \frac{1}{(\log 2R/k)^{B-2}} \sum_{b_{12}|k} \frac{\mu^2(b_{12})}{\phi(b_{12})} = \frac{k}{\phi(k)(\log 2R/k)^{B-2}},$$

which proves (5.4).

It is worth noting that we can also give a very short proof of (5.4) using Theorem 4.1 from the last section. With $\boldsymbol{j} = (0, k)$, we have

$$\begin{aligned} W_R(\boldsymbol{j}) &= \sum_{\substack{d_1\le R\\ (d_1,k)=1}} \frac{\mu(d_1)}{\phi(d_1)} \log(R/d_1) \\ &= \mathfrak{S}_2(k) + O(e^{-c_1\sqrt{\log R}}) \qquad (5.7) \end{aligned}$$

by Lemma 2.1. By Theorem 4.1 this proves (5.4) and also evaluates the mixed second correlation as well.

6. Triple correlation for $\Lambda_R(n)$

To prove Theorem 1.1 when $k = 3$ we need to evaluate the sums

$$\mathcal{S}_3(k) = \sum_{n=1}^{N} {\Lambda_R}^2(n)\Lambda_R(n + k), \qquad (6.1)$$

and, for non-zero $k_1 \ne k_2$,

$$\mathcal{S}_3(k_1, k_2) = \sum_{n=1}^{N} \Lambda_R(n)\Lambda_R(n + k_1)\Lambda_R(n + k_2) \qquad (6.2)$$

In the notation of Theorem 1.1 we have $\mathcal{S}_3(0)=\mathcal{S}_3(N, (0), (3))$, $\mathcal{S}_3(k)=\mathcal{S}_3(N, (0, k), (2, 1))$ if $k \ne 0$, and $\mathcal{S}_3(k_1, k_2) = \mathcal{S}_3(N, (0, k_1, k_2), (1, 1, 1))$ for non-zero $k_1 \ne k_2$. We will obtain the following results on these correlations.

Theorem 6.1 *We have*

$$\mathcal{S}_3(0) = \frac{3}{4} N \log^2 R + O(N \log N (\log\log R)^{18}) + O(R^3), \tag{6.3}$$

and, for $k \neq 0$, $|k| \leq R^{\frac{1}{2}-\epsilon}$,

$$\mathcal{S}_3(k) = \mathfrak{S}_2(k) N \log R + O(N(\log\log R)^{13}) + O(R^3), \tag{6.4}$$

and letting $\boldsymbol{k} = (0, k_1, k_2)$, $k_1 \neq k_2 \neq 0$, *if* $(k^*)^2 < R/2$, *where* $k^* = \max(|k_1|, |k_2|)$, *then*

$$\mathcal{S}_3(k_1, k_2) = \mathfrak{S}(\boldsymbol{k}) N + O\left(N e^{-c_1 \sqrt{\log\left(\frac{R}{2(k^*)^2}\right)}} \log^8 R \right) + O(R^3). \tag{6.5}$$

We consider the general situation and specialize later. Let

$$\mathcal{S}_3(k_1, k_2, k_3) = \sum_{n=1}^{N} \Lambda_R(n+k_1)\Lambda_R(n+k_2)\Lambda_R(n+k_3) \tag{6.6}$$

Expanding, we have

$$\mathcal{S}_3(k_1, k_2, k_3) = \sum_{d_1, d_2, d_3 \leq R} \mu(d_1)\log(R/d_1)\mu(d_2)\log(R/d_2)\mu(d_3)\log(R/d_3) \sum_{\substack{n \leq N \\ d_1|n+k_1 \\ d_2|n+k_2 \\ d_3|n+k_3}} 1.$$

The sum over n is zero unless $(d_1, d_2)|k_2 - k_1$, $(d_1, d_3)|k_3 - k_1$, and $(d_2, d_3)|k_3 - k_2$, in which case the sum runs through a residue class modulo $[d_1, d_2, d_3]$, and we have

$$\sum_{\substack{n \leq N \\ d_1|n+k_1 \\ d_2|n+k_2 \\ d_3|n+k_3}} 1 = \frac{N}{[d_1, d_2, d_3]} + O(1).$$

We conclude

$$\begin{aligned} \mathcal{S}_3(k_1, k_2, k_3) &= N \sum_{\substack{d_1, d_2, d_3 \leq R \\ (d_1, d_2)|k_2 - k_1 \\ (d_1, d_3)|k_3 - k_1 \\ (d_2, d_3)|k_3 - k_2}} \frac{\mu(d_1)\log(R/d_1)\mu(d_2)\log(R/d_2)\mu(d_3)\log(R/d_3)}{[d_1, d_2, d_3]} + O(R^3) \\ &= N T_3(k_1, k_2, k_3) + O(R^3). \end{aligned} \tag{6.7}$$

We now decompose d_1, d_2, and d_3 into relatively prime factors

$$\begin{aligned} d_1 &= a_1b_{12}b_{13}a_{123} \\ d_2 &= a_2b_{12}b_{23}a_{123} \\ d_3 &= a_3b_{13}b_{23}a_{123} \end{aligned}$$

where a_χ or b_χ is a divisor of the d_i's where i occurs in χ. Since the d_i's are squarefree, these new variables are pairwise relatively prime. The letters a and b reflect the parity of the number of d_i's that the new variable divides. We will let $\mathcal{D}$ denote the set of a_χ's and b_χ's which satisfy the conditions

$$\begin{aligned} a_1b_{12}b_{13}a_{123} &\leq R \\ a_2b_{12}b_{23}a_{123} &\leq R \\ a_3b_{13}b_{23}a_{123} &\leq R \end{aligned} \tag{6.8}$$

$$b_{12}a_{123}|k_2-k_1, \quad b_{13}a_{123}|k_3-k_1, \qquad b_{23}a_{123}|k_3-k_2. \tag{6.9}$$

Letting

$$L_i(R) = \log\frac{R}{d_i},$$

we have

$$\begin{aligned} T_3(k_1,k_2,k_3) &= \sum_{\mathcal{D}}{}' \frac{\mu(a_1)\mu(a_2)\mu(a_3)\mu^2(b_{12})\mu^2(b_{13})\mu^2(b_{23})\mu(a_{123})}{a_1a_2a_3b_{12}b_{13}b_{23}a_{123}} L_1(R)L_2(R)L_3(R) \\ &= \sum_{\mathcal{D}}{}' f_R(d_1,d_2,d_3). \end{aligned} \tag{6.10}$$

We now will sum over a_1, a_2, and a_3 using Lemma 2.1 and Lemma 2.3. In order to apply these lemmas we need each a_i to range over a long enough interval, and therefore we need to restrict the ranges of some of the other variables. The excluded ranges will later be shown to make a lower order contribution. If D is a product of some of the variables in $\mathcal{D}$, we let $\mathcal{D}(D)$ denote the subset of $\mathcal{D}$ where the variables not occuring in D are eliminated from the inequalities in (6.8) and divisibility conditions in (6.9). Thus, letting $D_1 = a_2a_3b_{12}b_{13}b_{23}a_{123}$, we have that $\mathcal{D}(D_1)$ no longer includes the variable a_1 and we take $a_1 = 1$ in (6.8). We now obtain on summing over a_1 using Lemma 2.1 and taking $R_1 < R$,

$$\begin{aligned} T_3(k_1,k_2,k_3) &= \sum_{\substack{\mathcal{D} \\ b_{12}b_{13}a_{123}\leq R_1}}{}' f_R(d_1,d_2,d_3) + \sum_{\substack{\mathcal{D} \\ R_1<b_{12}b_{13}a_{123}\leq R}}{}' f_R(d_1,d_2,d_3) \\ &= \sum_{\substack{\mathcal{D}(D_1) \\ b_{12}b_{13}a_{123}\leq R_1}}{}' \frac{\mu(a_2)\mu(a_3)\mu(a_{123})\mu^2(D_1)}{\phi(D_1)} L_2(R)L_3(R) + E_1(R) \\ &+ \sum_{\substack{\mathcal{D} \\ R_1<b_{12}b_{13}a_{123}\leq R}}{}' f_R(d_1,d_2,d_3) \\ &= U_3(k_1,k_2,k_3) + E_1(R) + \mathcal{E}_f(\mathcal{D}), \end{aligned} \tag{6.11}$$

where

$$\begin{aligned} E_1(R) &= \sideset{}{'}\sum_{\substack{\mathcal{D} \\ b_{12}b_{13}a_{123}\le R_1}} \frac{\mu(a_2)\mu(a_3)\mu(a_{123})\mu^2(D_1)}{D_1} L_2(R)L_3(R)r_0(\frac{R}{b_{12}b_{13}a_{123}}, D_1) \\ &\ll e^{-c_1\sqrt{\log(R/R_1)}}\log^8 R. \end{aligned}$$

Hence

$$T_3(k_1,k_2,k_3) = U_3(k_1,k_2,k_3) + O(e^{-c_1\sqrt{\log(R/R_1)}}\log^8 R) + \mathcal{E}_f(\mathcal{D}). \qquad (6.12)$$

Denote the summand for $U_3(k_1,k_2,k_3)$ by $g_R(d_1,d_2,d_3)$, which does not depend on a_1. Because of the symmetry in our original variables in (6.10), we could equally well have summed over a_2 or a_3 above and obtained the same expression for $g_R(d_1,d_2,d_3)$ with the appropriate change in variables and renumbering of the $k_i's$. We will later make use of this fact for some of our error terms, and will let the summation conditions $\mathcal{D}(D)$ determine which variables appear in g_R and subsequent summands. Returning to (6.11), we obtain on summing over a_2 using Lemma 2.1 that, with $D_2 = a_3b_{12}b_{13}b_{23}a_{123}$ and $R_2 < R$,

$$\begin{aligned} U_3(k_1,k_2,k_3) &= \sideset{}{'}\sum_{\substack{\mathcal{D}(D_1) \\ b_{12}b_{13}a_{123}\le R_1 \\ b_{12}b_{23}a_{123}\le R_2}} g_R(d_1,d_2,d_3) + \sideset{}{'}\sum_{\substack{\mathcal{D}(D_1) \\ b_{12}b_{13}a_{123}\le R_1 \\ R_2<b_{12}b_{23}a_{123}\le R}} g_R(d_1,d_2,d_3) \\ &= \sideset{}{'}\sum_{\substack{\mathcal{D}(D_2) \\ b_{12}b_{13}a_{123}\le R_1 \\ b_{12}b_{23}a_{123}\le R_2}} \frac{\mu(a_3)\mu(a_{123})\mu^2(D_2)}{\phi(D_2)} L_3(R)\mathfrak{S}_2(D_2) \\ &+ O(e^{-c_1\sqrt{\log(R/R_2)}}\log^6 R) + \sideset{}{'}\sum_{\substack{\mathcal{D}(D_1) \\ b_{12}b_{13}a_{123}\le R_1 \\ R_2<b_{12}b_{23}a_{123}\le R}} g_R(d_1,d_2,d_3) \\ &= V_3(k_1,k_2,k_3) + O(e^{-c_1\sqrt{\log(R/R_2)}}\log^6 R) + \mathcal{E}_g(\mathcal{D}(D_1)). \qquad (6.13) \end{aligned}$$

Finally, we denote the summand in $V_3(k_1,k_2,k_3)$ by $h_R(d_1,d_2,d_3)$ and we let $D_3 = b_{12}b_{13}b_{23}a_{123}$ and $R_3 < R$. Then by Lemma 2.3 with $j=1$ we obtain

$$\begin{aligned}
V_3(k_1,k_2,k_3) &= \sideset{}{'}\sum_{\substack{\mathcal{D}(D_2)\\ b_{12}b_{13}a_{123}\le R_1\\ b_{12}b_{23}a_{123}\le R_2\\ b_{13}b_{23}a_{123}\le R_3}} h_R(d_1,d_2,d_3) + \sideset{}{'}\sum_{\substack{\mathcal{D}(D_2)\\ b_{12}b_{13}a_{123}\le R_1\\ b_{12}b_{23}a_{123}\le R_2\\ R_3<b_{13}b_{23}a_{123}\le R}} h_R(d_1,d_2,d_3)\\
&= -\sideset{}{'}\sum_{\substack{\mathcal{D}(D_3)\\ b_{12}b_{13}a_{123}\le R_1\\ b_{12}b_{23}a_{123}\le R_2\\ b_{13}b_{23}a_{123}\le R_3}} \frac{\mu^2(D_3)\mu(a_{123})\mu((D_3,2))}{\phi(D_3)}\mathfrak{S}_2(2D_3)\mathfrak{S}_3(2D_3)\\
&+ O(e^{-c_1\sqrt{\log(R/R_3)}}\log^4 R) + \sideset{}{'}\sum_{\substack{\mathcal{D}(D_2)\\ b_{12}b_{13}a_{123}\le R_1\\ b_{12}b_{23}a_{123}\le R_2\\ R_3<b_{13}b_{23}a_{123}\le R}} h_R(d_1,d_2,d_3)\\
&= W_3(k_1,k_2,k_3) + O(e^{-c_1\sqrt{\log(R/R_3)}}\log^4 R) + \mathcal{E}_h(\mathcal{D}(D_2)). \qquad (6.14)
\end{aligned}$$

We now prove Theorem 6.1 by considering each case separately. We first prove (6.5) which is the case where the error terms are the easiest to handle.

7. Evaluation of $S_3(k_1,k_2)$

We consider $\mathcal{S}_3(k_1,k_2)$ by taking $k_1\ne k_2\ne 0$ and $k_3=0$ in (6.9) which therefore becomes

$$b_{12}a_{123}|k_2-k_1, \quad b_{13}a_{123}|k_1, \quad b_{23}a_{123}|k_2. \qquad (7.1)$$

These conditions imply, letting $k^* = \max(|k_1|,|k_2|)$, that

$$b_{12}b_{13}a_{123} < 2(k^*)^2, \quad b_{12}b_{23}a_{123} < 2(k^*)^2, \quad b_{13}b_{23}a_{123} < 2(k^*)^2. \qquad (7.2)$$

Hence, taking $R_1=R_2=R_3=2(k^*)^2<R$, we see that the error terms $\mathcal{E}_f$, $\mathcal{E}_g$, and $\mathcal{E}_h$ in (6.11), (6.13), and (6.14) are identically zero, and therefore

$$T_3(k_1,k_2) = W_3(k_1,k_2) + O(e^{-c_1\sqrt{\log\left(\frac{R}{2(k^*)^2}\right)}}\log^8 R). \qquad (7.3)$$

Now that the variables a_1, a_2, and a_3 have been eliminated, the bounds on the variables in $\mathcal{D}$ are automatically satisfied from (7.2), and, provided $(k^*)^2 < R/2$, we have

$$W_3(k_1,k_2) = -\sideset{}{'}\sum_{\substack{b_{13}a_{123}|k_1\\ b_{23}a_{123}|k_2\\ b_{12}a_{123}|k_2-k_1}} \frac{\mu^2(D_3)\mu(a_{123})\mu((D_3,2))}{\phi(D_3)}\mathfrak{S}_2(2D_3)\mathfrak{S}_3(2D_3). \qquad (7.4)$$

This sum is over square-free divisors, and therefore we let

$$k_1 = s_1K_1, \quad k_2 = s_2K_2, \tag{7.5}$$

where K_1 and K_2 are the largest square-free divisors of k_1 and k_2, and let

$$\kappa = (k_1, k_2), \quad K_{12} = (K_1, K_2). \tag{7.6}$$

Then we may rewrite W_3 as

$$W_3(k_1, k_2) = -\sideset{}{'}\sum_{\substack{b_{13}a_{123}|K_1\\ b_{23}a_{123}|K_2\\ b_{12}a_{123}|k_2-k_1}} \frac{\mu^2(D_3)\mu(a_{123})\mu((D_3, 2))}{\phi(D_3)}\mathfrak{S}_2(2D_3)\mathfrak{S}_3(2D_3). \tag{7.7}$$

The proof of (6.5) will follow from (6.7), (7.3) and Lemma 2.5 once we prove that

$$W_3(k_1, k_2) = \mathfrak{S}_2(K_{12})\mathfrak{S}_3(\Delta) = \mathfrak{S}_2(\kappa)\mathfrak{S}_3(\Delta) = \mathfrak{S}(\boldsymbol{k}). \tag{7.8}$$

We now let

$$K_1 = K_{12}j_1, \quad K_2 = K_{12}j_2, \quad k_2 - k_1 = K_{12}(s_2j_2 - s_1j_1) \tag{7.9}$$

where K_{12}, j_1, and j_2 are square-free and pairwise relatively prime ($s_2j_2 - s_1j_1$ may not be square-free or relatively prime with K_{12}.) Next, let

$$b_{13} = c_{13}d_{13}, \quad b_{23} = c_{23}d_{23} \tag{7.10}$$

where $c_{13}, c_{23}|K_{12}$, $d_{13}|j_1$, $d_{23}|j_2$, and c_{13}, d_{13}, c_{23}, d_{23}, b_{12}, a_{123} are thus all pairwise relatively prime. We also see that $a_{123}|K_{12}$. Thus, with $D_3 = b_{12}c_{13}c_{23}d_{13}d_{23}a_{123}$,

$$W_3(k_1, k_2) = -\sideset{}{'}\sum_{c_{13}c_{23}a_{123}|K_{12}} \mu(a_{123}) \sideset{}{'}\sum_{\substack{b_{12}|k_2-k_1\\ d_{13}|j_1\\ d_{23}|j_2}} \frac{\mu^2(D_3)\mu((D_3, 2))}{\phi(D_3)}\mathfrak{S}_2(2D_3)\mathfrak{S}_3(2D_3). \tag{7.11}$$

We first sum over d_{13} in the inner sum. To do this, we take $D_4 = b_{12}c_{13}c_{23}d_{23}a_{123}$ and have

$$\begin{aligned}
&\sum_{\substack{d_{13}|j_1\\ (d_{13},D_4)=1}} \frac{\mu^2(D_3)\mu((D_3, 2))}{\phi(D_3)}\mathfrak{S}_2(2D_3)\mathfrak{S}_3(2D_3)\\
&\quad = 2C_2\frac{\mu^2(D_4)}{\phi(D_4)} \sum_{\substack{d_{13}|j_1\\ (d_{13},D_4)=1}} \frac{\mu^2(d_{13})}{\phi(d_{13})}\mu((d_{13}D_4, 2))H_2(2d_{13}D_4)\mathfrak{S}_3(2d_{13}D_4).
\end{aligned}$$

We now break the sum on the right into two sums according to whether d_{13} is even or odd, in the former case we let $d_{13} = 2d$, and on using Lemma 2.4 we obtain that the right-hand side is

$$= \frac{2C_2\mu^2(D_4)H_2(D_4)}{\phi(D_4)}\Bigg\{ -[(D_4,2)=1]\sum_{\substack{2d|j_1\\(d,2D_4)=1}}\frac{\mu^2(d)}{\phi_2(d)}\mathfrak{S}_3(2dD_4)$$

$$+\mu((D_4,2))\sum_{\substack{d_{13}|j_1\\(d_{13},2D_4)=1}}\frac{\mu^2(d_{13})}{\phi_2(d_{13})}\mathfrak{S}_3(2d_{13}D_4)\Bigg\}$$

$$= \frac{2C_2\mu^2(D_4)H_2(D_4)}{\phi(D_4)}\Big(-[(D_4,2)=1][2|j_1]+\mu((D_4,2))\Big)\mathfrak{S}_3(2D_4j_1).$$

We will denote

$$B(d) = \frac{2C_2\mu^2(d)H_2(d)}{\phi(d)}. \tag{7.12}$$

Now substituting into (7.11), and letting $D_5 = b_{12}c_{13}c_{23}a_{123}$, the sum over d_{23} is equal to

$$B(D_5)\sum_{\substack{d_{23}|j_2\\(d_{23},D_5)=1}}\frac{\mu^2(d_{23})H_2(d_{23})}{\phi(d_{23})}\Big(-[(d_{23}D_5,2)=1][2|j_1]+\mu((d_{23}D_5,2))\Big)\mathfrak{S}_3(2d_{23}D_5j_1).$$

Since $(j_1,j_2)=1$, and $d_{23}|j_2$, we may replace the condition $(d_{23},D_5)=1$ in the sum by (d_{23},j_1D_5) and divide the sum into two sums with even or odd terms as above to see that this expression is

$$= -B(D_5)[(D_5,2)=1][2|j_2]\sum_{\substack{2d|j_2\\(d,2j_1D_5)=1}}\frac{\mu^2(d)}{\phi_2(d)}\mathfrak{S}_3(2dD_5j_1)$$

$$+B(D_5)\Big(-[(D_5,2)=1][2|j_1]+\mu((D_5,2))\Big)\sum_{\substack{d_{23}|j_2\\(d_{23},2j_1D_5)=1}}\frac{\mu^2(d_{23})}{\phi_2(d_{23})}\mathfrak{S}_3(2d_{23}D_5j_1)$$

$$= B(D_5)\Big(-[(D_5,2)=1]\big([2|j_1]+[2|j_2]\big)+\mu((D_5,2))\Big)\mathfrak{S}_3(2D_5j_1j_2).$$

Now let

$$\Delta = k_1k_2(k_2-k_1) = s_1s_2{K_{12}}^2j_1j_2(k_2-k_1) = s_1s_2{K_{12}}^3j_1j_2(s_2j_2-s_1j_1). \tag{7.13}$$

We substitute the last result into (7.11) and sum over b_{12}. Let $D_6 = c_{13}c_{23}a_{123}$. We claim that the relatively prime condition $(b_{12},D_6)=1$ may be replaced by $(b_{12},D_6j_1j_2)=1$. To see this, note that j_1, j_2, and K_{12} are pairwise relatively prime, and further $(s_1,j_2)=1$ and $(s_2,j_1)=1$. Thus

$$(k_2-k_1,j_1) = (K_{12}(s_2j_2-s_1j_1),j_1) = (s_2,j_1) = 1$$

and similarly $(k_2 - k_1, j_2) = 1$. Hence, since $b_{12}|k_2 - k_1$, we have $(b_{12}, j_1 j_2) = 1$. Now summing over b_{12} our sum is

$$= B(D_6) \sum_{\substack{b_{12}|k_2-k_1 \\ (b_{12},D_6 j_1 j_2)=1}} \Bigg(- [(b_{12}D_6, 2) = 1]\big([2|j_1] + [2|j_2]\big) + \mu((b_{12}D_6, 2))\Bigg) \times \frac{\mu^2(b_{12})H_2(b_{12})}{\phi(b_{12})} \mathfrak{S}_3(2b_{12}D_6 j_1 j_2)$$

$$= B(D_6) \sum_{\substack{2b|k_2-k_1 \\ (b,2D_6 j_1 j_2)=1}} \frac{\mu^2(b)}{\phi_2(b)} \Big(- [2|k_2 - k_1][(D_6, 2) = 1]\Big) \mathfrak{S}_3(2bD_6 j_1 j_2)$$

$$+B(D_6) \sum_{\substack{b_{12}|k_2-k_1 \\ (b_{12},2D_6 j_1 j_2)=1}} \frac{\mu^2(b_{12})}{\phi_2(b_{12})} \Big(- [(D_6, 2) = 1]\big([2|j_1] + [2|j_2]\big) + \mu((D_6, 2))\Big) \mathfrak{S}_3(2b_{12}D_6 j_1 j_2)$$

$$= B(D_6)\Big(- [(D_6, 2) = 1]\big([2|j_1]+[2|j_2]+[2|k_2 - k_1]\big)+\mu((D_6, 2))\Big) \mathfrak{S}_3(2D_6 j_1 j_2(k_2 - k_1)).$$

Now $D_6|K_{12}|\Delta$, and hence

$$\mathfrak{S}_3(2D_6 j_1 j_2(k_2 - k_1)) = \mathfrak{S}_3(2D_6\Delta) = \mathfrak{S}_3(\Delta).$$

We conclude that

$$W_3(k_1, k_2) = -\mathfrak{S}_3(\Delta) \sideset{}{'}\sum_{D_6|K_{12}} \mu(a_{123})B(D_6)\Big(-[(D_6, 2)=1]\big([2|j_1]+[2|j_2]+[2|k_2 - k_1]\big)+\mu((D_6, 2))\Big). \tag{7.14}$$

If K_{12} is odd, then exactly one of the variables j_1, j_2, and $k_2 - k_1$ is even and the other two are odd. If K_{12} is even, then j_1 and j_2 are odd and $k_2 - k_1$ is even. Hence in either case

$$[2|j_1] + [2|j_2] + [2|k_2 - k_1] = 1,$$

and therefore

$$W_3(k_1, k_2) = -\mathfrak{S}_3(\Delta) \sideset{}{'}\sum_{D_6|K_{12}} \mu(a_{123})B(D_6)\big(- [(D_6, 2) = 1] + \mu((D_6, 2))\big). \tag{7.15}$$

When K_{12} is odd the expression in parentheses is zero, and hence

$$W_3(k_1, k_2) = 0, \quad \text{if } K_{12} \text{ is odd.}$$

If K_{12} is even, then the expression in parentheses in (7.15) is zero if D_6 is odd, and is equal to -1 when D_6 is even. Hence, we conclude

$$W_3(k_1, k_2) = [2|K_{12}]\mathfrak{S}_3(\Delta) \sideset{}{'}\sum_{\substack{D_6|K_{12} \\ 2|D_6}} \mu(a_{123})B(D_6). \tag{7.16}$$

We could now evaluate this sum as before by summing over each variable in turn, but there is an easier approach, based on the observation that if a square-free number is a product of some factors, then necessarily those factors must be relatively prime with each other. Let $z = vwy$, and $\mathcal{A}$ be a set of natural numbers. Then for any arithmetic function $a(z)$ we have

$$\begin{aligned} {\sum_{z\in\mathcal{A}}}' \mu^2(z)a(z)\mu(y) &= \sum_{z\in\mathcal{A}} \mu^2(z)a(z) \sum_{w|z} \sum_{y|\frac{z}{w}} \mu(y) \\ &= \sum_{z\in\mathcal{A}} \mu^2(z)a(z) \sum_{\substack{w|z \\ w=z}} 1 \\ &= \sum_{z\in\mathcal{A}} \mu^2(z)a(z). \end{aligned} \tag{7.17}$$

The sum in (7.16) is of this form, and therefore we have

$$\begin{aligned} W_3(k_1,k_2) &= [2|K_{12}]\mathfrak{S}_3(\Delta) \sum_{\substack{z|K_{12} \\ 2|z}} B(z) \\ &= 2C_2[2|K_{12}]\mathfrak{S}_3(\Delta) \sum_{\substack{2z'|K_{12} \\ (z',2)=1}} \frac{\mu^2(z')}{\phi_2(z')} \\ &= 2C_2[2|K_{12}]\mathfrak{S}_3(\Delta)H_2(K_{12}), \end{aligned}$$

where we used (2.3) in the last line. We conclude by (2.5) that

$$W_3(k_1,k_2) = \mathfrak{S}_2(K_{12})\mathfrak{S}_3(\Delta) = \mathfrak{S}_2(\kappa)\mathfrak{S}_3(\Delta), \tag{7.18}$$

which completes the proof of (7.8).

8. Evaluation of $S_3(k)$, $k \neq 0$

We now take $k_1 = k \neq 0$, $k_2 = k_3 = 0$ so that the divisibility conditions in (6.9) become

$$b_{12}a_{123}|k, \quad b_{13}a_{123}|k,$$

and since our variables are relatively prime

$$b_{12}b_{13}a_{123}|k. \tag{8.1}$$

Taking $R_1 = k < R$ we see that $\mathcal{E}_f$ in (6.12) is identically zero. We will take $R_2 = R_3 \geq k^2$ to be chosen later as a function of R, and conclude from (6.12), (6.13), and (6.14) that

$$T_3(k) = W_3(k) + \mathcal{E}_g(\mathcal{D}(D_1)) + \mathcal{E}_h(\mathcal{D}(D_2)) + O(e^{-c_1\sqrt{\log(R/R_2)}}\log^8 R). \tag{8.2}$$

Now, summing over a_3 using Lemma 2.1, we have

$$\begin{aligned}\mathcal{E}_g(\mathcal{D}(D_1)) &= \sum\nolimits'_{\substack{\mathcal{D}(D_1)\\ b_{12}b_{13}a_{123}\le R_1\\ R_2<b_{12}b_{23}a_{123}\le R\\ b_{13}b_{23}a_{123}\le R_3}} g_R(d_1,d_2,d_3) + \sum\nolimits'_{\substack{\mathcal{D}(D_1)\\ b_{12}b_{13}a_{123}\le R_1\\ R_2<b_{12}b_{23}a_{123}\le R\\ R_3<b_{13}b_{23}a_{123}\le R}} g_R(d_1,d_2,d_3)\\ &= \sum\nolimits'_{\substack{\mathcal{D}(a_2D_3)\\ b_{12}b_{13}a_{123}\le R_1\\ R_2<b_{12}b_{23}a_{123}\le R\\ b_{13}b_{23}a_{123}\le R_3}} h_R(d_1,d_2,d_3) + O(e^{-c_1\sqrt{\log(R/R_3)}}\log^6 R)\\ &+ \sum\nolimits'_{\substack{\mathcal{D}(D_1)\\ b_{12}b_{13}a_{123}\le R_1\\ R_2<b_{12}b_{23}a_{123}\le R\\ R_3<b_{13}b_{23}a_{123}\le R}} g_R(d_1,d_2,d_3). \qquad (8.3)\end{aligned}$$

The first sum above is the same as $\mathcal{E}_h(\mathcal{D}(D_2))$ with the appropriate relabeling of variables, and the estimate we now obtain applies to both expressions. We see in this sum that $a_2 \le \frac{R}{b_{12}b_{23}a_{123}} \le \frac{R}{R_2}$, and hence the sum is

$$\begin{aligned}&\ll \log(R/R_2) \sum_{b_{12}b_{13}a_{123}|k} \sum_{\frac{R_2}{b_{12}a_{123}}<b_{23}\le\frac{R}{b_{12}a_{123}}} \frac{\mu^2(D_3)}{\phi(D_3)} \sum_{\substack{a_2\le\frac{R}{R_2}\\ (a_2,D_3)=1}} \frac{\mu^2(a_2)}{\phi(a_2)}\mathfrak{S}_2(a_2D_3)\\ &\ll \log^2(R/R_2)\log\log R \sum_{b_{12}b_{13}a_{123}|k} \sum_{\frac{R_2}{b_{12}a_{123}}<b_{23}\le\frac{R}{b_{12}a_{123}}} \frac{\mu^2(D_3)}{\phi(D_3)}\\ &\ll \log^3(R/R_2)\log\log R \left(\sum_{b|k}\frac{\mu^2(b)}{\phi(b)}\right)^3\\ &\ll \log^3(R/R_2)\log\log R(\frac{k}{\phi(k)})^3,\end{aligned}$$

by Lemma 2.4, and hence, since $\frac{k}{\phi(k)} \ll \log\log 3k$,

$$\mathcal{E}_h(\mathcal{D}(D_2)) \ll \log^3(R/R_2)(\log\log R)^4. \qquad (8.4)$$

Similarly, for the second sum in (8.3) both $a_2 \le \frac{R}{R_2}$ and $a_3 \le \frac{R}{R_2}$, so that the sum is

$$\ll \log^5(R/R_2)(\log\log 3k)^3$$

and hence

$$\mathcal{E}_g(\mathcal{D}(D_1)) \ll \log^5(R/R_2)(\log\log R)^3 + e^{-c_1\sqrt{\log(R/R_2)}}\log^6 R. \qquad (8.5)$$

We take

$$R_2 = R_3 = Re^{-c_2(\log\log R)^2} \qquad (8.6)$$

where c_2 is a sufficiently large constant, and conclude by (8.2), (8.4), and (8.5) that, for $|k| \le R^{\frac{1}{2}-\epsilon}$,

$$T_3(k) = W_3(k) + O\big((\log\log R)^{13}\big), \tag{8.7}$$

where

$$W_3(k) = -\sideset{}{'}\sum_{\substack{b_{12}b_{23}a_{123}\le R_2\\ b_{13}b_{23}a_{123}\le R_2\\ b_{12}b_{13}a_{123}|k}} \frac{\mu^2(D_3)\mu(a_{123})\mu((D_3,2))}{\phi(D_3)}\mathfrak{S}_2(2D_3)\mathfrak{S}_3(2D_3), \tag{8.8}$$

and as before $D_3 = b_{12}b_{13}b_{23}a_{123}$. Since b_{23} is the only variable not constrained by the divisibility condition, we have on letting

$$R_4 = \frac{R_2}{a_{123}}\min(\frac{1}{b_{12}}, \frac{1}{b_{13}}), \tag{8.9}$$

and $E = b_{12}b_{13}a_{123}$, that

$$W_3(k) = -\sideset{}{'}\sum_{E|k} \frac{\mu^2(E)}{\phi(E)}\mu(a_{123}) \sum_{\substack{b_{23}\le R_4\\ (b_{23},E)=1}} \frac{\mu^2(b_{23})}{\phi(b_{23})}\mu((b_{23}E,2))\mathfrak{S}_2(2b_{23}E)\mathfrak{S}_3(2b_{23}E).$$

We break the inner sum into two subsums according to whether b_{23} is even or odd, the former case forcing $(E,2)=1$. We thus obtain, on taking $b_{23}=2b$ in the first subsum and applying (2.4)

$$\begin{aligned} W_3(k) &= \sideset{}{'}\sum_{\substack{E|k\\ (E,2)=1}} \frac{\mu^2(E)\mathfrak{S}_2(2E)}{\phi(E)}\mu(a_{123}) \sum_{\substack{b\le\frac{R_4}{2}\\ (b,2E)=1}} \frac{\mu^2(b)}{\phi_2(b)}\mathfrak{S}_3(2bE) \\ &\quad - \sideset{}{'}\sum_{E|k} \frac{\mu^2(E)\mathfrak{S}_2(2E)}{\phi(E)}\mu((E,2))\mu(a_{123}) \sum_{\substack{b_{23}\le R_4\\ (b_{23},2E)=1}} \frac{\mu^2(b_{23})}{\phi_2(b_{23})}\mathfrak{S}_3(2b_{23}E). \end{aligned}$$

The inner sums in both sums above are by Lemma 2.4 and the estimate (2.19)

$$= \log R_4 + O(\log\log 3k) + O(\frac{\exp\left(\frac{c\sqrt{\log k}}{\log\log 3k}\right)}{\sqrt{R_4}}).$$

Substituting we now see that the main term from the first sum is canceled by the part of the main term in the second sum when $(E,2)=1$, and hence we obtain

$$\begin{aligned}
&W_3(k)\\
&= \sum_{E|k}{}' \frac{\mu^2(E)\mathfrak{S}_2(2E)}{\phi(E)}\Big(\mu(a_{123})[2|E]\log R_4 + +O(\log\log 3k) + O(\frac{\exp\left(\frac{c\sqrt{\log k}}{\log\log 3k}\right)}{\sqrt{R_4}})\Big)\\
&= \log R_2 \sum_{\substack{E|k\\ 2|E}}{}' \frac{\mu^2(E)\mathfrak{S}_2(2E)}{\phi(E)}\mu(a_{123})\\
&\quad + O\left(\sum_{E|k}{}' \frac{\mu^2(E)\mathfrak{S}_2(2E)}{\phi(E)}\Big(\log E + \log\log 3k + \frac{\exp\left(\frac{c\sqrt{\log k}}{\log\log 3k}\right)}{\sqrt{\frac{R_2}{E}}}\right)\\
&= Y_3(k)\log R_2 + O\left((\sum_{b|k}\frac{\mu^2(b)\log b}{\phi(b)})^3 \log\log 3k\right)\\
&\quad + O\left(\frac{\exp\left(\frac{c\sqrt{\log k}}{\log\log 3k}\right)}{\sqrt{R_2}}(\sum_{b|k}\frac{\mu^2(b)\sqrt{b}}{\phi(b)})^3\right),
\end{aligned}$$

where we used the estimate $\mathfrak{S}_2(k) \ll \log\log 3k$ in the last two error terms. By Lemma 5 of [11], the sum in the first error term is

$$= \frac{kh_1(k)}{\phi(k)} \ll (\log\log 3k)^2$$

where $h_1(k)$ is given by (2.17). The sum in the second error term is

$$\ll \exp\left(c\sum_{n<\log k}\frac{\mu^2(n)\sqrt{n}}{\phi(n)}\right) \ll e^{c\sqrt{\log k}}$$

and hence we conclude that

$$W_3(k) = Y_3(k)\log R_2 + O((\log\log R)^7). \tag{8.10}$$

We complete the evaluation of $W_3(k)$ by proving

$$Y_3(k) = \mathfrak{S}_2(k). \tag{8.11}$$

We note first that if $2 \not| k$ then the sum defining $Y_3(k)$ is empty and therefore $Y_3(k) = 0$, in agreement with $\mathfrak{S}_2(k)$. Therefore by (7.17)

$$\begin{aligned} Y_3(k) &= [2|k] \sum_{\substack{E|k \\ 2|E}} \frac{\mu^2(E)\mathfrak{S}_2(2E)}{\phi(E)} \\ &= 2C_2[2|k] \sum_{\substack{2E'|k \\ (E',2)=1}} \frac{\mu^2(E')}{\phi_2(E')} \\ &= 2C_2[2|k]H_2(k) = \mathfrak{S}_2(k), \end{aligned}$$

which proves (8.11).

9. Evaluation of $S_3(0)$: Main Term

We now consider $k_1 = k_2 = k_3 = 0$. On applying equations (6.12), (6.13), and (6.14) and taking

$$S = R_1 = R_2 = R_3$$

we have that

$$T_3(0) = W_3(0) + O(e^{-c_1\sqrt{\log(R/S)}} \log^8 R) + \mathcal{E}_f(\mathcal{D}) + \mathcal{E}_g(\mathcal{D}(D_1)) + \mathcal{E}_h(\mathcal{D}(D_2)), \tag{9.1}$$

where we relabel variables for simplicity and have

$$W_3(0) = -\sideset{}{'}\sum_{\substack{uvy\le S \\ uwy\le S \\ vwy\le S}} \frac{\mu(y)\mu^2(D)\mu((D,2))}{\phi(D)}\mathfrak{S}_2(2D)\mathfrak{S}_3(2D), \tag{9.2}$$

and $D = uvwy$. We will prove in this section that

$$W_3(0) = \frac{3}{4}\log^2 S + O(\log S). \tag{9.3}$$

In the next section we will prove that $\mathcal{E}_f(\mathcal{D})$, $\mathcal{E}_g(\mathcal{D}(D_1))$, and $\mathcal{E}_h(\mathcal{D}(D_2))$ all satisfy the bound

$$\ll \log^9(\frac{R}{S}) \log R \tag{9.4}$$

from which we conclude from (9.1) and (9.3) on taking

$$S = Re^{-c_2(\log\log R)^2}$$

that

$$T_3(0) = \frac{3}{4}\log^2 R + O(\log R(\log\log R)^{18}) \tag{9.5}$$

and therefore by (6.7) we obtain (6.3).

The first step in evaluating $W_3(0)$ is to sum over u in (9.2) by separating into sums according to whether u is even or odd, and apply Lemma 2.4. We let $z = vwy$, and have

$$\begin{aligned}
&\sum_{\substack{u\le \min(\frac{S}{vy},\frac{S}{wy})\\(u,z)=1}} \frac{\mu^2(u)}{\phi(u)}\mu((uz,2))\mathfrak{S}_2(2uz)\mathfrak{S}_3(2uz)\\
&\qquad = 2C_2H_2(2z)\sum_{\substack{u\le \frac{S}{z}\min(v,w)\\(u,z)=1}} \frac{\mu^2(u)}{\phi(u)}\mu((uz,2))H_2(2u)\mathfrak{S}_3(2uz)\\
&\qquad = 2C_2H_2(2z)\Bigg(-[(z,2)=1]\sum_{\substack{2u'\le \frac{S}{z}\min(v,w)\\(u',2z)=1}} \frac{\mu^2(u')}{\phi_2(u')}\mathfrak{S}_3(2u'z)\\
&\qquad\quad + \mu((z,2))\sum_{\substack{u\le \frac{S}{z}\min(v,w)\\(u,2z)=1}} \frac{\mu^2(u)}{\phi_2(u)}\mathfrak{S}_3(2uz).\Bigg)\\
&\qquad = 2C_2H_2(2z)\Bigg(-[(z,2)=2]\Bigg(\log\Big(\frac{S(z,3)}{3z}\min(v,w)\Big)+D_3\\
&\qquad\quad + h_3(6z)\Bigg)+[(z,2)=1]\log 2 + O\Big(\frac{m(6z)}{\sqrt{\frac{S}{z}\min(v,w)}}\Big)\Bigg).
\end{aligned}$$

Substituting this result into (9.2) we obtain

$$\begin{aligned}
W_3(0) =& 2C_2{\sum_{\substack{z\le S\\2|z}}}' \frac{\mu^2(z)\mu(y)}{\phi(z)}H_2(2z)\Big(\log\Big(\frac{S(z,3)}{3z}\min(v,w)\Big)+D_3+h_3(6z)\Big)\\
&- 2C_2\log 2{\sum_{\substack{z\le S\\(z,2)=1}}}' \frac{\mu^2(z)\mu(y)}{\phi(z)}H_2(2z) + O\Big({\sum_{z\le S}}' \frac{\mu^2(z)m(6z)}{\phi(z)\sqrt{\frac{S}{z}\min(v,w)}}H_2(2z)\Big).
\end{aligned} \tag{9.6}$$

Except for the factor of $\log\min(v,w)$, the sums above are of the form considered in (7.17).

Hence we have

$$\begin{aligned} 2C_2 \sideset{}{'}\sum_{\substack{z\le S\\ 2|z}} \frac{\mu^2(z)\mu(y)}{\phi(z)} H_2(2z)\log\frac{S}{z} &= 2C_2 \sum_{\substack{z\le S\\ 2|z}} \frac{\mu^2(z)}{\phi(z)} H_2(2z)\log\frac{S}{z} \\ &= 2C_2 \sum_{\substack{2z'\le S\\ (z',2)=1}} \frac{\mu^2(z')}{\phi_2(z')}\log\frac{S}{2z'} \\ &= \frac{1}{2}\log^2 S + O(\log S), \end{aligned} \tag{9.7}$$

where we used (2.21). Similarly we have, by Lemma 2.2

$$2C_2 \sideset{}{'}\sum_{\substack{z\le S\\ 2|z}} \frac{\mu^2(z)\mu(y)}{\phi(z)} H_2(2z) = \log S + O(1). \tag{9.8}$$

We also see that the condition $2|z$ in (9.8) may be replaced by $(z,2)=1$ and (9.8) will still hold. Similarly it is clear that

$$2C_2 \sideset{}{'}\sum_{\substack{z\le S\\ 2|z}} \frac{\mu^2(z)\mu(y)}{\phi(z)} H_2(2z)\log((z,3)) \ll \log S, \tag{9.9}$$

and also that

$$\begin{aligned} &2C_2 \sideset{}{'}\sum_{\substack{z\le S\\ 2|z}} \frac{\mu^2(z)\mu(y)}{\phi(z)} H_2(2z)h_3(6z) \\ &\quad\ll \sum_{\substack{z\le S\\ 2|z}} \frac{\mu^2(z)}{\phi(z)} H_2(2z)h_3(6z) \\ &\quad\ll \sum_{p\le S}\frac{\log p}{p}\sum_{\substack{z\le S\\ p|z}} \frac{\mu^2(z)}{\phi(z)} H_2(2z) \\ &\quad\ll \sum_{3\le p\le S}\frac{\log p}{p\phi_2(p)}\sum_{\substack{m\le \frac{S}{p}\\ (m,p)=1}} \frac{\mu^2(m)}{\phi(m)} H_2(2m) \\ &\quad\ll \log S. \end{aligned} \tag{9.10}$$

We conclude that

$$\begin{aligned} W_3(0) =&\frac{1}{2}\log^2 S + O(\log S) + 2C_2 \sideset{}{'}\sum_{\substack{z\le S\\ 2|z}} \frac{\mu^2(z)\mu(y)}{\phi(z)} H_2(2z)\log(\min(v,w)) \\ &+ O\Big(\sideset{}{'}\sum_{z\le S} \frac{\mu^2(z)m(6z)}{\phi(z)\sqrt{\frac{S}{z}\min(v,w)}} H_2(2z)\Big). \end{aligned} \tag{9.11}$$

We now evaluate the first sum in (9.11), and show that

$$2C_2 \sideset{}{'}\sum_{\substack{z\le S\\ 2|z}} \frac{\mu^2(z)\mu(y)}{\phi(z)} H_2(2z)\log(\min(v,w)) = \frac{1}{4}\log^2 S + O(\log S). \tag{9.12}$$

By the symmetry in the variables v and w we have that the sum above is

$$= 4C_2 \sideset{}{'}\sum_{\substack{vwy\le S\\ 2|vwy\\ v<w}} \frac{\mu^2(v)\mu^2(w)\mu(y)}{\phi(vwy)} H_2(2vwy)\log v, \tag{9.13}$$

since the condition $(v,w)=1$ implies that the only term with $v=w$ is when $v=w=1$, and this term is zero. We will sum over w and apply Lemma 2.2 according to the parity of w, which gives

$$\begin{aligned}
\sum_{\substack{v<w\le \frac{S}{vy}\\ 2|vwy\\ (w,vy)=1}} \frac{\mu^2(w)}{\phi(w)}H_2(w) &= [(vy,2)=1]\sum_{\substack{v<2w'\le \frac{S}{vy}\\ (w',2vy)=1}} \frac{\mu^2(w')}{\phi_2(w')} + [2|vy]\sum_{\substack{v<w\le \frac{S}{vy}\\ (w,2vy)=1}} \frac{\mu^2(w)}{\phi_2(w)}\\
&= \Big([(vy,2)=1]+[2|vy]\Big)\left(\frac{1}{\mathfrak{S}_2(2vy)}\log\frac{S}{v^2y}\right) + O(\frac{m(2vy)}{\sqrt{v}})\\
&= \frac{1}{\mathfrak{S}_2(2vy)}\log\frac{S}{v^2y} + O(\frac{m(2vy)}{\sqrt{v}}).
\end{aligned}$$

Substituting this result into (9.13) we find that this sum is

$$\begin{aligned}
=&4C_2\sum_{v\le\sqrt{S}}\frac{\mu^2(v)}{\phi(v)}H_2(v)\log v\sum_{\substack{y\le\frac{S}{v^2}\\ (y,v)=1}}\frac{\mu(y)}{\phi(y)}H_2(y)\left(\frac{1}{\mathfrak{S}_2(2vy)}\log\frac{S}{v^2y}+O(\frac{m(2vy)}{\sqrt{v}})\right)\\
=&2\sum_{v\le\sqrt{S}}\frac{\mu^2(v)}{\phi(v)}\log v\sum_{\substack{y\le\frac{S}{v^2}\\ (y,v)=1}}\frac{\mu(y)}{\phi(y)}\log\frac{S}{v^2y}\\
&\quad + O\Big(\sum_{v\le\sqrt{S}}\frac{\mu^2(v)H_2(v)}{\sqrt{v}\phi(v)}m(v)\log v\sum_{y\le\frac{S}{v^2}}\frac{\mu^2(y)H_2(y)}{\phi(y)}m(y)\Big).
\end{aligned}$$

The inner sum in the error term here is

$$\begin{aligned}
&\ll \sum_{y\le S}\frac{\mu^2(y)H_2(y)}{\phi(y)}\sum_{d|y}\frac{\mu^2(d)}{\sqrt{d}}\\
&= \sum_{d\le S}\frac{\mu^2(d)H_2(d)}{\sqrt{d}\phi(d)}\sum_{\substack{m\le S/d\\ (m,d)=1}}\frac{\mu^2(m)H_2(m)}{\phi(m)}\\
&\ll \log S,
\end{aligned}$$

and by (2.19) the sum over v in the error term converges; hence the error term is $O(\log S)$. Thus by Lemma 2.1 and partial summation in Lemma 2.4 the above expression is

$$\begin{aligned} &= 2\sum_{v\leq\sqrt{S}}\frac{\mu^2(v)}{\phi(v)}\mathfrak{S}_2(v)\log v + O(\sum_{v\leq\sqrt{S}}\left(\frac{\mu^2(v)}{\phi(v)}\log v\right)e^{-c_1\sqrt{\log(\frac{S}{v^2})}}) + O(\log S) \\ &= \log^2\sqrt{S} + O(\log S), \end{aligned}$$

which proves (9.12).

It remains to deal with the error term in (9.11). By symmetry we may assume that $v \leq w$, and thus the error term is

$$\begin{aligned} &\ll \frac{1}{\sqrt{S}}\sideset{}{'}\sum_{\substack{vwy\leq S\\(vwy,2)=1}}\frac{\mu^2(vwy)\sqrt{wy}}{\phi_2(vwy)}m(vwy) \\ &\ll \frac{1}{\sqrt{S}}\sum_{w\leq S}\frac{\mu^2(w)\sqrt{w}}{\phi_2(w)}m(w)\sideset{}{'}\sum_{\substack{v\leq\frac{S}{w}\\(v,2)=1}}\frac{\mu^2(v)}{\phi_2(v)}m(v)\sideset{}{'}\sum_{\substack{y\leq\frac{S}{vw}\\(y,2)=1}}\frac{\mu^2(y)\sqrt{y}}{\phi_2(y)}m(y). \end{aligned}$$

The inner sum over y is

$$\begin{aligned} &\ll \sum_{\substack{y\leq\frac{S}{vw}\\(y,2)=1}}\frac{\mu^2(y)\sqrt{y}}{\phi_2(y)}\sum_{d|y}\frac{\mu^2(d)}{\sqrt{d}} \\ &\ll \sum_{\substack{d\leq\frac{S}{vw}\\(d,2)=1}}\frac{\mu^2(d)}{\phi_2(d)}\sum_{\substack{m\leq\frac{S}{dvw}\\(m,2)=1}}\frac{\mu^2(m)\sqrt{m}}{\phi_2(m)} \\ &\ll \sqrt{\frac{S}{vw}}\sum_{\substack{d\leq\frac{S}{vw}\\(d,2)=1}}\frac{\mu^2(d)}{\phi_2(d)\sqrt{d}} \\ &\ll \sqrt{\frac{S}{vw}}. \end{aligned}$$

Substituting we see the sum over v now converges, and on summing over w and treating $m(w)$ as in the previous estimate we see the error is $O(\log S)$. This completes the proof of (9.3).

10. Evaluation of $S_3(0)$: Error Terms

We now treat the error terms $\mathcal{E}_f(\mathcal{D})$, $\mathcal{E}_g(\mathcal{D}(D_1))$, and $\mathcal{E}_h(\mathcal{D}(D_2))$. We proceed as we did before in (8.3); in $\mathcal{E}_f(\mathcal{D})$ we break the sum into two sums according to whether

$b_{12}b_{23}a_{123} \leq R_1$ or $R_1 < b_{12}b_{23}a_{123} \leq R$, in the former sum we sum over a_2 using Lemma 2.1 and obtain a sum of the same form as $\mathcal{E}_g(\mathcal{D}(D_1))$ and an error term

$$\ll e^{-c_1\sqrt{\log(R/R_1)}} \log^8 R. \tag{10.1}$$

In the second sum where $R_1 < b_{12}b_{23}a_{123}$ we have $a_2 \leq \frac{R}{R_1}$ and we do not sum over a_2. We continue this process with regard to a_3, and likewise deal with the error term $\mathcal{E}_g(\mathcal{D}(D_1))$. The result of this process is that we are left with errors bounded by (10.1) and three types of sums of the forms

$$\mathcal{E}_1 = \sum_{a \leq R/R_1} \frac{\mu(a)}{\phi(a)} \sideset{}{'}\sum_{\substack{uvy \leq R_1 \\ uwy \leq R_1 \\ R_1 < vwy \leq R/a}} \frac{\mu(y)\mu^2(D)}{\phi(D)} \mathfrak{S}_2(aD) \log \frac{R}{avwy}, \tag{10.2}$$

$$\mathcal{E}_2 = \sideset{}{'}\sum_{a,b \leq R/R_1} \frac{\mu(ab)}{\phi(ab)} \sideset{}{'}\sum_{\substack{uvy \leq R_1 \\ R_1 < uwy \leq R/a \\ R_1 < vwy \leq R/b}} \frac{\mu(y)\mu^2(D)}{\phi(D)} \log \frac{R}{auwy} \log \frac{R}{bvwy}, \tag{10.3}$$

and

$$\mathcal{E}_3 = \sideset{}{'}\sum_{a,b,c \leq R/R_1} \frac{\mu(abc)}{abc} \sideset{}{'}\sum_{\substack{R_1 < uvy \leq R/a \\ R_1 < uwy \leq R/b \\ R_1 < vwy \leq R/c}} \frac{\mu(y)\mu^2(D)}{D} \log \frac{R}{auvy} \log \frac{R}{buwy} \log \frac{R}{cvwy}, \tag{10.4}$$

where $D = uvwy$.

We can handle $\mathcal{E}_3$ immediately. Estimating trivially, we have

$$\mathcal{E}_3 \ll (\log R/R_1)^6 \sum_{\substack{R_1 < uvy \leq R \\ R_1 < uwy \leq R \\ R_1 < vwy \leq R}} \frac{1}{uvwy}.$$

The top two inequalities in the summation conditions imply

$$\left(\frac{R_1}{u}\right)^2 < vwy^2 \leq \left(\frac{R}{u}\right)^2$$

and hence

$$\left(\frac{R_1}{u}\right)^2 \frac{1}{vwy} < y \leq \left(\frac{R}{u}\right)^2 \frac{1}{vwy}$$

Thus the bottom inequality in the summation conditions implies

$$\left(\frac{R_1}{u}\right)^2 \frac{1}{R} < y \leq \left(\frac{R}{u}\right)^2 \frac{1}{R_1},$$

and therefore the sum above is

$$\begin{aligned} &\ll \sum_{u\le R}\frac{1}{u}\sum_{\left(\frac{R_1}{u}\right)^2\frac{1}{R}<y\le\left(\frac{R}{u}\right)^2\frac{1}{R_1}}\frac{1}{y}\sum_{\frac{R_1}{uy}<w\le\frac{R}{uy}}\frac{1}{w}\sum_{\frac{R_1}{uy}<v\le\frac{R}{uy}}\frac{1}{v} \\ &\ll \log R(\log R/R_1)^3. \end{aligned}$$

Thus

$$\mathcal{E}_3 \ll \log R(\log R/R_1)^9. \tag{10.5}$$

Consider next $\mathcal{E}_2$. The trivial estimate used for $\mathcal{E}_3$ would give the bound $\ll \log^2 R(\log R/R_1)^6$, and therefore we need to save a factor of $\log R$, which will occur when we sum over y. We first note that the conditions on the summation variables for the sum in (10.3) imply that $u, v \le w$. Next, we extend the summation range $uvy \le R_1$ to $uvy \le R$, which may be done with an error $\ll \log R(\log R/R_1)^7$ in the same way that (10.5) was obtained. Finally, the terms with $R_1 < wy$ also contribute this same error, since this condition implies with the other summation conditions that $u, v \le \frac{R}{R_1}$ and $\frac{R_1}{uy} < w < \frac{R}{uy}$, so that only y has a full summation range. Hence we have

$$\mathcal{E}_2 = {\sum_{a,b\le R/R_1}}' \frac{\mu(ab)}{\phi(ab)} {\sum_{\substack{u,v\le w\\ wy\le R_1\\ R_1<uwy\le R/a\\ R_1<vwy\le R/b}}}' \frac{\mu(y)\mu^2(D)}{\phi(D)}\log\frac{R}{auwy}\log\frac{R}{bvwy} + O\big(\log R(\log(R/R_1))^7\big). \tag{10.6}$$

We now sum over u, which satisfies

$$\frac{R_1}{wy} < u \le \min\Big(\frac{R}{awy}, w\Big).$$

If $\min(\frac{R}{awy}, w) = w$ then $w^2 \le \frac{R}{ay}$ and there will only be terms when $\frac{R_1}{y} < w^2$. We conclude in this case that

$$\sqrt{\frac{R_1}{y}} < w \le \sqrt{\frac{R}{ay}}.$$

Hence, as in the estimate to obtain (10.5), these terms contribute at most $\ll \log R(\log(R/R_1))^7$ since only the variable y runs through a full summation range. We conclude, with $z = vwy$,

$$\begin{aligned} \mathcal{E}_2 = {\sum_{a,b\le R/R_1}}' \frac{\mu(ab)}{\phi(ab)} {\sum_{\substack{v\le w\\ wy\le R_1\\ R_1<z\le R/b}}}' \frac{\mu(y)\mu^2(z)}{\phi(z)}\log\frac{R}{bz}\sum_{\substack{\frac{R_1}{wy}<u\le\frac{R}{awy}\\ (u,abz)=1}}\frac{\mu^2(u)}{\phi(u)}\log\frac{R}{auwy} \\ + O\big(\log R(\log(R/R_1))^7\big). \end{aligned} \tag{10.7}$$

To evaluate the inner sum, we use the relation, for $1 \leq S \leq R$, $p(j)|k$,

$$\sum_{\substack{\frac{R}{S}<d\leq R \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} \log \frac{R}{d} = \frac{1}{2\mathfrak{S}_j(k)} \log^2 S + O\Big(\frac{m(k)}{\sqrt{\frac{R}{S}}} \log S\Big). \tag{10.8}$$

This result follows immediately on writing the sum on the left-hand side above as

$$= \sum_{\substack{d\leq R \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} \log \frac{R}{d} - \sum_{\substack{d\leq \frac{R}{S} \\ (d,k)=1}} \frac{\mu^2(d)}{\phi_j(d)} \Big(\log \frac{R}{Sd} + \log S \Big)$$

and applying (2.21) and Lemma 2.2. Thus we have

$$\sum_{\substack{\frac{R_1}{wy}<u\leq \frac{R}{awy} \\ (u,abz)=1}} \frac{\mu^2(u)}{\phi(u)} \log \frac{R}{auwy} = \frac{\phi(abz)}{2abz} \log^2 \frac{R}{aR_1} + O\Big(\frac{m(abz)}{\sqrt{\frac{R_1}{wy}}} \log \frac{R}{aR_1}\Big). \tag{10.9}$$

Substituting this expression into (10.7) we obtain

$$\mathcal{E}_2 = \frac{1}{2} \sideset{}{'}\sum_{a,b\leq R/R_1} \frac{\mu(ab)}{ab} \log^2 \frac{R}{aR_1} \sideset{}{'}\sum_{\substack{v\leq w \\ wy\leq R_1 \\ R_1<z\leq R/b}} \frac{\mu(y)\mu^2(z)}{z} \log \frac{R}{bz} + O\Big(\log R (\log(R/R_1))^7 \Big), \tag{10.10}$$

once we show that the contribution

$$\ll \frac{\log^2 \frac{R}{R_1}}{\sqrt{R_1}} \sideset{}{'}\sum_{a,b\leq R/R_1} \frac{\mu^2(ab)m(ab)}{\phi(ab)} \sideset{}{'}\sum_{\substack{v\leq w \\ wy\leq R_1 \\ R_1<z\leq R/b}} \frac{\mu^2(z)\sqrt{wy}m(z)}{\phi(z)}$$

from the error term in (10.9) is covered by the error term in (10.10). This expression is of the same form as the error term in (9.11) estimated at the end of the last section, except it is over a more restricted summation range. The factors $m(ab)$ and $m(z)$ are handled as in that argument, and make no contribution, so we may ignore them. Hence the expression above is, by Lemma 2.2,

$$\begin{aligned}
&\ll \frac{\log^4 \frac{R}{R_1}}{\sqrt{R_1}} \sum_{\substack{v\leq w \\ wy\leq R_1 \\ R_1<vwy\leq R}} \frac{\mu^2(vwy)\sqrt{wy}}{\phi(vwy)} \\
&\ll \frac{\log^4 \frac{R}{R_1}}{\sqrt{R_1}} \sum_{w\leq R_1} \frac{\sqrt{w}}{\phi(w)} \sum_{y\leq \frac{R_1}{w}} \frac{\sqrt{y}}{\phi(y)} \sum_{\frac{R_1}{wy}<v\leq \frac{R}{wy}} \frac{\mu^2(v)}{\phi(v)} \\
&\ll \log R_1 \log^5 \frac{R}{R_1},
\end{aligned}$$

which is acceptable. Thus we have established (10.10).

We now treat the sum in (10.10) and show it is also bounded by the error term, from which we conclude that

$$\mathcal{E}_2 \ll \log R(\log R/R_1)^7. \tag{10.11}$$

To see this, consider the sum over y in equation (10.10)

$$\sum_{\substack{\frac{R_1}{vw}<y\le\min(\frac{R_1}{w},\frac{R}{bvw})\\ (y,abvw)=1}} \frac{\mu(y)}{y}\log\frac{R}{bvwy}.$$

For the terms with $\min\left(\frac{R_1}{w},\frac{R}{bvw}\right)=\frac{R_1}{w}$, we have $v\le\frac{R}{bR_1}$, and the sum in (10.10) is

$$\begin{aligned}
&\ll \log^5\frac{R}{R_1}\sum_{v\le\frac{R}{R_1}}\frac{\mu^2(v)}{v}\sum_{w\le R_1}\frac{\mu^2(w)}{w}\sum_{\frac{R_1}{vw}<y\le\frac{R_1}{w}}\frac{\mu^2(y)}{y}\\
&\ll \log^5\frac{R}{R_1}\sum_{v\le\frac{R}{R_1}}\frac{\mu^2(v)}{v}\sum_{w\le R_1}\frac{\mu^2(w)}{w}\left([vw\le R_1]\log v+[vw>R_1]\log\frac{R_1}{w}\right)\\
&\ll \log R\log^5\frac{R}{R_1}\left(\sum_{v\le\frac{R}{R_1}}\frac{\mu^2(v)\log v}{v}+\sum_{v\le\frac{R}{R_1}}\frac{\mu^2(v)}{v}\sum_{\frac{R_1}{v}<w\le R_1}\frac{\mu^2(w)}{w}\right)\\
&\ll \log R\log^7\frac{R}{R_1},
\end{aligned}$$

which is acceptable. For the remaining terms when $\min\left(\frac{R_1}{w},\frac{R}{bvw}\right)=\frac{R}{bvw}$, we have $\frac{R}{bR_1}<v$, and by Lemma 2.1 with the error term estimate (2.14) the sum over y is

$$\begin{aligned}
&=\sum_{\substack{y\le\frac{R}{bvw}\\ (y,abvw)=1}}\frac{\mu(y)}{y}\log\frac{R}{bvwy}-[vw\le R_1]\sum_{\substack{y\le\frac{R_1}{vw}\\ (y,abvw)=1}}\frac{\mu(y)}{y}\left(\log\frac{R_1}{vwy}+\log\frac{R}{bR_1}\right)\\
&=(1-[vw\le R_1])\mathfrak{S}_1(abvw)+O(m(abvw)e^{-c_1\sqrt{\log(\frac{R}{bvw})}})\\
&\quad+[vw\le R_1]O(m(abvw)e^{-c_1\sqrt{\log(\frac{R_1}{vw})}}\log\frac{R}{R_1}).
\end{aligned}$$

Hence the sum in (10.10) is in this case

$$\begin{aligned}
&\ll \log^5(R/R_1)\sum_{vw\le R_1}\frac{m(vw)}{vw}e^{-c_1\sqrt{\log\frac{R_1}{vw}}}\\
&\quad+\log^4(R/R_1)\sum_{R_1<vw\le R}\frac{1}{vw}(\mathfrak{S}_1(vw)+m(vw))\\
&\ll \log R(\log(R/R_1))^5,
\end{aligned}$$

where as before the factors of $m(vw)$ and $\mathfrak{S}(vw)$ make no contribution to the error when they are summed. This finishes the proof of (10.11).

Finally consider $\mathcal{E}_1$. The inner sum in (10.2) is, with $E = uwy$,

$$\sum_{\substack{uvy\le R_1\\ uwy\le R_1\\ R_1<vwy\le R/a\\ (D,a)=1}}{}' \frac{\mu(y)\mu^2(D)}{\phi(D)}\mathfrak{S}_2(aD)\log\frac{R}{avwy} = 2C_2\sum_{\substack{E\le R_1\\ (E,a)=1}}{}' \frac{\mu(y)\mu^2(E)}{\phi(E)}H_2(aE)S_R(E),$$

where

$$S_R(E) = \sum_{\substack{\frac{R_1}{wy}<v\le\min(\frac{R}{awy},\frac{R_1}{uy})\\ (v,aE)=1}} \chi(aE,v)\frac{\mu^2(v)H_2(v)}{\phi(v)}\log\frac{R}{avwy}$$

and

$$\chi(aE,v) = [2|v][(aE,2)=1] + [(v,2)=1][2|aE].$$

We apply (10.8) and Lemma 2.2 to evaluate $S_R(E)$. When $\min(\frac{R}{awy},\frac{R_1}{uy}) = \frac{R}{awy}$ then in (10.8) $S = \frac{R}{aR_1}$, and

$$S_R(E) = \frac{1}{2\mathfrak{S}_2(2aE)}\log^2\frac{R}{aR_1} + O(\frac{m(2aE)\log\frac{R}{aR_1}}{\sqrt{\frac{R_1}{wy}}});$$

while if $\min(\frac{R}{awy},\frac{R_1}{uy}) = \frac{R_1}{uy}$ then $1 < S = \frac{w}{u} \le \frac{R}{aR_1}$, whence $u < w \le \frac{R}{aR_1}u$, and thus we obtain

$$S_R(E) = \frac{1}{2\mathfrak{S}_2(2aE)}\log\frac{w}{u}\log\left(\left(\frac{R}{aR_1}\right)^2\frac{u}{w}\right) + O(\frac{m(2aE)\log\frac{R}{aR_1}}{\sqrt{\frac{R_1}{wy}}}).$$

On substituting, the inner sum in (10.2) becomes

$$\begin{aligned}
=&\frac{1}{2}\log^2\frac{R}{aR_1}\sum_{\substack{uw\le R_1\\ \frac{R}{aR_1}\le\frac{w}{u}\\ (uw,a)=1}}{}' \frac{\mu^2(uw)}{\phi(uw)}\sum_{\substack{y\le\frac{R_1}{uw}\\ (y,auw)=1}}\frac{\mu(y)}{\phi(y)}\\
&+\frac{1}{2}\sum_{\substack{uw\le R_1\\ 1<\frac{w}{u}<\frac{R}{aR_1}\\ (uw,a)=1}}{}' \frac{\mu^2(uw)}{\phi(uw)}\log\frac{w}{u}\log\left(\left(\frac{R}{aR_1}\right)^2\frac{u}{w}\right)\sum_{\substack{y\le\frac{R_1}{uw}\\ (y,auw)=1}}\frac{\mu(y)}{\phi(y)}\\
&+O\left(\sum_{\substack{uw\le R_1\\ (uw,a)=1}}{}' \frac{\mu^2(uw)}{\phi(uw)}\Big(\sum_{\substack{y\le\frac{R_1}{uw}\\ (y,auw)=1}}\frac{\mu^2(y)H_2(2auwy)}{\phi(y)}\frac{m(2auwy)\log\frac{R}{aR_1}}{\sqrt{\frac{R_1}{wy}}}\Big)\right).
\end{aligned}$$

As before, in estimating the contribution of the error term above to (10.2) the factor $m(2auwy)$ may be ignored, and therefore this contribution is

$$\begin{aligned}
&\ll \frac{\log^2 \frac{R}{R_1}}{\sqrt{R_1}} \sum_{\substack{uwy\le R_1\\(uwy,2)=1}} \frac{\mu^2(uwy)\sqrt{wy}}{\phi_2(uwy)}\\
&\ll \frac{\log^2 \frac{R}{R_1}}{\sqrt{R_1}} \sum_{\substack{u\le R_1\\(u,2)=1}} \frac{\mu^2(u)}{\phi_2(u)} \sum_{\substack{w\le \frac{R_1}{u}\\(w,2)=1}} \frac{\mu^2(w)\sqrt{w}}{\phi_2(w)} \sum_{\substack{y\le \frac{R_1}{uw}\\(y,2)=1}} \frac{\mu^2(y)\sqrt{y}}{\phi_2(y)}\\
&\ll \log^2 \frac{R}{R_1} \sum_{\substack{u\le R_1\\(u,2)=1}} \frac{\mu^2(u)}{\sqrt{u}\phi_2(u)} \sum_{\substack{w\le \frac{R_1}{u}\\(w,2)=1}} \frac{\mu^2(w)}{\phi_2(w)}\\
&\ll \log R_1 \log^2 \frac{R}{R_1}.
\end{aligned}$$

Next, for the main terms above the sum over y is $\ll m(auw)e^{-c_2\sqrt{\log \frac{R_1}{uw}}}$ by Lemma 2.1 with the error term in (2.14), and hence both sums contribute to (10.2)

$$\begin{aligned}
&\ll \quad \log^3 \frac{R}{R_1} \sum_{u\le R_1} \frac{\mu^2(u)m(u)}{\phi(u)} \sum_{w\le R_1/u} \frac{\mu^2(w)m(w)}{\phi(w)} e^{-c_1\sqrt{\log \frac{R_1}{uw}}}\\
&\ll \quad \log R \log^3 \frac{R}{R_1}.
\end{aligned}$$

We conclude

$$\mathcal{E}_1 \ll \log R(\log R/R_1)^3. \tag{10.12}$$

By (10.1),(10.5),(10.11), and (10.12) we have proved (9.4) and thus completed the proof of (5.3).

11. Mixed Triple Correlations

The case $k = 2$ of Theorem 1.4 has already been handled by (1.5) and (5.7). In this section we evaluate $W_R(\boldsymbol{k})$ from Section 4 in the case $k = 3$. This will not only prove Theorem 1.4 but will also give an alternative and simpler proof of the second two parts of Theorem 6.1.

We let $\boldsymbol{k} = (k_1, k_2, 0)$ and

$$W_R(\boldsymbol{k}) = \sum_{\substack{d_1,d_2\le R\\(d_1,k_1)=1,\,(d_2,k_2)=1\\(d_1,d_2)|k_2-k_1}} \frac{\mu(d_1)\mu(d_2)}{\phi([d_1,d_2])} \log\frac{R}{d_1} \log\frac{R}{d_2}. \tag{11.1}$$

The results we obtain are contained in the following theorem. Recall $H_j(m)$ is defined in (2.3), $h_j(m)$ is defined in (2.17), and $H_j(m), h_j(m) \ll_j \log\log 3m$.

Theorem 11.1 *If* $\boldsymbol{k} = (k, k, 0)$, $k \neq 0$, *and* $\log|k| \ll \log R$, *we have*

$$W_R(\boldsymbol{k}) = \mathfrak{S}(\boldsymbol{k}) \log R + O(H_2(k)h_3(6k)), \tag{11.2}$$

and if $\boldsymbol{k} = (k_1, k_2, 0)$ *and* $k_1 \neq k_2 \neq 0$, *letting* $\Delta = k_1 k_2 (k_2 - k_1)$, *and assuming* $\Delta < R/2$, *we have*

$$W_R(\boldsymbol{k}) = \mathfrak{S}(\boldsymbol{k}) + O(\frac{k_1}{\phi(k_1)} H_2(\Delta) e^{-c_1\sqrt{\log R/2\Delta}}). \tag{11.3}$$

We decompose into relatively prime variables by letting $d_1 = a_1 b_{12}$ and $d_2 = a_2 b_{12}$ where $(d_1, d_2) = b_{12}$ and thus a_1, a_2, and b_{12} are pairwise relatively prime. Then we have

$$W_R(\boldsymbol{k}) = \sideset{}{'}\sum_{\substack{a_1 b_{12} \leq R \\ a_2 b_{12} \leq R \\ (a_1 b_{12}, k_1) = 1,\, (a_2 b_{12}, k_2) = 1 \\ b_{12} | k_2 - k_1}} \frac{\mu(a_1)\mu(a_2)\mu^2(b_{12})}{\phi(a_1 a_2 b_{12})} \log \frac{R}{a_1 b_{12}} \log \frac{R}{a_2 b_{12}}. \tag{11.4}$$

We first sum over a_1 and apply Lemma 2.1 to see that

$$\sum_{\substack{a_1 \leq R/b_{12} \\ (a_1, a_2 b_{12} k_1) = 1}} \frac{\mu(a_1)}{\phi(a_1)} \log \frac{R}{a_1 b_{12}} = \mathfrak{S}_2(a_2 b_{12} k_1) + r_1(\frac{R}{b_{12}}, a_2 b_{12} k_1), \tag{11.5}$$

and hence we have

$$\begin{aligned} W_R(\boldsymbol{k}) &= \sideset{}{'}\sum_{\substack{a_2 b_{12} \leq R \\ (b_{12}, k_1) = 1,\, (a_2 b_{12}, k_2) = 1 \\ b_{12} | k_2 - k_1}} \frac{\mu(a_2)\mu^2(b_{12})\mathfrak{S}_2(a_2 b_{12} k_1)}{\phi(a_2 b_{12})} \log \frac{R}{a_2 b_{12}} + E_1(R) \\ &= V_R(\boldsymbol{k}) + E_1(R), \end{aligned} \tag{11.6}$$

where

$$\begin{aligned} E_1(R) &= \sideset{}{'}\sum_{\substack{a_2 b_{12} \leq R \\ (b_{12}, k_1) = 1,\, (a_2 b_{12}, k_2) = 1 \\ b_{12} | k_2 - k_1}} \frac{\mu(a_2)\mu^2(b_{12}) r_1(\frac{R}{b_{12}}, a_2 b_{12} k_1)}{\phi(a_2 b_{12})} \log \frac{R}{a_2 b_{12}} \\ &\ll \sum_{\substack{b_{12} \leq R \\ b_{12} | k_2 - k_1}} \frac{\mu^2(b_{12})}{\phi(b_{12})} e^{-c_1\sqrt{\log(R/b_{12})}} \sum_{a_2 \leq R/b_{12}} \frac{\mu^2(a_2)}{\phi(a_2)} \log \frac{R}{a_2 b_{12}} \\ &\ll \sum_{\substack{b_{12} \leq R \\ b_{12} | k_2 - k_1}} \frac{\mu^2(b_{12})}{\phi(b_{12})} e^{-c_2\sqrt{\log(R/b_{12})}}. \end{aligned} \tag{11.7}$$

We first consider the case when $k_1 = k_2 = k$. Then from (11.7) we see that

$$E_1(R) \ll 1. \tag{11.8}$$

By Lemma 2.3 we have

$$\begin{aligned}
V_R(\boldsymbol{k}) &= \sideset{}{'}\sum_{\substack{a_2b_{12}\le R\\(a_2b_{12},k)=1}} \frac{\mu(a_2)\mu^2(b_{12})\mathfrak{S}_2(a_2b_{12}k)}{\phi(a_2b_{12})}\log\frac{R}{a_2b_{12}}\\
&= \sum_{\substack{b_{12}\le R\\(b_{12},k)=1}} \frac{\mu^2(b_{12})}{\phi(b_{12})}\sum_{\substack{a_2\le R/b_{12}\\(a_2,b_{12}k)=1}} \frac{\mu(a_2)\mathfrak{S}_2(a_2b_{12}k)}{\phi(a_2)}\log\frac{R}{a_2b_{12}}\\
&= -\sum_{\substack{b_{12}\le R\\(b_{12},k)=1}} \frac{\mu^2(b_{12})\mu((b_{12}k,2))}{\phi(b_{12})}\mathfrak{S}_2(2b_{12}k)\mathfrak{S}_3(2b_{12}k)\\
&\qquad + O\left(\sum_{b_{12}\le R}\frac{\mu^2(b_{12})H_2(b_{12})}{\phi(b_{12})\log^A(2R/b_{12})}\right)\\
&= Y_R(k)+O(1).
\end{aligned}\tag{11.9}$$

To evaluate $Y_R(k)$ we divide into even terms and odd terms and apply Lemma 2.4. Thus

$$\begin{aligned}
V_R(k) &= [(k,2)=1]\sum_{\substack{2b\le R\\(b,2k)=1}}\frac{\mu^2(b)}{\phi(b)}\mathfrak{S}_2(2bk)\mathfrak{S}_3(2bk)\\
&\qquad -\mu((k,2))\sum_{\substack{b_{12}\le R\\(b_{12},2k)=1}}\frac{\mu^2(b_{12})}{\phi(b_{12})}\mathfrak{S}_2(2b_{12}k)\mathfrak{S}_3(2b_{12}k)\\
&= 2C_2H_2(k)\Bigg([(k,2)=1]\sum_{\substack{b\le R/2\\(b,2k)=1}}\frac{\mu^2(b)}{\phi_2(b)}\mathfrak{S}_3(2bk)\\
&\qquad -\mu((k,2))\sum_{\substack{b_{12}\le R\\(b_{12},2k)=1}}\frac{\mu^2(b_{12})}{\phi_2(b_{12})}\mathfrak{S}_3(2b_{12}k)\Bigg)\\
&= 2C_2H_2(k)\left(\Big([(k,2)=1]-\mu((k,2))\Big)\log R+O(h_3(6k))+O\big(\frac{\mathfrak{S}_3(6k)m(k)}{\sqrt{R}}))\right)\\
&= 2C_2[2|k]H_2(k)\log R+O(H_2(k)h_3(6k))\\
&= \mathfrak{S}_2(k)\log R+O(H_2(k)h_3(6k)).
\end{aligned}$$

On combining these results we have proved the first part of Theorem 11.1.

We now turn to the case that $k_2\ne k_1$. As in Lemma 2.5 we let $\kappa=(k_1,k_2)$ and $\Delta=k_1k_2(k_2-k_1)$. As before let $k^*=\max(|k_1|,|k_2|)$. In this case we see from (11.7) that

$$E_1(R)\ll H_1(k_2-k_1)e^{-c_2\sqrt{\log(R/2k^*)}}.\tag{11.10}$$

We now let $k_1=s_1K_1$, $k_2=s_2K_2$, where K_1 and K_2 are the largest squarefree divisors of k_1 and k_2 respectively. Let $K_{12} = (K_1, K_2)$. Then we have

$$V_R(\boldsymbol{k}) = \sideset{}{'}\sum_{\substack{a_2 b_{12} \le R \\ (b_{12}, K_1)=1,\, (a_2 b_{12}, K_2)=1 \\ b_{12} | k_2 - k_1}} \frac{\mu(a_2)\mu^2(b_{12})\mathfrak{S}_2(a_2 b_{12} K_1)}{\phi(a_2 b_{12})} \log \frac{R}{a_2 b_{12}}$$

Next let $a_2 = c_2 d_2$ where $c_2 | K_1$ and $(d_2, K_1) = 1$ so that $(a_2, K_1) = c_2$, from which we see

$$V_R(\boldsymbol{k}) = \sideset{}{'}\sum_{\substack{c_2 d_2 b_{12} \le R \\ (d_2 b_{12}, K_1)=1,\, (c_2 d_2 b_{12}, K_2)=1 \\ b_{12} | k_2 - k_1,\, c_2 | K_1}} \frac{\mu(c_2)\mu(d_2)\mu^2(b_{12})\mathfrak{S}_2(d_2 b_{12} K_1)}{\phi(c_2 d_2 b_{12})} \log \frac{R}{c_2 d_2 b_{12}}$$

On summing over d_2 by dividing the sum according to whether d_2 is even or odd, we see on applying Lemma 2.1 that

$$\begin{aligned}
&\sum_{\substack{d_2 \le R/c_2 b_{12} \\ (d_2, b_{12} K_1 K_2)=1}} \frac{\mu(d_2)\mathfrak{S}_2(d_2 b_{12} K_1)}{\phi(d_2)} \log \frac{R}{c_2 d_2 b_{12}} \\
&\qquad = 2C_2 H_2(b_{12} K_1) \sum_{\substack{d_2 \le R/c_2 b_{12} \\ (d_2, b_{12} K_1 K_2)=1 \\ 2 | d_2 b_{12} K_1}} \frac{\mu(d_2)}{\phi(d_2)} H_2(d_2) \log \frac{R}{c_2 d_2 b_{12}} \\
&\qquad = 2C_2 H_2(b_{12} K_1) \Bigg(- [(K_1 K_2 b_{12}, 2) = 1] \sum_{\substack{2d \le R/c_2 b_{12} \\ (d, 2 b_{12} K_1 K_2)=1}} \frac{\mu(d)}{\phi_2(d)} \log \frac{R}{2 c_2 d b_{12}} \\
&\qquad\qquad + [2 | b_{12} K_1] \sum_{\substack{d_2 \le R/c_2 b_{12} \\ (d_2, 2 b_{12} K_1 K_2)=1}} \frac{\mu(d_2)}{\phi_2(d_2)} \log \frac{R}{c_2 d_2 b_{12}} \Bigg) \\
&\qquad = 2C_2 H_2(b_{12} K_1) \chi(b_{12}) \mathfrak{S}_3(2 b_{12} K_1 K_2) + O(H_2(b_{12} K_1) e^{-c_1 \sqrt{\log(R/2 c_2 b_{12})}}),
\end{aligned}$$

where

$$\chi(b_{12}) = -[(K_1 K_2 b_{12}, 2) = 1] + [2 | b_{12} K_1].$$

Since $c_2 b_{12} | K_1(k_2 - k_1)$, we see that the condition $|\Delta| < R/2$ implies $c_2 b_{12} \le R/2$ is automatically satisfied. Therefore we have

$$\begin{aligned}
V_R(\boldsymbol{k}) &= 2C_2 \sum_{\substack{b_{12} | k_2 - k_1,\, c_2 | K_1 \\ (b_{12}, K_1 K_2)=1,\, (c_2, K_2)=1}} \frac{\mu(c_2)\mu^2(b_{12}) H_2(b_{12} K_1) \chi(b_{12}) \mathfrak{S}_3(2 b_{12} K_1 K_2)}{\phi(c_2)\phi(b_{12})} + E_2(R) \\
&= Y_R(\boldsymbol{k}) + E_2(R),
\end{aligned} \tag{11.11}$$

where

$$E_2(R) \ll \frac{k_1}{\phi(k_1)} H_2(\Delta) e^{-c_1\sqrt{\log R/2\Delta}}.$$

Since

$$\sum_{\substack{c_2|K_1\\(c_2,K_2)=1}} \frac{\mu(c_2)}{\phi(c_2)} = \prod_{p|\frac{K_1}{K_{12}}}\left(1-\frac{1}{p-1}\right) = [(K_1/K_{12},2)=1]\frac{H_2(K_{12})}{H_2(K_1)},$$

we have

$$Y_R(\boldsymbol{k}) = 2C_2[(K_1/K_{12},2)=1]\frac{H_2(K_{12})}{H_2(K_1)} \sum_{\substack{b_{12}|k_2-k_1\\(b_{12},K_1K_2)=1}} \frac{\mu^2(b_{12})H_2(b_{12}K_1)\chi(b_{12})\mathfrak{S}_3(2b_{12}K_1K_2)}{\phi(b_{12})}. \tag{11.12}$$

We divide the sum in the equation above according to the parity of b_{12} and apply Lemma 2.4 to see that the sum is

$$\begin{aligned}
&= [(K_1K_2,2)=1] \sum_{\substack{2b|k_2-k_1\\(b,2K_1K_2)=1}} \frac{\mu^2(b)H_2(bK_1)\chi(2b)\mathfrak{S}_3(2bK_1K_2)}{\phi(b)}\\
&\qquad + \sum_{\substack{b_{12}|k_2-k_1\\(b_{12},2K_1K_2)=1}} \frac{\mu^2(b_{12})H_2(b_{12}K_1)\chi(b_{12})\mathfrak{S}_3(2b_{12}K_1K_2)}{\phi(b_{12})}\\
&= \eta(K_1,K_2)H_2(K_1) \sum_{\substack{b|k_2-k_1\\(b,2K_1K_2)=1}} \frac{\mu^2(b)\mathfrak{S}_3(2bK_1K_2)}{\phi_2(b)}\\
&= \eta(K_1,K_2)H_2(K_1)\mathfrak{S}_3(2K_1K_2(k_2-k_1)),
\end{aligned}$$

where

$$\eta(K_1,K_2) = [(K_1K_2,2)=1][2|k_2-k_1] - [(K_1K_2,2)=1] + [2|K_1]$$

On substituting we have that

$$Y_R(\boldsymbol{k}) = 2C_2[(K_1/K_{12},2)=1]\eta(K_1,K_2)H_2(K_{12})\mathfrak{S}_3(2K_1K_2(k_2-k_1)).$$

Now $[(K_1/K_{12},2)=1]=0$ if K_1 is even and K_2 is odd, and $\eta(K_1,K_2)=0$ if both K_1 and K_2 are odd or K_1 is odd and K_2 is even. Hence $Y_R(\boldsymbol{k})$ is zero unless K_{12} is even, and therefore

$$Y_R(\boldsymbol{k}) = [2|K_{12}]2C_2H_2(K_{12})\mathfrak{S}_3\big(2K_1K_2(k_2-k_1)\big) = \mathfrak{S}_2(\kappa)\mathfrak{S}_3(\Delta) = \mathfrak{S}(\boldsymbol{k}),$$

by Lemma 2.5. This completes the proof of Theorem 11.1

12. Application to primes

The use of correlations of short divisor sums to study primes goes back at least to Selberg's work on the sieve. Our mixed correlation result that, for $R \le N^{\frac{1}{4}-\epsilon}$,

$$\tilde{\mathcal{S}}_3(N,k) = \sum_{n=1}^{N} {\Lambda_R}^2(n)\Lambda(n+k) = \mathfrak{S}_2(k) N \log R + o(N\log N) \tag{12.1}$$

provides the upper bound for prime pairs in (1.39) with $\mathcal{B} = 4$, since for $n \ge R$

$$\mu^2(n)\Lambda(n) \le \frac{\log n}{\log^2 R}\Lambda_R^2(n),$$

and the prime powers make a contribution $\ll N^{1/2+\epsilon}$. The Selberg sieve provides the same information, and $\lambda_R(n)$ in (1.2) gives the optimal majorant in the Selberg sieve for this problem when the appropriate normalization is used.

To study primes in short intervals, we consider the modified moments

$$M_k'(N,h,\psi_R,C) = \sum_{n=N+1}^{2N} \big(\psi_R(n+h) - \psi_R(n) - C\log N\big)^k, \tag{12.2}$$

and

$$\tilde{M}_k'(N,h,\psi_R,C) = \sum_{n=N+1}^{2N} \big(\psi(n+h) - \psi(n)\big)\big(\psi_R(n+h) - \psi_R(n) - C\log N\big)^{k-1}, \tag{12.3}$$

where C is a function of h and R that will be chosen to optimize our applications. If we take $C = 0$ these moments reduce to the moments considered in Section 1. We will assume in this section that $h = \lambda \log N$, $\lambda \ll 1$, and thus

$$h \ll \log N, \tag{12.4}$$

which we will make free use of in our estimates. We now consider, for $\rho \ge 0$,

$$\begin{aligned} \mathcal{M}(h,\rho) &= \tilde{M}_3'(N,h,\psi_R,C) - \rho\log N M_2'(N,h,\psi_R,C) \\ &= \sum_{n=N+1}^{2N} \big(\psi(n+h) - \psi(n) - \rho\log N\big)\big(\psi_R(n+h) - \psi_R(n) - C\log N\big)^2. \end{aligned} \tag{12.5}$$

To evaluate $\mathcal{M}(h,\rho)$ we see first that

$$\tilde{M}_3'(N,h,\psi_R,C) = \tilde{M}_3'(N,h,\psi_R) - 2C\log N\tilde{M}_2'(N,h,\psi_R) + C^2\log^2 N\tilde{M}_1'(N,h,\psi_R).$$

We apply Corollary 1.5 (which as mentioned in Section 1 applies immediately to $\tilde{M}_k'$ as well as $\tilde{M}_k$), with $R = N^\theta$ and $0 < \theta < \frac{\vartheta}{2}$,

$$\tilde{M}_3'(N,h,\psi_R,C) \sim \Big((\theta^2\lambda + 3\theta\lambda^2 + \lambda^3) - 2C(\theta\lambda + \lambda^2) + C^2\lambda\Big)N\log^3 N$$

and by Corollary 1.2, for $0 < \theta \le \frac{1}{2}$,

$$M_2'(N,h,\psi_R,C) \sim \Big(\theta\lambda + \lambda^2 - 2C\lambda + C^2\Big) N \log^2 N.$$

We therefore see that $\mathcal{M}(h,\rho)$ is quadratic in C when $\lambda \ne \rho$, and therefore on completing the square we find that, for $\lambda \ne \rho$, and $0 < \theta < \frac{\vartheta}{2}$,

$$\mathcal{M}(h,\rho) \sim \Big((\lambda-\rho)\Big(C - \frac{\lambda(\lambda-\rho+\theta)}{\lambda-\rho}\Big)^2 + \frac{\lambda\theta}{\lambda-\rho}\big((\lambda-\rho)^2 - \theta\rho\big)\Big) N \log^3 N.$$

By choosing

$$C = \frac{\lambda(\lambda-\rho+\theta)}{\lambda-\rho} = \lambda\Big(1 + \frac{\theta}{\lambda-\rho}\Big) \tag{12.6}$$

we maximize $\mathcal{M}(h,\rho)$ if $\lambda < \rho$ and minimize it if $\lambda > \rho$. We conclude that with this choice of C, and $0 < \theta < \frac{\vartheta}{2}$,

$$\mathcal{M}(h,\rho) \sim \frac{\lambda\theta}{\lambda-\rho}\big((\lambda-\rho)^2 - \theta\rho\big) N \log^3 N. \tag{12.7}$$

We see that $\mathcal{M}(h,\rho)$ is positive (and $\gg N\log^3 N$) when λ is a fixed number in the range $\rho - \sqrt{\theta\rho} < \lambda < \rho$ but is negative when $\rho < \lambda < \rho + \sqrt{\theta\rho}$.

We now let $P_r(N,h)$ denote the number of integers $N < n \le 2N$ for which the interval $(n, n+h]$ contains exactly r primes. Thus

$$P_r(N,h) = \sum_{\substack{n=N+1 \\ \pi(n+h)-\pi(n)=r}}^{2N} 1. \tag{12.8}$$

The Poisson model for primes in short intervals (see [6]) is equivalent to the conjecture that

$$P_r(N,h) \sim \frac{\lambda^r e^{-\lambda}}{r!} N. \tag{12.9}$$

We let

$$Q_r^-(N,h) = \sum_{m=0}^{r} P_m(N,h) = \sum_{\substack{n=N+1 \\ \pi(n+h)-\pi(n)\le r}}^{2N} 1 \tag{12.10}$$

and

$$Q_r^+(N,h) = \sum_{m=r+1}^{\infty} P_m(N,h) = \sum_{\substack{n=N+1 \\ \pi(n+h)-\pi(n)> r}}^{2N} 1. \tag{12.11}$$

Thus we have

$$Q_r^-(N,h) + Q_r^+(N,h) = N. \tag{12.12}$$

We let $p_{j_0} = p_{j_0}(N)$ and $p_{j_1} = p_{j_1}(N)$ denote respectively the smallest and the largest primes in the interval $[N+1, 2N]$. For smaller than average gaps between primes, we use the relation, for $r \geq 1$,

$$Q_r^+(N,h) = \sum_{\substack{N+1\leq n<p_{j_0}\\ \pi(n+h)-\pi(n)>r}} 1 + \sum_{j=j_0}^{j_1} \sum_{\substack{p_j\leq n<p_{j+1}\\ p_{j+r+1}\leq n+h}} 1 - \sum_{\substack{2N<n<p_{j_1+1}\\ \pi(n+h)-\pi(n)>r}} 1.$$

The first and third sums are $O(Ne^{-c_1\sqrt{\log N}})$ by the prime number theorem with error term, and hence

$$\begin{aligned} Q_r^+(N,h) &= \sum_{N+1\leq p_j\leq 2N} \sum_{\substack{p_j\leq n<p_{j+1}\\ p_{j+r+1}\leq n+h}} 1 + O(Ne^{-c_1\sqrt{\log N}}) \\ &= \sum_{\substack{N+1\leq p_j\leq 2N\\ p_{j+r+1}-p_{j+1}<h}} \big(p_{j+1} - \max(p_j, p_{j+r+1} - \lfloor h\rfloor)\big) + O(Ne^{-c_1\sqrt{\log N}}) \\ &\leq h \sum_{\substack{N+1\leq p_j\leq 2N\\ p_{j+r+1}-p_{j+1}<h}} 1 + O(Ne^{-c_1\sqrt{\log N}}). \end{aligned} \tag{12.13}$$

For larger than average gaps between primes a similar argument shows, for $r \geq 0$,

$$\begin{aligned} Q_r^-(N,h) &= \sum_{N+1\leq p_j\leq 2N} \sum_{\substack{p_j\leq n<p_{j+1}\\ p_{j+r+1}>n+h}} 1 + O(Ne^{-c_1\sqrt{\log N}}) \\ &= \sum_{\substack{N+1\leq p_j\leq 2N\\ p_{j+r+1}-p_j>h}} \big(\min(p_{j+1}, p_{j+r+1} - \lfloor h\rfloor) - p_j\big) + O(Ne^{-c_1\sqrt{\log N}}) \\ &\leq \sum_{\substack{N+1\leq p_j\leq 2N\\ p_{j+r+1}-p_j>h}} (p_{j+r+1} - p_j) + O(Ne^{-c_1\sqrt{\log N}}). \end{aligned} \tag{12.14}$$

Next, we have

$$Q_r^+(N,h) = \sum_{\substack{n=N+1\\ \psi(n+h)-\psi(n)\geq\rho\log N}}^{2N} 1 + O(N^{\frac{1}{2}}), \tag{12.15}$$

where ρ can be taken to be any number in the range $r < \rho < r+1$, since the prime powers may be discarded with an error $\ll N^{\frac{1}{2}}$. Also

$$Q_r^-(N,h) = \sum_{\substack{n=N+1\\ \psi(n+h)-\psi(n)\leq\rho\log N}}^{2N} 1 + O(N^{\frac{1}{2}}), \tag{12.16}$$

where again ρ can be taken to be any number in the range $r < \rho < r+1$. Returning to (12.5), we see on applying Cauchy's inequality twice and using (12.15) that

$$\begin{aligned}\mathcal{M}(h,\rho) &\le \sum_{\substack{n=N+1\\ \psi(n+h)-\psi(n)\ge\rho\log N}}^{2N} \big(\psi(n+h)-\psi(n)\big)\big(\psi_R(n+h)-\psi_R(n)-C\log N\big)^2\\ &\le \left(\sum_{\substack{n=N+1\\ \psi(n+h)-\psi(n)\ge\rho\log N}}^{2N} 1\right)^{\frac14}\left(\sum_{n=N+1}^{2N}(\psi(n+h)-\psi(n))^4\right)^{\frac14}\\ &\qquad \left(\sum_{n=N+1}^{2N}(\psi_R(n+h)-\psi_R(n)-C\log N)^4\right)^{\frac12}\\ &= \big(Q_r^+(N,h)+O(N^{\frac12})\big)^{\frac14} M_4'(N,h,\psi)^{\frac14} M_4'(N,h,\psi_R,C)^{\frac12}.\end{aligned} \tag{12.17}$$

The same argument also shows that

$$-\mathcal{M}(h,\rho) \le \rho\log N\big(Q_r^-(N,h)+O(N^{\frac12})\big)^{\frac12} M_4'(N,h,\psi_R,C)^{\frac12}, \tag{12.18}$$

and therefore we conclude that for any $r < \rho < r+1$

$$\begin{aligned}-\big(Q_r^-(N,h)+O(N^{\frac12})\big)^{\frac12} &\le \frac{\mathcal{M}(h,\rho)}{\rho\log N M_4'(N,h,\psi_R,C)^{\frac12}}\\ &\le \big(Q_r^+(N,h)+O(N^{\frac12})\big)^{\frac14}\left(\frac{M_4'(N,h,\psi)^{\frac14}}{\rho\log N}\right)\end{aligned} \tag{12.19}$$

To prove the first part of Theorem 1.6 we estimate the moments $M_4'(N,h,\psi)$ and $M_4'(N,h,\psi_R,C)$ trivially when $h \gg 1$ using the inequality

$$|abcd| \le \frac14(a^4+b^4+c^4+d^4)$$

and the equation above (1.10) to see that

$$\begin{aligned}M_4'(N,h,\psi) &= \sum_{\substack{1\le m_i\le h\\ 1\le i\le 4}}\sum_{n=N+1}^{2N}\Lambda(n+m_1)\Lambda(n+m_2)\Lambda(n+m_3)\Lambda(n+m_4)\\ &\ll h^4\sum_{n\le 3N}\Lambda(n)^4\\ &\ll h^4 N\log^3 N,\end{aligned}$$

and similarly

$$\begin{aligned} M_4'(N,h,\psi_R,C) &\ll M_4'(N,h,\psi_R) + N\log^4 N \\ &\ll h^4 \sum_{n\le 3N} \Lambda_R(n)^4 + N\log^4 N \\ &\ll h^4 \log^4 R \sum_{n\le 3N} d(n)^4 \\ &\ll h^4 N \log^{19} N \end{aligned}$$

by (4.14). Hence, subject to (12.4), we see by (12.7), (12.13), and (12.19) that, for $r \ge 1$, and some positive constant C,

$$\sum_{\substack{N+1\le p_j\le 2N \\ p_{j+r+1}-p_{j+1}<h}} 1 \gg \frac{N}{\log^C N} \tag{12.20}$$

provided $\rho - \sqrt{\theta\rho} < \lambda < \rho$, $r < \rho < r+1$, and $0 < \theta < \frac{\vartheta}{2}$. Since ρ can be taken as close to r as we wish, we conclude that

$$\Xi_r \le r - \sqrt{\theta r},$$

where unconditionally we may take any $0 < \theta < 1/4$. This proves the first part of Theorem 1.6. If we assume $\vartheta = 1$ we can take $0 < \theta < \frac{1}{2}$. The corresponding result for larger than average gaps between primes is proved in the same way.

In order to obtain positive proportion results, we need to use the generalization of the sieve upper bound (1.39) for prime k-tuples. This result states that for the function $\psi_{\boldsymbol{j}}(N)$ defined in (1.8) where $\boldsymbol{j} = (j_1, j_2, \ldots, j_r)$ with the j_i's distinct and $\mathfrak{S}(\boldsymbol{j}) \ne 0$

$$\psi_{\boldsymbol{j}}(N) \le (2^r r! + \epsilon)\mathfrak{S}(\boldsymbol{j})N, \tag{12.21}$$

see Theorem 5.7 of [16]. On applying this bound to the formulas leading to (1.15) we see that, subject to (12.4),

$$M_k(N,h,\psi) \ll N(\log N)^k, \tag{12.22}$$

which implies the same estimate holds for $M_k'(N,h,\psi)$. Next, as above

$$M_4'(N,h,\psi_R,C) \ll M_4'(N,h,\psi_R) + N\log^4 N,$$

and therefore assuming (1.47) we have, for $0 < \theta < \frac{1}{4}$,

$$M_4'(N,h,\psi_R,C) \ll N\log^4 N. \tag{12.23}$$

Using these estimates in (12.19) we obtain

$$\sum_{\substack{N+1\le p_j\le 2N \\ p_{j+r+1}-p_{j+1}<h}} 1 \gg \frac{N}{\log N} \tag{12.24}$$

under the same conditions as (12.20) and $0 < \theta < \frac{1}{4}$. This proves the remaining part of Theorem 1.6.

In an identical fashion we see that if $\mathcal{M}(h, \rho) < 0$ then assuming (1.47) we have

$$\sum_{\substack{N+1 \leq p_j \leq 2N \\ p_{j+r+1}-p_j > h}} (p_{j+r+1} - p_j) \gg N, \tag{12.25}$$

where $r \geq 0$, $r < \rho < r+1$, $\rho < \lambda < \rho + \sqrt{\theta\rho}$, and $0 < \theta < \frac{1}{4}$. Since ρ can be taken as close to $r+1$ and θ as close to $\frac{1}{4}$ as we wish, this completes the proof of Theorem 1.7.

References

[1] E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. A, **293** (1966), 1–18.

[2] E. Bombieri, J. B. Friedlander, and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Math. **156** (1986), no. 3-4, 203–251.

[3] A. Y. Cheer and D. A. Goldston, *Longer than average intervals containing no primes*, Transactions of the AMS, **304** No.2 (1987), 469–486.

[4] J.-R. Chen, *On the Goldbach problem and the sieve methods*, Sci. Sinica **21** (1978), 701–739.

[5] J. B. Friedlander and D. A. Goldston, *Some singular series averages and the distribution of Goldbach numbers in short intervals*, Illinois J. of Math. **39** No. 1,(1995), 158–180.

[6] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), 4–9.

[7] D. A. Goldston, *The second moment for prime numbers*, Quart. J. Math. Oxford (2) **35** (1984), 153–163.

[8] D. A. Goldston, *On Bombieri and Davenport's theorem concerning small gaps between primes*, Mathematika **39** (1992), 10–17.

[9] D. A. Goldston, *A lower bound for the second moment of primes in short intervals*, Expo. Math. **13** (1995), 366–376.

[10] D. A. Goldston and H. L. Montgomery, *Pair correlation of zeros and primes in short intervals,* Analytic Number Theory and Diophantine Problems, Birkhaüser, Boston, Mass., 1987, 183–203.

[11] D. A. Goldston and C. Y. Yıldırım, *On the second moment for primes in an arithmetic progression*, Acta Arithmetica **C.1** (2001), 85–104.

[12] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics. A foundation for computer science*, Second Edition. Addison-Wesley Publishing Company, Reading, MA, 1994.

[13] S. Graham, *An asymptotic estimate related to Selberg's sieve*, J. Number Theory **10** (1978), 83–94.

[14] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum': III On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.

[15] A. Hildebrand, *Über die punktweise Konvergenz von Ramanujan-Entwicklungen zahlentheoretischer Funktionen*, Acta Arithmetica **XLIV** (1984), 109–140.

[16] H. Halberstam and H. -E. Richert, *Sieve methods* Academic Press, London-New York, 1974.

[17] D. R. Heath-Brown, *The ternary Goldbach problem*, Rev. Mat. Iberoamericana **1** (1985), no. 1, 45–59.

[18] C. Hooley, *On the Barban-Davenport-Halberstam theorem XIII*, Acta Arithmetica **XCIV** (2000), no. 1, 53–86.

[19] Loo Keng Hua, *Introduction to number theory*, Translated from the Chinese by Peter Shiu, Springer-Verlag, Berlin-New York, 1982.

[20] M. N. Huxley, *Small differences between consecutive primes II*, Mathematika **24** (1977), 142–152.

[21] M. N. Huxley, *An application of the Fouvry-Iwaniec theorem*, Acta Arithmetica **XLIII** (1984), 441–443.

[22] H. Maier, *Small differences between prime numbers*, Michigan Math. J. **35** (1988), 323–344.

[23] H. L. Montgomery and K. Soundararajan, *Beyond pair correlation*, preprint 1999, 1–20.

[24] J. Pintz, *Very large gaps between consecutive primes*, J. of Number Theory **63** (1997), 286–301.

[25] A. Selberg, *Contributions to the theory of Dirichlet's L-functions*, Skrifter utgitt ay Det Norske Videnskaps-Akademi i Oslo I. Mat.-Naturv. Klasse (1946), No. 3, 1–62.

[26] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, Second edition. Edited and with a preface by D. R. Heath-Brown. The Clarendon Press, Oxford University Press, New York, 1986.