Proc. London Math. Soc. (3) 95 (2007) 199–247
doi:10.1112/plms/pdm010

HIGHER CORRELATIONS OF DIVISOR SUMS RELATED TO PRIMES II: VARIATIONS OF THE ERROR TERM IN THE PRIME NUMBER THEOREM

D. A. GOLDSTON AND C. Y. YILDIRIM

Abstract

We calculate the triple correlations for the truncated divisor sum $\lambda_R(n)$. The $\lambda_R(n)$ behave over certain averages just as the prime counting von Mangoldt function $\Lambda(n)$ does or is conjectured to do. We also calculate the mixed (with a factor of $\Lambda(n)$) correlations. The results for the moments up to the third degree, and therefore the implications for the distribution of primes in short intervals, are the same as those we obtained (in the first paper with this title) by using the simpler approximation $\Lambda_R(n)$. However, when $\lambda_R(n)$ is used, the error in the singular series approximation is often much smaller than what $\Lambda_R(n)$ allows. Assuming the Generalized Riemann Hypothesis (GRH) for Dirichlet L-functions, we obtain an $\Omega_\pm$-result for the variation of the error term in the prime number theorem. Formerly, our knowledge under GRH was restricted to Ω-results for the absolute value of this variation. An important ingredient in the last part of this work is a recent result due to Montgomery and Soundararajan which makes it possible for us to dispense with a large error term in the evaluation of a certain singular series average. We believe that our results on the sums $\lambda_R(n)$ and $\Lambda_R(n)$ can be employed in diverse problems concerning primes.

1. *Introduction*

In this paper we calculate the triple correlations of the short divisor sum defined by

$$\lambda_R(n) = \sum_{r \leqslant R} \frac{\mu^2(r)}{\phi(r)} \sum_{d|(r,n)} d\mu(d), \qquad \text{for } n \geqslant 1, \tag{1.1}$$

and $\lambda_R(n) = 0$ if $n \leqslant 0$. In the previous paper of this series [7] we gave the calculation of the correlations of

$$\Lambda_R(n) = \sum_{\substack{d|n \\ d \leqslant R}} \mu(d) \log(R/d), \qquad \text{for } n \geqslant 1, \tag{1.2}$$

and $\Lambda_R(n) = 0$ if $n \leqslant 0$. As can be seen from our results, these divisor sums tend to behave similarly to the prime counting von Mangoldt function $\Lambda(n)$, and thus they may sometimes be used in place of $\Lambda(n)$ when it is not possible to work directly with $\Lambda(n)$ itself. Since

$$\Lambda(n) = \sum_{d|n} \mu(d) \log(R/d), \qquad \text{for } n > 1,$$

$\Lambda_R(n)$ comes about as a surrogate for $\Lambda(n)$ by truncation. We can relate $\lambda_R(n)$ to $\Lambda_R(n)$ by interchanging the order of the summations in (1.1), thereupon the new inner sum can be evaluated (equation (2.15) below) and the contribution of its main term gives $\Lambda_R(n)$.

Goldston in unpublished work found $\lambda_R(n)$ while remedying the failure of the circle method in an application to the related problems of twin primes and short gaps between primes for

Received 3 July 2005; revised 24 August 2006; published online 17 April 2007.

2000 *Mathematics Subject Classification* 11N05 (primary), 11P32 (secondary).

The research of Goldston was supported by the NSF; that of Yıldırım was supported by TÜBİTAK.

which a starting point is the observation that

$$\sum_{n\leqslant N}\Lambda(n)\Lambda(n+k)=\int_0^1|S(\alpha)|^2e(-k\alpha)\,d\alpha+O(k\log^2 N),\tag{1.3}$$

where

$$S(\alpha)=\sum_{n\leqslant N}\Lambda(n)e(n\alpha),\qquad e(u)=e^{2\pi iu}.$$

For α close to the rational number a/r, we write $\alpha=a/r+\beta$, and approximate $S(\alpha)$ throughout $[0,1]$ by a sum of local approximations

$$\sum_{r\leqslant R}\sum_{\substack{1\leqslant a\leqslant r\\(a,r)=1}}\frac{\mu(r)}{\phi(r)}I\left(\frac{a}{r}+\beta\right),\qquad \text{where } I(u)=\sum_{n\leqslant N}e(nu).$$

However, the above double sum is equal to

$$\sum_{n\leqslant N}\lambda_R(n)e(n\beta),$$

which suggests that we replace $\Lambda(m)$ by $\lambda_R(m)$ in sums such as (1.3). Furthermore, Goldston showed that among sums of the form

$$\sum_{\substack{r\leqslant R\\r|n}}a(R,r)\qquad \text{with } a(R,1)=1,\ a(R,r)\in\mathbb{R},$$

$\lambda_R(n)$ is the best approximation to $\Lambda(n)$ in an L^2 sense. The proof involves a minimization which was solved in a more general setting by Selberg [**17**] for his upper bound sieve. Hooley's recent use of $\lambda_R(n)$ in [**13**, **14**] leans greatly on its origin in the Selberg sieve. It should further be mentioned that, as far as we know, Heath-Brown [**11**] was the first to use $\lambda_R(n)$ in additive prime number theory.

The correlations that we are interested in evaluating are

$$\mathcal{S}_k(N,\boldsymbol{j},\boldsymbol{a})=\sum_{n=1}^{N}\lambda_R(n+j_1)^{a_1}\lambda_R(n+j_2)^{a_2}\dots\lambda_R(n+j_r)^{a_r}\tag{1.4}$$

and

$$\tilde{\mathcal{S}}_k(N,\boldsymbol{j},\boldsymbol{a})=\sum_{n=1}^{N}\lambda_R(n+j_1)^{a_1}\lambda_R(n+j_2)^{a_2}\dots\lambda_R(n+j_{r-1})^{a_{r-1}}\Lambda(n+j_r)\tag{1.5}$$

where $\boldsymbol{j}=(j_1,j_2,\dots,j_r)$ and $\boldsymbol{a}=(a_1,a_2,\dots a_r)$, the j_i are distinct integers, $a_i\geqslant 1$ and $\sum_{i=1}^r a_i=k$. In (1.5) we assume that $r\geqslant 2$ and take $a_r=1$. For later convenience we define

$$\tilde{\mathcal{S}}_1(N,\boldsymbol{j},\boldsymbol{a})=\sum_{n=1}^{N}\Lambda(n+j_1)=\psi(N)+O(|j_1|\log N)\sim N\tag{1.6}$$

for $|j_1|=o(N/\log N)$ by the prime number theorem (as usual $\psi(x)=\sum_{n\leqslant x}\Lambda(n)$).

For $k=1$ and $k=2$ these correlations have been evaluated before [**4**, **13**, **14**], and the more general cases of n running through arithmetic progressions were also worked out [**6**, **11**, **14**].

Correlations which include in their summands factors such as $\Lambda(n)\Lambda(n+j)$, with $j\neq 0$, cannot be evaluated unconditionally; they are the subject of the Hardy–Littlewood prime r-tuple conjecture [**8**]. This conjecture states that for $\boldsymbol{j}=(j_1,j_2,\dots,j_r)$ with the j_i distinct integers,

$$\psi_{\boldsymbol{j}}(N)=\sum_{n=1}^{N}\Lambda(n+j_1)\Lambda(n+j_2)\dots\Lambda(n+j_r)\sim\mathfrak{S}(\boldsymbol{j})N\tag{1.7}$$

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

$$\mathfrak{S}(\boldsymbol{j}) = \prod_p \left(1 - \frac{1}{p}\right)^{-r} \left(1 - \frac{\nu_p(\boldsymbol{j})}{p}\right) \tag{1.8}$$

and $\nu_p(\boldsymbol{j})$ is the number of distinct residue classes modulo p that the j_i occupy. If $r = 1$, we see that $\mathfrak{S}(\boldsymbol{j}) = 1$, and for $|j_1| \leqslant N$, (1.7) reduces to (1.6), which is the only case where (1.7) has been proved. The cases $r = 2, 3$ will be of particular interest to us in this paper, the explicit expressions have been shown in [7] to be

$$\mathfrak{S}((0, j)) = \mathfrak{S}_2(j), \qquad (j \neq 0), \tag{1.9}$$

$$\mathfrak{S}((0, j_1, j_2)) = \mathfrak{S}_2((j_1, j_2))\mathfrak{S}_3(j_1 j_2(j_1 - j_2)), \qquad (j_1 \neq j_2,\ j_1 j_2 \neq 0), \tag{1.10}$$

where writing

$$p(n) = \begin{cases} n & \text{if } n \text{ is a prime,} \\ 1 & \text{otherwise,} \end{cases} \tag{1.11}$$

we define the singular series for $n \geqslant 1$ and $j \neq 0$ as

$$\mathfrak{S}_n(j) = \begin{cases} C_n G_n(j) H_n(j) & \text{if } p(n) \mid j, \\ 0 & \text{otherwise,} \end{cases} \tag{1.12}$$

in which

$$C_n = \prod_{\substack{p \\ p \neq n-1,\ p \neq n}} \left(1 - \frac{n-1}{(p-1)(p-n+1)}\right), \tag{1.13}$$

$$G_n(j) = \prod_{\substack{p \mid j \\ p = n-1 \text{ or } p = n}} \left(\frac{p}{p-1}\right), \tag{1.14}$$

$$H_n(j) = \prod_{\substack{p \mid j \\ p \neq n-1,\ p \neq n}} \left(1 + \frac{1}{p-n}\right). \tag{1.15}$$

Note that since $\mathfrak{S}(\boldsymbol{j}) = \mathfrak{S}(\boldsymbol{j} - \boldsymbol{j_1})$ for $\boldsymbol{j_1}$ a vector with j_1 in every component, no loss of generality is incurred when the first components of the vectors in the arguments of $\mathfrak{S}$ in (1.9) and (1.10) are taken to be 0.

Gallagher [2] proved that the moments

$$M_k(N, h, \psi) = \sum_{n=1}^{N} (\psi(n+h) - \psi(n))^k \qquad (k \in \mathbb{Z}^+) \tag{1.16}$$

can be calculated from the prime r-tuple conjecture (1.7) for $h \sim \lambda \log N$ as $N \to \infty$, with λ a positive constant. For this purpose Gallagher showed that

$$\sum_{\substack{1 \leqslant j_1, j_2, \ldots, j_r \leqslant h \\ \text{distinct}}} \mathfrak{S}(\boldsymbol{j}) \sim h^r \qquad (h \to \infty). \tag{1.17}$$

The calculation of the moments (1.16) was carried out in [7] by expressing them in terms of the quantities (1.7) for which the prime r-tuple conjecture is assumed, with the result that

$$M_k(N, h, \psi) \sim N(\log N)^k \sum_{r=1}^{k} \left\{ {k \atop r} \right\} \lambda^r \qquad (N \to \infty,\ h \sim \lambda \log N,\ \lambda \ll 1), \tag{1.18}$$

where $\left\{ {k \atop r} \right\}$ denotes the Stirling numbers of the second type.

For larger h the appropriate moments to study are

$$\mu_k(N,h,\psi) = \sum_{n=1}^{N} (\psi(n+h) - \psi(n) - h)^k \,. \tag{1.19}$$

Assuming the Hardy–Littlewood conjecture in the strong form

$$\psi_{\boldsymbol{j}}(x) = \mathfrak{S}(\boldsymbol{j})x + O(N^{1/2+\epsilon}) \tag{1.20}$$

uniformly for $1 \leqslant r \leqslant k$, $1 \leqslant x \leqslant N$ and distinct j_i satisfying $1 \leqslant j_i \leqslant h$, Montgomery and Soundararajan [16] proved that

$$\mu_k(N,h,\psi) \sim (1\cdot 3\cdot \ldots \cdot (k-1))N\left(h\log\frac{N}{h}\right)^{k/2} \qquad \text{if } k \text{ is even,} \tag{1.21}$$

$$\mu_k(N,h,\psi) \ll N(h\log N)^{k/2}\left(\frac{h}{\log N}\right)^{-1/(8k)} + h^k N^{1/2+\epsilon} \qquad \text{if } k \text{ is odd,} \tag{1.22}$$

uniformly for $(\log N)^{1+\delta} \leqslant h \leqslant N^{1/k-\epsilon}$ (with any fixed $\delta > 0$). They also conjectured upon heuristics that

$$\mu_k(N,h,\psi) = ([2\mid k](1\cdot 3\cdot \ldots \cdot (k-1)) + o(1))N\left(h\log\frac{N}{h}\right)^{k/2}$$

holds uniformly for $(\log N)^{1+\delta} \leqslant h \leqslant N^{1-\delta}$ for each fixed k (see (1.49) below for the notation $[2\mid k]$). Their proof depends on the estimation of the quantities

$$R_r(h) = \sum_{\substack{1\leqslant j_1,j_2,\ldots,j_r\leqslant h\\ \text{distinct}}} \mathfrak{U}((j_1,\ldots,j_r)), \tag{1.23}$$

where

$$\mathfrak{U}((j_1,\ldots,j_r)) = \sum_{\mathcal{J}\subset\{j_1,\ldots,j_r\}} (-1)^{r-|\mathcal{J}|}\mathfrak{S}(\mathcal{J}) \tag{1.24}$$

($R_0(h)$ and $\mathfrak{S}(\emptyset)$ are taken to be 1), as

$$R_r(h) = (1\cdot 3\cdot \ldots \cdot (r-1))(-h\log h + Ah)^{r/2} + O_r(h^{r/2-1/(7r)+\epsilon}), \qquad \text{for } r \text{ even,} \tag{1.25}$$

$$R_r(h) \ll h^{r/2-1/(7r)+\epsilon}, \qquad \text{for } r \text{ odd} \tag{1.26}$$

($A = 2-\gamma-\log 2\pi$, and γ denotes Euler's constant). Gallagher's result (1.17) can be deduced from these. Note that it is easy to see that $R_1(h) = 0$, and for $r = 2$ we know from Goldston [3] that (1.25) holds with the much smaller error term $O(h^{1/2+\epsilon})$.

Only the first moment is known unconditionally as a simple consequence of the prime number theorem. The work of Goldston and Montgomery [5] reveals, upon assuming the Riemann Hypothesis, an equivalence between the asymptotic formulae for the second moment and the pair correlation conjecture for the zeros of the Riemann zeta-function.

From the surrogate prime-counting function $\lambda_R(n)$, we write

$$\psi_R(x) = \sum_{n\leqslant x} \lambda_R(n), \tag{1.27}$$

and we wish to examine the moments $M(N, h, \psi_R)$ defined as in (1.16). We have

$$M_k(N, h, \psi_R) = \sum_{n=1}^{N} \left(\sum_{1 \leqslant m \leqslant h} \lambda_R(n+m) \right)^k$$
$$= \sum_{\substack{1 \leqslant m_i \leqslant h \\ 1 \leqslant i \leqslant k}} \sum_{n=1}^{N} \lambda_R(n+m_1)\lambda_R(n+m_2)\dots\lambda_R(n+m_k).$$

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 $\sum_{1 \leqslant i \leqslant r} a_i = k$. Grouping the terms leads to the expression

$$M_k(N, h, \psi_R) = \sum_{r=1}^{k} \sum_{\substack{a_1, a_2, \dots, a_r \\ a_i \geqslant 1, \sum a_i = k}} \binom{k}{a_1, a_2, \dots, a_r} \sum_{1 \leqslant j_1 < j_2 < \dots < j_r \leqslant h} \mathcal{S}_k(N, \boldsymbol{j}, \boldsymbol{a}), \tag{1.28}$$

where $\mathcal{S}_k(N, \boldsymbol{j}, \boldsymbol{a})$ is the correlation given in (1.4). Our main result on these correlations is the following theorem.

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

$$\mathcal{S}_k(N, \boldsymbol{j}, \boldsymbol{a}) = \big(\mathcal{C}_k(\boldsymbol{a})\mathfrak{S}(\boldsymbol{j}) + o(1)\big)N(\log R)^{k-r} + O(R^k), \tag{1.29}$$

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

$$\mathcal{C}_1(1) = 1, \qquad \mathcal{C}_2(2) = 1, \qquad \mathcal{C}_2(1,1) = 1, \qquad \mathcal{C}_3(3) = \tfrac{3}{4}, \qquad \mathcal{C}_3(2,1) = 1, \qquad \mathcal{C}_3(1,1,1) = 1.$$

(As a notational convention extra parentheses have been dropped, so, for example, we write $\mathcal{C}_2(1,1)$ instead of $\mathcal{C}_2((1,1))$.) The method of proof used in this paper may be carried out for $k > 3$, but the calculation of the constants $\mathcal{C}_k(\boldsymbol{a})$ and controlling the error terms become extremely complicated even for $k = 4$. In the fourth paper in this series, employing $\Lambda_R(n)$, we resort to a different method in order to tackle the assertion of Theorem 1 for all k, wherewith a way for calculating the constants $\mathcal{C}_k(\boldsymbol{a})$ for small k is shown. We also believe that the error term $O(R^k)$ occurring in (1.29) can be reduced in size. In Hooley's method [**14**] for the special case $S_2(N, (0), (2))$, the error term $O(R^2)$ does not arise at all.

Letting $m = n + \min_i j_i$ in the sum of (1.4), and then shifting the summation range to extend from 1 to N again, we pick up an error $O(|\min_i j_i| N^{\epsilon})$ since $\lambda_R(n) \ll n^{\epsilon}$. This error is absorbed in the error term $o(N)$ under the conditions of the theorem. Also, as was remarked after (1.15), $\mathfrak{S}(\boldsymbol{j})$ is not affected by this shift. Hence in proving Theorem 1 we may take $j_1 = 0$, and $j_2, \dots, j_r$ all positive. To see the upper bound for $\lambda_R(n)$, note that with $n' = \prod_{p|n,\, p \leqslant R} p$, one has $\lambda_R(n) = \lambda_R(n')$, so

$$\lambda_R(n) = \sum_{r \leqslant R} \frac{\mu^2(r)\mu((r,n'))\phi((r,n'))}{\phi(r)} = \sum_{\substack{t \leqslant R \\ (t,n')=1}} \frac{\mu^2(t)}{\phi(t)} \sum_{\substack{s | n' \\ s \leqslant R/t}} \mu(s) \ll d(n') \log 2R. \tag{1.30}$$

We now apply Theorem 1 in (1.28), and obtain upon using (1.17) that for $h \ll N^{1-\epsilon}$ and $h \to \infty$, $R = N^{\theta_k}$ with $0 < \theta_k < 1/k$ for M_k,

$$\begin{gathered} M_1(N, h, \psi_R) \sim Nh, \qquad M_2(N, h, \psi_R) \sim Nh^2 + Nh\log R, \\ M_3(N, h, \psi_R) \sim Nh^3 + 3Nh^2 \log R + \tfrac{3}{4} Nh \log^2 R. \end{gathered} \tag{1.31}$$

The choice $h = \lambda \log N$ renders full meaning to (1.31) allowing us to state the following result.

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

$$M_1(N, h, \psi_R) \sim \lambda N \log N, \qquad M_2(N, h, \psi_R) \sim (\theta_2\lambda + \lambda^2) N \log^2 N,$$
$$M_3(N, h, \psi_R) \sim (\tfrac{3}{4}{\theta_3}^2\lambda + 3\theta_3\lambda^2 + \lambda^3) N \log^3 N. \tag{1.32}$$

We next consider the mixed moments

$$\tilde{M}_k(N, h, \psi_R) = \sum_{n=1}^{N} (\psi_R(n+h) - \psi_R(n))^{k-1} (\psi(n+h) - \psi(n)) \tag{1.33}$$

for $k \geqslant 2$, while for $k = 1$ we take $\tilde{M}_1(N, h, \psi_R) = M_1(N, h, \psi)$. Writing

$$\psi(x) = x + E(x), \tag{1.34}$$

we have for $1 \leqslant h \leqslant N$,

$$\begin{aligned} M_1(N, h, \psi) &= \sum_{n=1}^{N} \sum_{n<k\leqslant n+h} \Lambda(k) = \sum_{k=2}^{N+h} \Lambda(k) \sum_{n=\max(1,k-h)}^{\min(k-1,N)} 1 \\ &= \sum_{k=2}^{h} (k-1)\Lambda(k) + \sum_{k=h+1}^{N} h\Lambda(k) + \sum_{k=N+1}^{N+h} (N+h-k+1)\Lambda(k) \\ &= \psi(N+h) - \psi(N) - \psi(h) - \int_2^h \psi(t)\,dt + \int_N^{N+h} \psi(t)\,dt \\ &= Nh + E(N+h) - E(N) - E(h) - \int_2^h E(t)\,dt + \int_N^{N+h} E(t)\,dt + O(1), \end{aligned} \tag{1.35}$$

where partial summation on (1.34) has also been used. The prime number theorem says that $E(x) = o(x)$, so that we obtain for $1 \leqslant h \leqslant N$ as $N \to \infty$,

$$\tilde{M}_1(N, h, \psi_R) = M_1(N, h, \psi) \sim Nh. \tag{1.36}$$

If the Riemann Hypothesis is assumed, then it is known that $E(x) \ll x^{1/2} \log^2 x$, giving

$$\tilde{M}_1(N, h, \psi_R) = M_1(N, h, \psi) = Nh + O(N^{1/2} h \log^2 N). \tag{1.37}$$

For $k \geqslant 2$, leaving out the details that were included in [7], we have

$$\tilde{M}_k(N, h, \psi_R) = \sum_{r=2}^{k} \sum_{\substack{a_1, a_2, \ldots, a_{r-1} \\ a_i \geqslant 1, \sum a_i = k-1}} \frac{1}{(r-1)!} \binom{k-1}{a_1, a_2, \ldots, a_{r-1}} W_r(N, \boldsymbol{j}, \boldsymbol{a}) + O(RN^{\epsilon}), \tag{1.38}$$

where

$$W_r(N, \boldsymbol{j}, \boldsymbol{a}) = \sum_{i=1}^{r-1} (\mathcal{L}_1(R))^{a_i} \sum_{\substack{1 \leqslant j_1, j_2, \ldots, j_{r-1} \leqslant h \\ \text{distinct}}} \tilde{\mathcal{S}}_{k-a_i}(N, \boldsymbol{j_i}, \boldsymbol{a_i}) + \sum_{\substack{1 \leqslant j_1, j_2, \ldots, j_r \leqslant h \\ \text{distinct}}} \tilde{\mathcal{S}}_k(N, \boldsymbol{j}, \boldsymbol{a}),$$

with $\boldsymbol{j_i} = (j_1, j_2, \ldots, j_{i-1}, j_{i+1}, \ldots, j_{r-1}, j_i)$ and $\boldsymbol{a_i} = (a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{r-1}, 1)$, and

$$\mathcal{L}_k(R) := \sum_{\substack{r \leqslant R \\ (r,k)=1}} \frac{\mu^2(r)}{\phi(r)}. \tag{1.39}$$

(It is easily seen that $\mathcal{L}_1(R) \ll \log 2R$ for $R \geqslant 1$. A precise estimation of this sum due to Hildebrand [9] is given in (2.15) below.)

Formula (1.38) reduces the calculation of the mixed moments to the calculation of mixed correlations. Our results depend on the extent of uniformity in the distribution of primes in arithmetic progressions. We let

$$\psi(x;q,a) = \sum_{\substack{n\leqslant x\\ n\equiv a(q)}} \Lambda(n), \tag{1.40}$$

and on taking

$$E(x;q,a) = \psi(x;q,a) - [(a,q)=1]\frac{x}{\phi(q)} \tag{1.41}$$

(where we have used the Iverson notation (1.49) below) we see that the estimate we need is, for some fixed $0<\vartheta\leqslant 1$,

$$\sum_{1\leqslant q\leqslant x^{\vartheta-\epsilon}} \max_{\substack{a\\ (a,q)=1}} |E(x;q,a)| \ll \frac{x}{\log^{\mathcal{A}} x}, \tag{1.42}$$

for any $\epsilon>0$, any $\mathcal{A}=\mathcal{A}(\epsilon)>0$, and x sufficiently large (see [**1**, Chapter 28]). This is a weakened form of the Bombieri–Vinogradov theorem if $\vartheta=\frac{1}{2}$, and therefore (1.42) holds unconditionally if $\vartheta\leqslant\frac{1}{2}$. Elliott and Halberstam conjectured that (1.42) is true with $\vartheta=1$. The range of R where our results on mixed correlations hold depends on ϑ in (1.42). We prove the following result.

THEOREM 2. *Given $2\leqslant k\leqslant 3$, let $\boldsymbol{j}=(j_1,j_2,\ldots,j_r)$ and $\boldsymbol{a}=(a_1,a_2,\ldots,a_r)$, where $r\geqslant 2$, $a_r=1$, and where the j_i are distinct integers, and $a_i\geqslant 1$ with $\sum_{i=1}^{r} a_i = k$. Then we have, for $N^\epsilon \ll R \ll N^{\vartheta/(k-1)-\epsilon}$ where (1.42) holds with ϑ, and $\max_i |j_i| \ll N^{1/(k-1)-\epsilon}$,*

$$\tilde{\mathcal{S}}_k(N,\boldsymbol{j},\boldsymbol{a}) = \big(\mathfrak{S}(\boldsymbol{j})+o(1)\big)N(\log R)^{k-r}. \tag{1.43}$$

In proving Theorem 2 we may take the argument of Λ to be n, the error arising from arranging this by shifts of the range of summation being $O(|j_r|N^\epsilon)+O((\max|j_i|)^{1+\epsilon})$.

Using (1.43) in (1.38), and (1.17), we find that

$$\begin{aligned} M_2(N,h,\psi_R) &\sim Nh^2 + Nh\log R,\\ M_3(N,h,\psi_R) &\sim Nh^3 + 3Nh^2\log R + Nh\log^2 R, \end{aligned} \tag{1.44}$$

and similar to Corollary 1 we have the following.

COROLLARY 2. *For $h\sim\lambda\log N$, $\lambda\ll 1$, and $R=N^{\theta_k}$, where θ_k is fixed and $0<\theta_k<\vartheta/(k-1)$ for $k=2$ or 3, we have*

$$\begin{gathered} \tilde{M}_1(N,h,\psi_R)\sim \lambda N\log N, \qquad \tilde{M}_2(N,h,\psi_R)\sim(\theta_2\lambda+\lambda^2)N\log^2 N,\\ \tilde{M}_3(N,h,\psi_R)\sim({\theta_3}^2\lambda+3\theta_3\lambda^2+\lambda^3)N\log^3 N. \end{gathered} \tag{1.45}$$

The results for the correlations up to and including the third order of $\lambda_R(n)$ and $\Lambda_R(n)$ coincide asymptotically, thereby implying the results (Theorems 1.6 and 1.7) of [**7**] on primes in short intervals, in particular,

$$\liminf_{n\to\infty}\left(\frac{p_{n+r}-p_n}{\log p_n}\right) \leqslant r-\tfrac{1}{2}\sqrt{r}. \tag{1.46}$$

For longer intervals, instead of (1.42) we shall have recourse to Hooley's [**12**] bound depending on GRH that, for all $q\leqslant x$,

$$\sum_{\substack{1\leqslant a\leqslant q\\ (a,q)=1}} \max_{u\leqslant x} |E(u;q,a)|^2 \ll x(\log x)^4. \tag{1.47}$$

It is easy to see that the same bound holds when the sum is taken over all $1 \leqslant a \leqslant q$. For $(a,q) > 1$ we have

$$E(u;q,a) = \psi(u;q,a) \leqslant \sum_{\substack{n \leqslant x \\ n \equiv a (\mathrm{mod}\, q)}} \Lambda(n),$$

where it is seen that only those n which are powers of a prime divisor of (a,q) contribute. The sum is not void only if (a,q) has just one prime factor, say p, in which case its value is at most

$$(\log p) \left\lfloor \frac{\log x}{\log p} \right\rfloor \leqslant \log x.$$

So the addition of $(a,q) > 1$ terms in the sum of (1.47) brings in at most $q \log^2 x = o(x \log^4 x)$.

We prove the following.

THEOREM 3. *Assume the Generalized Riemann Hypothesis. For any arbitrarily small but fixed $\eta > 0$, and for sufficiently large N, with $\log^{14} N \ll h \ll N^{1/7-\epsilon}$ and writing $h = N^{\alpha}$, there exist $n_1, n_2 \in [N+1, 2N]$ such that*

$$\begin{aligned} \psi(n_1 + h) - \psi(n_1) - h &> \left(\tfrac{1}{2}\sqrt{1-5\alpha} - \eta\right)(h \log N)^{1/2} \\ \psi(n_2 + h) - \psi(n_2) - h &< -\left(\tfrac{1}{2}\sqrt{1-5\alpha} - \eta\right)(h \log N)^{1/2}. \end{aligned} \tag{1.48}$$

This is a new development in the sense that formerly our knowledge under GRH was restricted to lower-bound estimates for the absolute value of the variation of the error term in the prime number theorem. The strongest such results were attained in [**6**] in the more general case of primes in an arithmetic progression which yielded as a special case

$$\max_{x \leqslant y \leqslant 2x} |\psi(y+h) - \psi(y) - h| \gg_{\epsilon} (h \log x)^{1/2}$$

for $1 \leqslant h \leqslant N^{1/3-\epsilon}$. A proof of this is included in § 10. In fact the general case was also obtained by using the correlations of $\lambda_R(n)$. There only the first and second level correlations were employed; nevertheless, in the more general case of $n \in [N+1, 2N]$ running through an arithmetic progression, we have $n \equiv a (\mathrm{mod}\, q)$.

NOTATION. In this paper N is always a large natural number and p is a prime number. The largest square-free positive integer divisor of a non-zero integer j will be denoted by j^*. If a lower limit is unspecified in a summation it will be understood that the sum starts at 1. When a sum is denoted with a dash as $\sum'$ this always indicates that we will sum over all variables expressed by inequalities in the conditions of summation and these variables will all be pairwise coprime. 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 each time it occurs. We will also use the Iverson notation of putting brackets around a truth-valued statement $P(x)$ which means that

$$[P(x)] = \begin{cases} 1 & \text{if } P(x) \text{ is true}, \\ 0 & \text{if } P(x) \text{ is false}. \end{cases} \tag{1.49}$$

As usual, (a,b) denotes the greatest common divisor of a and b and $[a_1, a_2, \ldots, a_n]$ denotes the least common multiple of $a_1, a_2, \ldots, a_n$. If $k = 0$, the condition $d \mid k$ means d can be any positive integer; and we will take $(0,a) = 0$ for $a \neq 0$. We define $\phi_2(p) = p - 2$ on the primes, $\phi_2(1) = 1$, and extend the definition to square-free integers multiplicatively. For arithmetical functions α and β, we will sometimes write $\alpha \cdot \beta(n)$ for the product $\alpha(n)\beta(n)$, and

$$\frac{\alpha}{\beta}(n) \quad \text{for the quotient} \quad \frac{\alpha(n)}{\beta(n)}.$$

2. *Lemmas*

Let us recall some well-known facts to be used in this paper.

We shall need the elementary estimates (see [6]), for an integer $k \neq 0$,

$$\sum_{p|k} \frac{\log p}{p} \ll \log\log 3|k|, \tag{2.1}$$

$$m(k) := \sum_{d|k} \frac{\mu^2(d)}{\sqrt{d}} = \prod_{p|k}\left(1+\frac{1}{\sqrt{p}}\right) \ll \exp\left(\frac{c\sqrt{\log|k|}}{\log\log 3|k|}\right), \tag{2.2}$$

$$\prod_{p|k}\left(1+\frac{1}{p}\right) \ll \log\log 3|k|, \qquad \mathfrak{S}_n(k) \ll \log\log 3|k|, \tag{2.3}$$

which follow from the prime number theorem.

For a multiplicative function $f(n)$ we have

$$\sum_{d|n} \mu^2(d) f(d) \log d = \left(\sum_{p|n} \frac{f(p)\log p}{1+f(p)}\right) \prod_{p|n}(1+f(p)) \qquad (n \neq 0). \tag{2.4}$$

If $f : \mathbb{N} \to \mathbb{R}$ is a multiplicative function satisfying $0 \leqslant f(p^m) \leqslant \alpha_1\, \alpha_2^m$ at all prime powers, with constants $\alpha_1 > 0$ and $0 < \alpha_2 < 2$, then we have, uniformly for $x \geqslant 2$,

$$\sum_{n \leqslant x} f(n) \ll_{\alpha_1,\alpha_2} \frac{x}{\log x} \exp \sum_{p \leqslant x} \frac{f(p)}{p}. \tag{2.5}$$

This result, quoted from [9] (which refers to [10] for the proof of a sharper version), helps us see that for monic polynomials P_i,

$$\prod_{p|n} \frac{P_1(p)}{P_2(p)}$$

behaves on average the same as $n^{\deg P_1 - \deg P_2}$. In particular, we have

$$\sum_{n\leqslant x}\prod_{p|n}\frac{P_1(p)}{P_2(p)} \ll x \qquad (\deg P_1 = \deg P_2), \tag{2.6}$$

$$\sum_{n\leqslant x}\prod_{p|n}\frac{P_1(p)}{P_2(p)} \ll \log x \qquad (1+\deg P_1 = \deg P_2), \tag{2.7}$$

$$\sum_{x_1<n\leqslant x_2}\prod_{p|n}\frac{P_1(p)}{P_2(p)} \ll 1+\log\left(\frac{x_2}{x_1}\right) \qquad (1+\deg P_1 = \deg P_2), \tag{2.8}$$

$$\sum_{n>x}\frac{1}{n^\alpha}\prod_{p|n}\frac{P_1(p)}{P_2(p)} \ll_\alpha \frac{1}{x^\alpha} \qquad (1+\deg P_1 = \deg P_2,\ \text{fixed } \alpha>0), \tag{2.9}$$

$$\sum_{n>x}\log n\prod_{p|n}\frac{1}{P_2(p)} \ll \frac{\log x}{x} \qquad (\deg P_2 = 2). \tag{2.10}$$

Here (2.6) follows from a direct application of (2.5); (2.7) and (2.8) can be obtained by partial summation on (2.6); and to get (2.9) one may split the sum into ranges $(x, 2x]$, $(2x, 4x]$, $\ldots$ and apply (2.8) to each part. Then (2.10) is shown by partial summation on (2.9) with $\alpha = 1$. We will also need

$$\sum_{n\leqslant x}\frac{\mu^2(n)m(n)}{\sqrt{n}} \ll \sqrt{x}; \tag{2.11}$$

to see this we apply (2.5) with $f(n) = m(n)$ and then do partial summation.

We also quote a Perron-formula type of a result of Titchmarsh [**18**, § 3.12]. It will be used in proving Lemmas 2 and 3 which are needed in § 7. Let

$$A(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}, \qquad \text{for } \sigma = \Re s > 1.$$

Assume that $a_n = O(\mathfrak{a}(n))$, with $\mathfrak{a}(n)$ non-decreasing, and

$$A(s) = \sum_{n=1}^{\infty} \frac{|a_n|}{n^\sigma} = O\left(\frac{1}{(\sigma-1)^\alpha}\right) \qquad (\sigma \to 1^+). \tag{2.12}$$

For $c > 0$ and $c + \sigma > 1$, we have, for $x \notin \mathbb{N}$ and X the nearest integer to x,

$$\sum_{1 \leqslant n < x} \frac{a_n}{n^s} = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} A(s+w) \frac{x^w}{w}\, dw + O\left(\frac{x^c}{T(\sigma+c-1)^\alpha}\right) + O\left(\frac{\mathfrak{a}(2x) x^{1-\sigma} \log x}{T}\right) + O\left(\frac{\mathfrak{a}(X) x^{1-\sigma}}{T|x-X|}\right). \tag{2.13}$$

Our first lemma is a generalization of a result of Hildebrand [**9**].

Lemma 1. *Let P_1 and P_2 be monic polynomials such that $\deg P_2 = 1 + \deg P_1$ and $P_2(p) \neq 0$ for prime p. We have for each positive integer k, uniformly for $x \geqslant 1$,*

$$\sum_{\substack{n \leqslant x \\ (n,k)=1}} \mu^2(n) \prod_{p|n} \frac{P_1(p)}{P_2(p)} = \prod_p \left(1 + \frac{(p-1)P_1(p) - P_2(p)}{pP_2(p)}\right) \prod_{p|k} \left(\frac{P_2(p)}{P_1(p)+P_2(p)}\right) \times \left[\log x + \gamma + \sum_p \frac{P_2(p) - (p-2)P_1(p)}{(p-1)(P_1(p)+P_2(p))} \log p + \sum_{p|k} \frac{P_1(p)\log p}{P_1(p)+P_2(p)}\right] + O\left(\frac{m(k)}{\sqrt{x}}\right). \tag{2.14}$$

Hildebrand's result is the special case

$$\mathcal{L}_k(x) = \sum_{\substack{n \leqslant x \\ (n,k)=1}} \frac{\mu^2(n)}{\phi(n)} = \frac{\phi(k)}{k} \left(\log x + \gamma + \sum_p \frac{\log p}{p(p-1)} + \sum_{p|k} \frac{\log p}{p}\right) + O\left(\frac{m(k)}{\sqrt{x}}\right). \tag{2.15}$$

Note for future use that, by partial summation on (2.15),

$$\sum_{n \leqslant x} \frac{\mu^2(n)}{\phi(n)} \log\left(\frac{x}{n}\right) = \tfrac{1}{2} \log^2 x + O(\log x). \tag{2.16}$$

Another special case that we will use is

$$\sum_{\substack{n\leqslant x\\(n,k)=1}} \mu^2(n)\prod_{p\mid n}\frac{p^2-p-1}{(p-1)^3}$$
$$= \prod_p\left(1+\frac{p-2}{p(p-1)^2}\right)\prod_{p\mid k}\left(\frac{(p-1)^3}{p^3-2p^2+2p-2}\right)$$
$$\times\left\{\log x+\gamma+\sum_p\frac{(2p-3)\log p}{(p-1)(p^3-2p^2+2p-2)}+\sum_{p\mid k}\frac{(p^2-p-1)\log p}{p^3-2p^2+2p-2}\right\}$$
$$+O\left(\frac{m(k)}{\sqrt{x}}\right). \tag{2.17}$$

Lemma 2. *There exists a constant C such that for all $x\geqslant 1$,*

$$\left|\sum_{n\leqslant x}\frac{\mu(n)\phi_2(n)}{n\phi(n)}\right|\leqslant C. \tag{2.18}$$

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

$$\sum_{n\leqslant x}\mu^2(n)\prod_{p\mid n}\frac{3p-4}{(p-1)(\sqrt{p}-1)} = P(1)x^{1/2}\log^2 x + Dx^{1/2}\log x + (E+o(1))x^{1/2}, \tag{2.19}$$

where $P(s)$ is defined below in (3.20), and D and E are constants specified in (3.24).

In Sections 4, 5 and 8 we will need the following.

Lemma 4. *For non-zero integers j and k, we have, uniformly in $x\geqslant 1$,*

$$\sum_{\substack{n\leqslant x\\(n,k)=1}}\frac{\mu(n)\mu\cdot\phi((n,j))}{\phi^2(n)}$$
$$=\{1-[2\nmid k]\mu((2,j))\}C_2\prod_{\substack{p\mid k\\p>2}}\frac{(p-1)^2}{p(p-2)}\prod_{\substack{p\mid j\\p\nmid k\\p>2}}\left(\frac{p-1}{p-2}\right)+O\left(\frac{d(j')j'}{x\phi(j')}\right), \tag{2.20}$$

where $j'=j^*/(j^*,k)$.

From (2.20) we derive

$$-\sum_{n\leqslant x}\frac{\mu(n)\,\mu\cdot\phi((n,j))\log n}{\phi^2(n)}$$
$$=\begin{cases}\mathfrak{S}_2(j)\left[\displaystyle\sum_{p\nmid j}\frac{\log p}{p(p-2)}-\sum_{p\mid j}\frac{\log p}{p}\right]+O\left(\dfrac{j^*d(j^*)\log 2x}{\phi(j^*)x}\right) & \text{if } 2\mid j,\\ \mathfrak{S}_2(2j)\dfrac{\log 2}{2}+O\left(\dfrac{j^*d(j^*)\log 2x}{\phi(j^*)x}\right) & \text{if } 2\nmid j.\end{cases} \tag{2.21}$$

For use in § 9 we prove the following.

LEMMA 5. *For even $J \neq 0$ with $k \mid J$, and $J \ll x^A$ (for any fixed $A > 0$), we have, uniformly in $x \geqslant 1$,*

$$\sum_{n\leqslant x} \frac{\mu(n)d(n)}{\phi(n)\phi_2\left(\frac{n}{(n,2)}\right)} \frac{\mu}{d}((n,J))\frac{\mu}{\phi}((n,k))\phi_2\left(\left(\frac{n}{(n,2)},J\right)\right)$$
$$= 2\,[2\nmid k]\prod_{p\nmid J}\left(1-\frac{2}{(p-1)(p-2)}\right)\prod_{\substack{p>2\\ p\mid J\\ p\nmid k}}\left(1+\frac{1}{p-1}\right)\prod_{\substack{p>2\\ p\mid k}}\left(1-\frac{1}{(p-1)^2}\right)+O(x^{-1+\epsilon}). \quad (2.22)$$

3. *Proofs of the lemmas*

Proof of Lemma 1. We follow Hildebrand's way [9] of obtaining (2.15). Let, for $n \in \mathbb{N}$,

$$f_k(n) = \begin{cases} n\mu^2(n)\prod_{p\mid n}\dfrac{P_1(p)}{P_2(p)} & \text{if } (n,k)=1,\\ 0 & \text{if } (n,k)>1. \end{cases} \quad (3.1)$$

Also define g_k, the Möbius transform of f_k, through

$$f_k(n) = \sum_{d\mid n} g_k(d) \qquad (n\in\mathbb{N}). \quad (3.2)$$

We have

$$\sum_{\substack{n\leqslant x\\ (n,k)=1}} \mu^2(n)\prod_{p\mid n}\frac{P_1(p)}{P_2(p)} = \sum_{n\leqslant x}\frac{f_k(n)}{n} = \sum_{n\leqslant x}\frac{g_k(n)}{n}\sum_{m\leqslant x/n}\frac{1}{m}$$
$$= \sum_{n\leqslant x}\frac{g_k(n)}{n}\left(\log\frac{x}{n}+\gamma+O\left(\frac{n}{x}\right)\right). \quad (3.3)$$

The arithmetical functions f_k and g_k are multiplicative, their values at the prime powers are

$$f_k(p^m) = \begin{cases} \dfrac{pP_1(p)}{P_2(p)} & \text{for } m=1 \text{ and } p\nmid k,\\ 0 & \text{otherwise,} \end{cases} \quad (3.4)$$

and since $g_k(p^m) = f_k(p^m) - f_k(p^{m-1})$,

$$g_k(p^m) = \begin{cases} \dfrac{pP_1(p)}{P_2(p)} - 1 & \text{for } m=1 \text{ and } p\nmid k,\\ \dfrac{-pP_1(p)}{P_2(p)} & \text{for } m=2 \text{ and } p\nmid k,\\ -1 & \text{for } m=1 \text{ and } p\mid k,\\ 0 & \text{for } m=2 \text{ and } p\mid k, \text{ or } m>2. \end{cases} \quad (3.5)$$

We see that $\sum_{p,\, m\geqslant 1} |g_k(p^m)|/p^m$ is convergent, so that $\sum_{n=1}^{\infty} g_k(n)/n$ is absolutely convergent, giving

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{g_k(n)}{n} &= \prod_p\left(1+\sum_{m\geqslant 1}\frac{g_k(p^m)}{p^m}\right)\\ &= \prod_{p\mid k}\left(1-\frac{1}{p}\right)\prod_{p\nmid k}\left(1+\frac{pP_1(p)/P_2(p)-1}{p}-\frac{P_1(p)}{pP_2(p)}\right)\\ &= \prod_p\left(1+\frac{(p-1)P_1(p)-P_2(p)}{pP_2(p)}\right)\prod_{p\mid k}\frac{P_2(p)}{P_1(p)+P_2(p)}. \end{aligned}\tag{3.6}$$

Now $g_k(n)\neq 0$ only when n is of the form $n=n_1n_2n_3^2$, with pairwise coprime n_i $(i=1,2,3)$ satisfying $\mu^2(n_i)=1$, $n_1\mid k$ and $n_2n_3\nmid k$, in which case

$$g_k(n)=\mu(n_1)\mu^2(n_2)\prod_{p\mid n_2}\left(\frac{pP_1(p)-P_2(p)}{P_2(p)}\right)\mu(n_3)n_3\prod_{p\mid n_3}\frac{P_1(p)}{P_2(p)}.\tag{3.7}$$

Hence, for $t\geqslant 1$, we have

$$\begin{aligned}\sum_{n\leqslant t}|g_k(n)| &\leqslant \sum_{n_1\mid k}\mu^2(n_1)\sum_{n_2\leqslant t/n_1}\prod_{p\mid n_2}\left(\frac{pP_1(p)-P_2(p)}{P_2(p)}\right)\sum_{n_3\leqslant\sqrt{t/(n_1n_2)}}n_3\prod_{p\mid n_3}\frac{P_1(p)}{P_2(p)}\\ &\ll \sqrt{t}\sum_{n_1\mid k}\frac{\mu^2(n_1)}{\sqrt{n_1}}\sum_{n_2\leqslant t/n_1}\frac{1}{\sqrt{n_2}}\prod_{p\mid n_2}\frac{pP_1(p)-P_2(p)}{P_2(p)}\ll\sqrt{t}\,m(k),\end{aligned}\tag{3.8}$$

where we have made use of (2.6) and (2.9) for the sums over n_3 and n_2 respectively. Next, for $u\geqslant 1$, we have

$$\left|\sum_{n>u}\frac{g_k(n)}{n}\right| = \left|\int_u^{\infty}\frac{1}{t^2}\sum_{u<n\leqslant t}g_k(n)\,dt\right| \ll m(k)\int_u^{\infty}t^{-3/2}\,dt\ll\frac{m(k)}{\sqrt{u}},\tag{3.9}$$

and therefore

$$\sum_{n\leqslant x}\frac{g_k(n)}{n}=\sum_{n=1}^{\infty}\frac{g_k(n)}{n}+O\left(\frac{m(k)}{\sqrt{x}}\right).\tag{3.10}$$

For the main term observe that

$$\begin{aligned}\sum_{n\leqslant x}\frac{g_k(n)}{n}\log\frac{x}{n} &= \int_1^x\frac{1}{u}\sum_{n\leqslant u}\frac{g_k(n)}{n}\,du\\ &= \sum_{n=1}^{\infty}\frac{g_k(n)}{n}\int_1^x\frac{du}{u}-\int_1^{\infty}\frac{1}{u}\sum_{n>u}\frac{g_k(n)}{n}\,du+\int_x^{\infty}\frac{1}{u}\sum_{n>u}\frac{g_k(n)}{n}\,du\\ &= \sum_{n=1}^{\infty}\frac{g_k(n)}{n}\log x-\int_1^{\infty}\frac{1}{u}\sum_{n>u}\frac{g_k(n)}{n}\,du+O\left(\frac{m(k)}{\sqrt{x}}\right).\end{aligned}\tag{3.11}$$

The last integral here is

$$\begin{aligned}
&\int_1^\infty \frac{1}{u}\sum_{n>u}\frac{g_k(n)}{n}\,du\\
&\quad=\sum_{n=1}^\infty \frac{g_k(n)\log n}{n} = \sum_{n=1}^\infty \frac{g_k(n)}{n}\sum_{p^m||n}\log p^m\\
&\quad=\sum_{p,\,m\geqslant 1}\log p^m \sum_{\substack{n=1\\ p^m||n}}^\infty \frac{g_k(n)}{n} = \sum_{p,\,m\geqslant 1}\frac{g_k(p^m)\log p^m}{p^m}\sum_{\substack{n=1\\ p\nmid n}}^\infty \frac{g_k(n)}{n}\\
&\quad=\sum_{n=1}^\infty \frac{g_k(n)}{n}\sum_{p,\,m\geqslant 1}\frac{g_k(p^m)\log p^m}{p^m}\left(1+\sum_{l=1}^\infty \frac{g_k(p^l)}{p^l}\right)^{-1}\\
&\quad=\sum_{n=1}^\infty \frac{g_k(n)}{n}\left\{\sum_{p\nmid k}\frac{[(p-2)P_1(p)-P_2(p)]\log p}{(p-1)(P_1(p)+P_2(p))} - \sum_{p|k}\frac{\log p}{p-1}\right\}\\
&\quad=\sum_{n=1}^\infty \frac{g_k(n)}{n}\left\{\sum_{p}\frac{[(p-2)P_1(p)-P_2(p)]\log p}{(p-1)(P_1(p)+P_2(p))} - \sum_{p|k}\frac{P_1(p)\log p}{(P_1(p)+P_2(p))}\right\}. \qquad (3.12)
\end{aligned}$$

Plugging (3.12) in (3.11), and then using (3.11), (3.10), (3.8) and (3.6) in (3.3), we complete the proof of Lemma 1. □

Proof of Lemma 2. Let

$$A(s) := \sum_{n=1}^\infty \frac{\mu(n)\phi_2(n)}{\phi(n)}\frac{1}{n^s}, \qquad (3.13)$$

the series being absolutely convergent for $\Re s > 1$. From (2.13) with x half an odd integer, we have

$$\sum_{n\leqslant x}\frac{\mu(n)\phi_2(n)}{n\phi(n)} = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} A(1+w)\frac{x^w}{w}\,dw + O\left(\frac{x^c}{cT}\right) + O\left(\frac{\log x}{T}\right). \qquad (3.14)$$

Taking $c = 1/\log x$, which minimizes x^c/c, we find that the error of (3.14) is $O((\log x)/T)$. Next note that

$$A(s) = \prod_p \left(1 - \frac{p-2}{(p-1)p^s}\right) = \frac{1}{\zeta(s)}\prod_p\left(1+\frac{1}{(p-1)(p^s-1)}\right), \qquad (3.15)$$

where the last product is expressible as a Dirichlet series which is absolutely convergent for $\Re s > 0$. So we have

$$\begin{aligned}
&\sum_{n\leqslant x}\frac{\mu(n)\phi_2(n)}{n\phi(n)}\\
&\quad= \frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{1}{\zeta(w+1)}\prod_p\left(1+\frac{1}{(p-1)(p^{1+w}-1)}\right)\frac{x^w}{w}\,dw + O\left(\frac{\log x}{T}\right). \qquad (3.16)
\end{aligned}$$

Now we pull the line of integration to

$$w = -\frac{K}{\log T} + it \qquad (-T\leqslant t\leqslant T)$$

in accordance with the well-known zero-free region for $\zeta(s)$, so that the integrand has no poles in the region thus formed. Here

$$\frac{1}{\zeta(s)} = O(\log(|t|+2)) \qquad \left(\sigma \geqslant 1 - \frac{K}{\log(|t|+2)}\right) \tag{3.17}$$

holds (see Titchmarsh [18, Theorem 3.8 and equation (3.11.8)]), so that

$$\int_{-K/\log T - iT}^{-K/\log T + iT} \frac{1}{\zeta(1+w)} \prod_p \left(1 + \frac{1}{(p-1)(p^{w+1}-1)}\right) \frac{x^w}{w}\, dw \\ \ll x^{-K/\log T} \log T \int_{-T}^{T} \frac{dt}{\sqrt{(K/\log T)^2 + t^2}} \ll x^{-K/\log T} \log^2 T. \tag{3.18}$$

For the integrals over the horizontal sides of the contour we have

$$\int_{-K/\log T}^{1/\log x} \frac{1}{\zeta(1+u+iT)} \prod_p \left(1 + \frac{1}{(p-1)(p^{1+u+iT}-1)}\right) \frac{x^{u+iT}}{(u+iT)}\, du \\ \ll \frac{\log T}{T} \int_{-K/\log T}^{1/\log x} x^u \, du \ll \frac{\log T}{T}. \tag{3.19}$$

By taking $\log T = \sqrt{\log x}$, we deduce that all the error terms in (3.16), (3.18) and (3.19) are made to tend to 0 as $x \to \infty$. So, as $x \to \infty$, the sum of (2.18) tends to 0, and this completes the proof of Lemma 2. □

Proof of Lemma 3. Consider

$$A(s) := \sum_{n=1}^{\infty} \frac{\mu^2(n)}{n^s} \prod_{p|n} \frac{(3p-4)\sqrt{p}}{(p-1)(\sqrt{p}-1)} = \prod_p \left(1 + \frac{3}{p^s} \frac{(p-\frac{4}{3})\sqrt{p}}{(p-1)(\sqrt{p}-1)}\right) \\ = \zeta^3(s) \prod_p \left[1 + \frac{3}{p^{s+1/2}} \left(\frac{p - \sqrt{p}/3 - 1}{p - \sqrt{p} - 1 + 1/\sqrt{p}}\right) + \frac{3}{p^{2s}} \left(1 - \frac{(3p-4)\sqrt{p}}{(p-1)(\sqrt{p}-1)}\right) \right. \\ \left. + \frac{1}{p^{3s}} \left(\frac{(9p-12)\sqrt{p}}{(p-1)(\sqrt{p}-1)} - 1\right) - \frac{1}{p^{4s}} \frac{(3p-4)\sqrt{p}}{(p-1)(\sqrt{p}-1)}\right]. \tag{3.20}$$

The series $A(s)$ converges absolutely for $\Re s > 1$, and the last product, call it $P(s)$, is absolutely convergent for $\Re s > \frac{1}{2}$. Equation (2.13) can be applied with $\alpha = 3$, x half an odd integer, and $c = \frac{1}{2} + 1/\log x$ to have

$$\sum_{n \leqslant x} \mu^2(n) \prod_{p|n} \frac{(3p-4)}{(p-1)(\sqrt{p}-1)} \\ = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} A(\tfrac{1}{2}+w) \frac{x^w}{w}\, dw + O\left(\frac{x^{1/2}\log^3 x}{T}\right) + O\left(\frac{e^{K\log x/\log\log x} x^{1/2} \log x}{T}\right). \tag{3.21}$$

In writing the very last error term (with an appropriate constant K) we have used

$$\prod_{p|n} \frac{(3p-4)\sqrt{p}}{(p-1)(\sqrt{p}-1)} \ll 4^{\omega(n)} m(n), \tag{3.22}$$

and elementary deductions from the prime number theorem, namely (2.2) and

$$\omega(n) \leqslant c_1 \frac{\log n}{\log\log n}. \tag{3.23}$$

Now we pull the line of integration to $\Re w = \delta$ with $0 < \delta < \frac{1}{2}$. In doing so we pass the triple pole of the integrand at $w = \frac{1}{2}$, where the residue is

$$P(1)x^{1/2}\log^2 x + Dx^{1/2}\log x + Ex^{1/2} \tag{3.24}$$

(with the constants D and E made up of the Stieltjes constants and the values of $P(s)$ and its first two derivatives at $s = 1$). On the left vertical side of the contour we will have

$$\begin{aligned}\int_{\delta-iT}^{\delta+iT} \zeta^3(w+\tfrac{1}{2})P(w+\tfrac{1}{2})\frac{x^w}{w}\,dw &\ll x^\delta \int_0^T \frac{|\zeta(\frac{1}{2}+\delta+it)|^3}{\sqrt{\delta^2+t^2}}\,dt \\ &\ll x^\delta\left(\frac{1}{(\frac{1}{2}-\delta)^2} + \log^3 T\right), \end{aligned}\tag{3.25}$$

if we take $\frac{1}{2} - U/\log^{2/3} T < \delta < \frac{1}{2}$, U being an appropriate constant. To see this we employ the estimate (see Karatsuba and Voronin [**15**, p. 116])

$$\zeta(\sigma+it) = O(\log^{2/3}|t|) \qquad \left(\sigma \geqslant 1 - \frac{U}{\log^{2/3}|t|},\ |t| \geqslant 2\right), \tag{3.26}$$

which implies that

$$x^\delta \int_2^T \frac{|\zeta(\frac{1}{2}+\delta+it)|^3}{\sqrt{\delta^2+t^2}}\,dt \ll x^\delta \int_2^T \frac{\log^2 t}{t}\,dt \ll x^\delta \log^3 T. \tag{3.27}$$

We also have

$$x^\delta \int_0^2 \frac{|\zeta(\frac{1}{2}+\delta+it)|^3}{\sqrt{\delta^2+t^2}}\,dt \ll x^\delta \int_0^2 \frac{1}{|\delta-\frac{1}{2}+it|^3}\,dt \ll \frac{x^\delta}{(\frac{1}{2}-\delta)^2}. \tag{3.28}$$

For the horizontal sides of the contour we have, by (3.26),

$$\int_{\delta+iT}^{c+iT} \zeta^3(w+\tfrac{1}{2})P(w+\tfrac{1}{2})\frac{x^w}{w}\,dw \ll \frac{1}{T}\int_\delta^c |\zeta(\tfrac{1}{2}+\sigma+iT)|^3 x^\sigma\,d\sigma \ll \frac{x^{1/2}\log^2 T}{T\log x}. \tag{3.29}$$

Choosing

$$\delta = \frac{1}{2} - \frac{1}{(\log T)^{3/4}}, \qquad T = e^{2K\log x/\log\log x}\log x\log\log x, \tag{3.30}$$

we make all the error terms in (3.21), (3.25) and (3.29) to be $o(\sqrt{x})$. This completes the proof of Lemma 3. □

Proof of Lemma 4. We have

$$\sum_{\substack{n=1\\(n,k)=1}}^{\infty} \frac{\mu(n)\mu\cdot\phi((n,j))}{\phi^2(n)} = \prod_{\substack{p|j\\p\nmid k}}\left(1+\frac{1}{p-1}\right)\prod_{\substack{p\nmid j\\p\nmid k}}\left(1-\frac{1}{(p-1)^2}\right), \tag{3.31}$$

where we notice that the last product is 0 if $2 \nmid jk$. If $2 \mid jk$, the products are re-organized to give the main term of (2.20), which has been expressed so as to be valid whether or not $2 \mid jk$. The O-term of (2.20) is the tail of the series

$$\begin{aligned}\sum_{\substack{n>x\\(n,k)=1}} \frac{\mu(n)\mu\cdot\phi((n,j))}{\phi^2(n)} &= \sum_{\substack{d|j^*\\(d,k)=1}} \mu\cdot\phi(d) \sum_{\substack{n>x\\(n,k)=1\\(n,j)=d}} \frac{\mu(n)}{\phi^2(n)} \\ &= \sum_{\substack{d|j^*\\(d,k)=1}} \frac{\mu^2(d)}{\phi(d)} \sum_{\substack{t>x/d\\(t,kj)=1}} \frac{\mu(t)}{\phi^2(t)} \ll \frac{1}{x}\sum_{\substack{d|j^*\\(d,k)=1}} \frac{d}{\phi(d)} \ll \frac{d(j')j'}{x\phi(j')}. \end{aligned}\tag{3.32}$$

To obtain (2.21), let

$$B(s) := \sum_{n=1}^{\infty} \frac{\mu(n)\mu\cdot\phi((n,j))}{\phi^2(n)n^s} = \prod_{p|j}\left(1+\frac{1}{(p-1)p^s}\right)\prod_{p\nmid j}\left(1-\frac{1}{(p-1)^2p^s}\right), \tag{3.33}$$

where for $\Re s > -1$ the product is absolutely convergent. Then

$$B'(0) = \begin{cases} \mathfrak{S}_2(j)\left[\sum_{p\nmid j}\frac{\log p}{p(p-2)} - \sum_{p|j}\frac{\log p}{p}\right] & \text{if } 2 \mid j, \\ \mathfrak{S}_2(2j)\frac{\log 2}{2} & \text{if } 2\nmid j\ . \end{cases} \tag{3.34}$$

However,

$$\sum_{n\leqslant x}\frac{\mu(n)\mu\cdot\phi((n,j))}{\phi^2(n)}(-\log n) = B'(0) + \sum_{n>x}\frac{\mu(n)\mu\cdot\phi((n,j))}{\phi^2(n)}(\log n),$$

and the very last sum is shown to be

$$\ll \frac{d(j^*)j^*\log 2x}{\phi(j^*)x},$$

similarly to (3.32). □

Proof of Lemma 5. We extend the sum to all n, introducing as error the tail of the series for $n > x$. We have

$$\begin{aligned}\sum_{n=1}^{\infty}&\frac{\mu(n)d(n)}{\phi(n)\phi_2\left(\frac{n}{(n,2)}\right)}\frac{\mu}{d}((n,J))\frac{\mu}{\phi}((n,k))\phi_2\left(\left(\frac{n}{(n,2)},J\right)\right)\\ &= (1+\mu((2,k)))\prod_{p\nmid J}\left(1-\frac{2}{(p-1)(p-2)}\right)\prod_{\substack{p>2\\ p|J\\ p\nmid k}}\left(1+\frac{1}{p-1}\right)\prod_{\substack{p>2\\ p|k}}\left(1-\frac{1}{(p-1)^2}\right),\end{aligned} \tag{3.35}$$

and the introduced error is bounded as

$$\begin{aligned}&\ll \sum_{n>x}\frac{\mu^2(n)d(n)\phi_2\left(\left(\frac{n}{(n,2)},J\right)\right)}{\phi(n)\phi_2\left(\frac{n}{(n,2)}\right)d((n,J))} \ll \sum_{m|J}\frac{\phi_2}{d}(m)\sum_{\substack{n>x\\ (n,J/2)=m}}\frac{\mu^2(n)d(n)}{\phi(n)\phi_2\left(\frac{n}{(n,2)}\right)}\\ &\ll \sum_{m|J}\frac{1}{\phi(m)}\sum_{\substack{t\geqslant x/m\\ (t,2J)=1}}\frac{\mu^2(t)d(t)}{\phi(t)\phi_2(t)} \ll \sum_{m|J}\frac{1}{\phi(m)}\sum_{t\geqslant x/m}\frac{1}{t^{2-\epsilon}}\\ &\ll \frac{1}{x^{1-\epsilon}}\sum_{m|J}\frac{m^{1-\epsilon}}{\phi(m)} \ll \frac{1}{x^{1-\epsilon}},\end{aligned} \tag{3.36}$$

which proves Lemma 5. □

4. *Pair correlations of $\lambda_R(n)$*

In the case $k=1$ of Theorem 1 we have

$$\begin{aligned}\mathcal{S}_1(N,(j_1),(1)) &= \sum_{n\leqslant N}\lambda_R(n+j_1) = \sum_{r\leqslant R}\frac{\mu^2(r)}{\phi(r)}\sum_{d|r}d\mu(d)\sum_{\substack{\max(1,1+j_1)\leqslant n\leqslant N+j_1\\ d|n}}1\\ &= \sum_{r\leqslant R}\frac{\mu^2(r)}{\phi(r)}\sum_{d|r}d\mu(d)\left(\frac{N+j_1-\max(0,j_1)}{d}+O(1)\right)\\ &= N+\min(0,j_1)+O\Bigg(\sum_{r\leqslant R}\frac{\mu^2(r)\sigma(r)}{\phi(r)}\Bigg)\\ &= N(1+o(1))+O(R),\end{aligned}\tag{4.1}$$

where we refer to (2.6) for

$$\sum_{r\leqslant R}\frac{\mu^2(r)\sigma(r)}{\phi(r)}\ll R.\tag{4.2}$$

To examine the case $k=2$ of Theorem 1 we need to consider

$$\mathcal{S}_2(j)=\sum_{n=1}^{N}\lambda_R(n)\lambda_R(n+j).\tag{4.3}$$

In our earlier notation, $\mathcal{S}_2(j)=\mathcal{S}_2(N,(0,j),(1,1))$ if $j\neq 0$, and $\mathcal{S}_2(0)=\mathcal{S}_2(N,(0),(2))$. We have for any j,

$$\mathcal{S}_2(j)=\sum_{r_1,r_2\leqslant R}\frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)}\sum_{\substack{d|r_1\\ e|r_2}}d\mu(d)e\mu(e)\sum_{\substack{n\leqslant N\\ d|n\\ e|n+j}}1.\tag{4.4}$$

The innermost sum is over the values of n in a unique residue class modulo $[d,e]$ whenever $(d,e)\mid j$, in which case its value is $N/[d,e]+O(1)$, otherwise the innermost sum is void. By (4.2) the last $O(1)$ leads to a contribution of $O(R^2)$ in (4.4). Hence

$$\mathcal{S}_2(j)=N\sum_{r_1,r_2\leqslant R}\frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)}\sum_{\substack{d|r_1\\ e|r_2\\ (d,e)|j}}\mu(d)\mu(e)(d,e)+O(R^2).\tag{4.5}$$

Let $(d,e)=\delta$, $d=d'\delta$, and $e=e'\delta$, so that $(d',e')=1$. Then the inner sums over d and e become

$$\sum_{\substack{\delta|(r_1,r_2)\\ \delta|j}}\delta\sum_{d'|(r_1/\delta)}\mu(d')\sum_{\substack{e'|(r_2/\delta)\\ (e',d')=1}}\mu(e').$$

Here the innermost sum is

$$\sum_{\substack{e'|(r_2/\delta)\\ (e',d')=1}}\mu(e')=\prod_{\substack{p|(r_2/\delta)\\ p\nmid d'}}(1+\mu(p))=\begin{cases}1 & \text{if } (r_2/\delta)\mid d',\\ 0 & \text{otherwise.}\end{cases}$$

Next the sum over d' becomes

$$\sum_{\substack{d'|(r_1/\delta)\\ (r_2/\delta)|d'}}\mu(d')=\begin{cases}\mu(r_1/\delta) & \text{if } r_1=r_2,\\ 0 & \text{otherwise.}\end{cases}$$

Hence we have, for any j,

$$\begin{aligned}\mathcal{S}_2(j) &= N\sum_{r_1\leqslant R}\frac{\mu(r_1)}{\phi^2(r_1)}\sum_{\delta|(r_1,j)}\delta\mu(\delta)+O(R^2)\\ &= N\sum_{r_1\leqslant R}\frac{\mu(r_1)\mu((j,r_1))\phi((j,r_1))}{\phi^2(r_1)}+O(R^2). \end{aligned}\tag{4.6}$$

Now, if $j=0$, then (4.6) reduces to

$$\sum_{n\leqslant N}(\lambda_R(n))^2 = N\mathcal{L}_1(R)+O(R^2),\tag{4.7}$$

where $\mathcal{L}_1(R)$ was defined in (1.39), and by (2.15) this proves Theorem 1 for the case $\mathcal{S}_2(N,(0),(2))$. If $j\neq 0$ then, by Lemma 4,

$$\sum_{r_1\leqslant R}\frac{\mu(r_1)\mu((j,r_1))\phi((j,r_1))}{\phi^2(r_1)} = \mathfrak{S}_2(j)+O\left(\frac{j^*d(j^*)}{R\phi(j^*)}\right),\tag{4.8}$$

and therefore

$$\sum_{n\leqslant N}\lambda_R(n)\lambda_R(n+j) = N\mathfrak{S}_2(j)+O\left(\frac{Nj^*d(j^*)}{R\phi(j^*)}\right)+O(R^2)\qquad (j\neq 0),\tag{4.9}$$

which completes the proof of Theorem 1 for the case $\mathcal{S}_2(N,(0,j),(1,1))$.

5. *The mixed correlations*

We now turn our attention to the mixed correlations $\tilde{\mathcal{S}}_k(N,\boldsymbol{j},\boldsymbol{a})$ defined in (1.5). The $k=1$ case was noted in (1.6), and the first mixed moment was treated in (1.35)–(1.37). We consider

$$\tilde{\mathcal{S}}_2(j) = \sum_{n=1}^{N}\lambda_R(n+j)\Lambda(n)\qquad (j\neq 0)\tag{5.1}$$

in the case of mixed second level correlations. In the notation of (1.5),

$$\tilde{\mathcal{S}}_2(j) = \tilde{\mathcal{S}}_2(N,(j,0),(1,1)).$$

We have

$$\tilde{\mathcal{S}}_2(j) = \sum_{r\leqslant R}\frac{\mu^2(r)}{\phi(r)}\sum_{d|r}d\mu(d)\sum_{\substack{\max(2,1-j)\leqslant n\leqslant N\\ d|n+j}}\Lambda(n)\qquad (j\neq 0).\tag{5.2}$$

The innermost sum of (5.2) is equal to

$$\begin{aligned}&\psi(N;d,-j)-\psi(\max(2,1-j);d,-j)\\ &\qquad = [(d,j)=1]\frac{N}{\phi(d)}+E(N;d,-j)+[j<0]\cdot O\left(\left(1+\frac{|j|}{d}\right)\log|j|\right),\end{aligned}\tag{5.3}$$

by (1.41). Hence (5.2) becomes

$$\begin{aligned}\tilde{\mathcal{S}}_2(j) = {}& N\sum_{r\leqslant R}\frac{\mu^2(r)}{\phi(r)}\sum_{\substack{d|r\\(d,j)=1}}\frac{d\mu(d)}{\phi(d)}+O\left(\sum_{d\leqslant R}\frac{d\mu^2(d)}{\phi(d)}|E(N;d,-j)|\sum_{\substack{s\leqslant R/d\\(s,d)=1}}\frac{\mu^2(s)}{\phi(s)}\right)\\ &+[j<0]\cdot O\left(\log|j|\sum_{r\leqslant R}\frac{\mu^2(r)}{\phi(r)}\sum_{d|r}d\left(1+\frac{|j|}{d}\right)\right).\end{aligned}\tag{5.4}$$

Since

$$\sum_{\substack{d|r \\ (d,j)=1}} \frac{d\mu(d)}{\phi(d)} = \mu\left(\frac{r}{(r,j)}\right) / \phi\left(\frac{r}{(r,j)}\right), \tag{5.5}$$

the main term of (5.4) is the same as that of (4.6), so (4.8) settles it. The last error term is easily bounded as $O((R + |j|\log^2 R)\log|j|)$. As for the error term with E in (5.4), it is

$$\ll \log R \log\log R \left(\sum_{\substack{d\leqslant R \\ (d,j)=1}} |E(N; d, -j)| + \sum_{\substack{d\leqslant R \\ (d,j)>1}} \mu^2(d)\psi(N; d, -j) \right).$$

Here the first sum is estimated by the Bombieri–Vinogradov theorem (1.42), provided that $R \ll N^{1/2-\epsilon}$. In the second sum only values of n which are powers of those primes that are divisors of (d,j) contribute to ψ. Hence the second sum is

$$\ll \sum_{p|j} \sum_{\substack{d\leqslant R \\ p|d}} \sum_{\substack{p^a\leqslant N \\ d|p^a+j}} \log p \ll \log N \sum_{p|j} \sum_{\substack{d\leqslant R \\ p|d}} 1 \ll R \log N \log\log 3|j|.$$

Thus we obtain, for $j \neq 0$,

$$\tilde{\mathcal{S}}_2(j) = \sum_{n=1}^{N} \lambda_R(n+j)\Lambda(n) = N\mathfrak{S}_2(j) + O\left(\frac{Nj^* d(j^*)}{R\phi(j^*)}\right) + O\left(\frac{N}{\log^{\mathcal{A}} N}\right), \tag{5.6}$$

completing the proof of Theorem 2 in the case $k = 2$.

For mixed correlations of the third level we may begin with

$$\tilde{\mathcal{S}}_3(j_1, j_2, 0) = \sum_{n=1}^{N} \lambda_R(n+j_1)\lambda_R(n+j_2)\Lambda(n) \qquad (j_1 j_2 \neq 0). \tag{5.7}$$

In the beginning the two cases, $j_1 = j_2$ or not, will undergo a common treatment. Using the definition of $\lambda_R(n)$ we can rewrite this as

$$\tilde{\mathcal{S}}_3(j_1, j_2, 0) = \sum_{r_1, r_2 \leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{d|r_1 \\ e|r_2}} d\mu(d) e\mu(e) \sum_{\substack{\max(2, 1-j_1, 1-j_2)\leqslant n\leqslant N \\ n\equiv -j_1 (\mathrm{mod}\, d) \\ n\equiv -j_2 (\mathrm{mod}\, e)}} \Lambda(n). \tag{5.8}$$

The congruence conditions in the innermost sum are compatible if and only if $(d,e) \mid j_1 - j_2$, in which case there is a unique residue class j such that $n \equiv j (\mathrm{mod}\, [d,e])$, and similar to (5.3) the innermost sum is equal to

$$[([d,e], j) = 1] \frac{N}{\phi([d,e])} + E(N; [d,e], j) + O\left(\max_i |j_i| \log N\right). \tag{5.9}$$

Placed in (5.8), the error part of (5.9) contributes (with the last O-term called M)

$$\begin{aligned}
&\ll \sum_{\substack{d,e\leqslant R\\(d,e)|j_1-j_2}} \frac{de}{\phi(d)\phi(e)}\big[|E(N;[d,e],j)|+O(M)\big] \sum_{\substack{s_1\leqslant R/d\\(s_1,d)=1}} \frac{\mu^2(s_1)}{\phi(s_1)} \sum_{\substack{s_2\leqslant R/e\\(s_2,d)=1}} \frac{\mu^2(s_2)}{\phi(s_2)}\\
&\ll (\log R)^2(\log\log R)^2 \sum_{\substack{d,e\leqslant R\\(d,e)|j_1-j_2}} \big[|E(N;[d,e],j)|+O(M)\big]\\
&\ll (\log R)^3 \sum_{D\leqslant R^2}\Big[\max_{a(\bmod D)}|E(N;D,a)|+O(M)\Big]\sum_{\substack{d,e\leqslant R\\ [d,e]=D}} 1\\
&\ll (\log R)^3 \sum_{D\leqslant R^2}\Big[\max_{a(\bmod D)}|E(N;D,a)|+O(M)\Big] d_3(D)\\
&\ll (\log R)^3 \sqrt{\sum_{D\leqslant R^2}\frac{d_3(D)^2}{D}}\sqrt{\sum_{D\leqslant R^2} D\Big[\max_{a(\bmod D)}|E(N;D,a)|+O(M)\Big]^2}\\
&\ll (\log R)^{15/2}\sqrt{\sum_{D\leqslant R^2} D\Big[\frac{N\log N}{D}\max_{a(\bmod D)}|E(N;D,a)|+O(M^2)\Big]}\\
&\ll (\log R)^{15/2}\sqrt{\frac{N^2}{\log^{\mathcal{A}-1}N}+NR^2\log^2 N+\Big(\max_i|j_i|\Big)^2R^4\log^2 N}\\
&\ll \frac{N}{\log^{\mathcal{B}} N},
\end{aligned} \tag{5.10}$$

where $\mathcal{B}$ can be as large as wished by taking $\mathcal{A}$ large enough, provided that $R \ll N^{1/4-\epsilon}$ for the applicability of the Bombieri–Vinogradov estimate (1.42), and that $R^2 \max_i |j_i| \ll N^{1-\epsilon}$. In the above sequence of inequalities we have used some well-known features of the divisor function $d_3(n)$, and we have employed the trivial estimate

$$|E(N;D,a)| \ll \frac{N\log N}{D},$$

and the $(a,D)>1$ part of $\max_{a(\bmod D)}$ is estimated to be $\ll NR^2\log^2 N$ as was done for (5.6).

From (5.8)–(5.10) we have

$$\tilde{\mathcal{S}}_3(j_1,j_2,0) = N\sum_{r_1,r_2\leqslant R}\frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)}\sum_{\substack{d|r_1,e|r_2\\(d,j_1)=1,(e,j_2)=1\\(d,e)|j_1-j_2}}\frac{d\mu(d)e\mu(e)}{\phi([d,e])}+O\left(\frac{N}{\log^{\mathcal{B}}N}\right), \tag{5.11}$$

for $(j,[d,e])=1$ if and only if $(j_1,d)=1$ and $(j_2,e)=1$. Letting

$$r_1' = \frac{r_1}{(r_1,j_1)},\qquad r_2' = \frac{r_2}{(r_2,j_2)},\qquad (d,e)=\delta,\qquad d=d'\delta,\qquad e=e'\delta, \tag{5.12}$$

we see that the inner sums over d and e in (5.11) take the form

$$\sum_{\substack{\delta|(r_1',r_2')\\ \delta|j_1-j_2}}\frac{\delta^2}{\phi(\delta)}\sum_{d'|(r_1'/\delta)}\frac{d'\mu(d')}{\phi(d')}\sum_{\substack{e'|(r_2'/\delta)\\(e',d')=1}}\frac{e'\mu(e')}{\phi(e')}.$$

Here the innermost sum is

$$\sum_{\substack{e'|(r_2'/\delta)\\(e',d')=1}} \frac{e'\mu(e')}{\phi(e')} = \prod_{\substack{p|(r_2'/\delta)\\p\nmid d'}} \left(1-\frac{p}{p-1}\right) = \prod_{\substack{p|(r_2'/\delta)\\p\nmid d'}} \frac{-1}{p-1} = \frac{\mu\left(\frac{r_2'/\delta}{(r_2'/\delta,d')}\right)}{\phi\left(\frac{r_2'/\delta}{(r_2'/\delta,d')}\right)}.$$

So the inner sums over d and e in (5.11) become

$$\frac{\mu(r_2')}{\phi(r_2')}\sum_{\substack{\delta|(r_1',r_2')\\\delta|j_1-j_2}}\delta^2\mu(\delta)\sum_{d'|(r_1'/\delta)}\frac{d'\mu(d')}{\phi(d')}\mu\left(\left(\frac{r_2'}{\delta},d'\right)\right)\phi\left(\left(\frac{r_2'}{\delta},d'\right)\right),$$

in which we can evaluate the sum over d' as

$$\prod_{p|(r_1'/\delta,r_2'/\delta)}(1+p)\prod_{\substack{p|(r_1'/\delta)\\p\nmid(r_2'/\delta)}}\frac{-1}{p-1} = \sigma\left(\left(\frac{r_1'}{\delta},\frac{r_2'}{\delta}\right)\right)\frac{\mu\left(\frac{r_1'/\delta}{(r_1'/\delta,r_2'/\delta)}\right)}{\phi\left(\frac{r_1'/\delta}{(r_1'/\delta,r_2'/\delta)}\right)}.$$

Now the inner sums over d and e in (5.11) have been simplified to

$$\begin{aligned}&\frac{\mu(r_1')}{\phi(r_1')}\frac{\mu(r_2')}{\phi(r_2')}\mu\cdot\phi\cdot\sigma((r_1',r_2'))\sum_{\substack{\delta|(r_1',r_2')\\\delta|j_1-j_2}}\frac{\delta^2\mu(\delta)}{\sigma(\delta)}\\&\qquad=\frac{\mu(r_1')}{\phi(r_1')}\frac{\mu(r_2')}{\phi(r_2')}\mu\cdot\phi\cdot\sigma((r_1',r_2'))\prod_{p|(r_1',r_2',j_1-j_2)}\left(1-\frac{p^2}{p+1}\right).\end{aligned}$$

Plugging this into (5.11) we can now express the main term of $\tilde{\mathcal{S}}_3(j_1,j_2,0)$ as

$$N\sum_{r_1,r_2\leqslant R}\frac{\mu^2(r_1)\mu^2(r_2)\mu(r_1')\mu(r_2')}{\phi(r_1)\phi(r_2)\phi(r_1')\phi(r_2')}\mu\cdot\phi\cdot\sigma((r_1',r_2'))\prod_{p|(r_1',r_2',j_1-j_2)}\left(\frac{1+p-p^2}{p+1}\right). \tag{5.13}$$

We now express r_1 and r_2 as products of coprime factors and transform (5.13) into a sum over these new variables. Let

$$r_1=(r_1,j_1)r_1'=(r_1,j_1)(r_1',j_2)r_1''=(r_1,j_1)(r_1',j_2)(r_1'',j_1-j_2)r_1'''=s_{11}s_{12}s_{13}r_1''', \tag{5.14}$$

say, so that

$$s_{11}\mid j_1;\qquad s_{12}\mid j_2,\quad (s_{12},j_1)=1;\qquad s_{13}\mid(j_1-j_2),\quad (s_{13},j_1j_2)=1;$$

and $(r_1''',j_1j_2(j_1-j_2))=1$. Similarly let

$$r_2=(r_2,j_2)r_2'=(r_2,j_2)(r_2',j_1)r_2''=(r_2,j_2)(r_2',j_1)(r_2'',j_1-j_2)r_2'''=s_{22}s_{21}s_{23}r_2''', \tag{5.15}$$

with

$$s_{22}\mid j_2;\qquad s_{21}\mid j_1,\quad (s_{21},j_2)=1;\qquad s_{23}\mid(j_1-j_2),\quad (s_{23},j_1j_2)=1;$$

and $(r_2''',j_1j_2(j_1-j_2))=1$. Then we let

$$(s_{13},s_{23})=s_3,\quad s_{13}=t_1s_3,\quad s_{23}=t_2s_3,\quad (r_1''',r_2''')=r,\quad r_1'''=s_{14}r,\quad r_2'''=s_{24}r, \tag{5.16}$$

so that t_1, t_2 and s_3 are each divisors of j_1-j_2 which are coprime to j_1j_2, and s_{14}, s_{24} and r are relatively prime to $j_1j_2(j_1-j_2)$. We now have the factorizations

$$\begin{gathered}r_1=s_{11}s_{12}t_1s_3s_{14}r,\qquad r_2=s_{22}s_{21}t_2s_3s_{24}r,\\ r_1'=\frac{r_1}{s_{11}},\qquad r_2'=\frac{r_2}{s_{22}},\qquad (r_1',r_2')=s_3r,\qquad (r_1',r_2',j_1-j_2)=s_3,\end{gathered} \tag{5.17}$$

with the just-stated coprimality and divisibility conditions on these variables to be specified by a star on the summation signs below. Now the sum of (5.11) has been transformed into

$$\sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3s_{14}r\leqslant R\\ s_{22}s_{21}t_2s_3s_{24}r\leqslant R}} \frac{\mu^2(s_{11})\mu(s_{12})\mu(t_1)\mu(t_2)\mu^2(s_{22})\mu(s_{21})}{\phi(s_{11})\phi^2(s_{12})\phi^2(t_1)\phi^2(t_2)\phi(s_{22})\phi^2(s_{21})}\frac{\mu^2(s_3)}{\phi^3(s_3)}\prod_{p\mid s_3}(p^2-p-1) \times \frac{\mu(s_{14})\mu(s_{24})\mu(r)\sigma(r)}{\phi^2(s_{14})\phi^2(s_{24})\phi^3(r)}. \tag{5.18}$$

If $j_1=j_2=j\neq 0$, we have $s_{12}=s_{21}=r_1'''=r_2'''=1$, so that (5.18) reduces to

$$\sideset{}{^*}\sum_{\substack{s_{11}t_1s_3\leqslant R\\ s_{22}t_2s_3\leqslant R}} \frac{\mu^2(s_{11})\mu^2(s_{22})\mu(t_1)\mu(t_2)}{\phi(s_{11})\phi(s_{22})\phi^2(t_1)\phi^2(t_2)}\mu^2(s_3)\prod_{p\mid s_3}\frac{(p^2-p-1)}{(p-1)^3}. \tag{5.19}$$

We first deal with (5.19) by starting to sum over s_3 as

$$\begin{aligned}\sum_{\substack{s_3\leqslant \min(R/(s_{11}t_1),R/(s_{22}t_2))\\ (s_3,t_1t_2j)=1}} &\mu^2(s_3)\prod_{p\mid s_3}\frac{(p^2-p-1)}{(p-1)^3}\\ &=\prod_p\left(1+\frac{p-2}{p(p-1)^2}\right)\prod_{p\mid t_1t_2j}\left(\frac{(p-1)^3}{p^3-2p^2+2p-2}\right)\\ &\quad\times\left\{\log\min\left(\frac{R}{s_{11}t_1},\frac{R}{s_{22}t_2}\right)+\gamma+\sum_p\frac{(2p-3)\log p}{(p-1)(p^3-2p^2+2p-2)}\right.\\ &\qquad\left.+\sum_{p\mid t_1t_2j}\frac{(p^2-p-1)\log p}{(p^3-2p^2+2p-2)}\right\}\\ &\quad+O\left(\frac{m(t_1t_2j)}{\sqrt{\min(R/(s_{11}t_1),R/(s_{22}t_2))}}\right)\end{aligned} \tag{5.20}$$

according to (2.14) and (2.17). Trivially $\max(s_{11}t_1,s_{22}t_2)\leqslant s_{11}t_1s_{22}t_2$, so that the contribution of the last error term to (5.19) can be estimated as

$$\begin{aligned}&\ll\frac{m(j)}{\sqrt{R}}\sum_{\substack{s_{11}\mid j\\ s_{22}\mid j}}\frac{\mu^2(s_{11})\sqrt{s_{11}}\mu^2(s_{22})\sqrt{s_{22}}}{\phi(s_{11})\phi(s_{22})}\sum_{\substack{t_1\leqslant R/s_{11}\\ (t_1,j)=1}}\frac{\mu^2(t_1)m(t_1)\sqrt{t_1}}{\phi^2(t_1)}\sum_{\substack{t_2\leqslant R/s_{22}\\ (t_2,t_1j)=1}}\frac{\mu^2(t_2)m(t_2)\sqrt{t_2}}{\phi^2(t_2)}\\ &\ll R^{-1/2+\epsilon}\end{aligned} \tag{5.21}$$

in view of (2.2).

Before calculating the main contribution from the log min term, we find the contributions from the other terms of (5.20). Since

$$\gamma+\sum_p\frac{(2p-3)\log p}{(p-1)(p^3-2p^2+2p-2)}+\sum_{p\mid t_1t_2j}\frac{(p^2-p-1)\log p}{(p^3-2p^2+2p-2)}\ll\log\log 3t_1t_2j^* \tag{5.22}$$

as in (2.1), the sums over t_1 and t_2 are each $O(1)$, and the sums over s_{11} and s_{22} running through the divisors of j are each of value $j/\phi(j)$, and by (2.3),

$$\left(\frac{j}{\phi(j)}\right)^2\prod_{p\mid j}\frac{(p-1)^3}{p^3-2p^2+2p-2}\ll\prod_{p\mid j}\left(1+\frac{1}{p}\right)\ll\log\log 3j^*, \tag{5.23}$$

the secondary terms of (5.20) contribute to (5.19)

$$\ll(\log\log 3j^*)^2. \tag{5.24}$$

As for the main term of the brackets of (5.20), $\log R - \log\max(s_{11}t_1, s_{22}t_2)$, the log max part will contribute little. This follows because the t_1 and t_2 sums are again $O(1)$, and the s_{11} and s_{22} sums are

$$\ll \sum_{\substack{s_{11}|j \\ s_{22}|j}} \frac{\mu^2(s_{11})\log s_{11}}{\phi(s_{11})}\frac{\mu^2(s_{22})}{\phi(s_{22})} \ll \frac{j}{\phi(j)}\sum_{s_{11}|j}\frac{\mu^2(s_{11})\log s_{11}}{\phi(s_{11})} = \left(\frac{j}{\phi(j)}\right)^2 \sum_{p|j}\frac{\log p}{p}, \tag{5.25}$$

where we have employed (2.4). Using (5.23), and (2.1) to bound the very last sum, we see that the contribution from the log max term is also majorized as in (5.24). We note that it is possible to carry out the calculation (5.22)–(5.24) more precisely, but this would not be significant since we have not been able to give a better evaluation of the contribution of the log max term.

The main term of (5.19) has now been reduced to

$$\log R \prod_p \left(1+\frac{p-2}{p(p-1)^2}\right)\prod_{p|j}\frac{(p-1)^3}{p^3-2p^2+2p-2}\sum_{\substack{s_{11}|j \\ s_{22}|j}}\frac{\mu^2(s_{11})\mu^2(s_{22})}{\phi(s_{11})\phi(s_{22})}$$
$$\times \sum_{\substack{t_1\leqslant R/s_{11} \\ (t_1,j)=1}}\mu(t_1)\prod_{p|t_1}\frac{p-1}{p^3-2p^2+2p-2}\sum_{\substack{t_2\leqslant R/s_{22} \\ (t_2,t_1j)=1}}\mu(t_2)\prod_{p|t_2}\frac{p-1}{p^3-2p^2+2p-2}. \tag{5.26}$$

Now observe that, by (2.9),

$$\sum_{\substack{n\leqslant x \\ (n,k)=1}}\mu(n)\prod_{p|n}\frac{p-1}{p^3-2p^2+2p-2}$$
$$= \sum_{\substack{n=1 \\ (n,k)=1}}^{\infty}\mu(n)\prod_{p|n}\left(\frac{p-1}{p^3-2p^2+2p-2}\right)+O\left(\frac{1}{x}\right)$$
$$= \prod_{p\nmid k}\left(1-\frac{p-1}{p^3-2p^2+2p-2}\right)+O\left(\frac{1}{x}\right)$$
$$= \prod_{p}\left(1-\frac{p-1}{p^3-2p^2+2p-2}\right)\prod_{p|k}\left(\frac{p^3-2p^2+2p-2}{p^3-2p^2+p-1}\right)+O\left(\frac{1}{x}\right). \tag{5.27}$$

Upon calculating the t_2-sum according to (5.27), we see that the error of the last kind contributes to (5.26)

$$\ll \frac{\log R}{R}\sum_{\substack{s_{11}|j \\ s_{22}|j}}\frac{\mu^2(s_{11})\mu^2(s_{22})s_{22}}{\phi(s_{11})\phi(s_{22})}\sum_{\substack{t_1\leqslant R/s_{11} \\ (t_1,j)=1}}\mu^2(t_1)\prod_{p|t_1}\frac{p-1}{p^3-2p^2+2p-2}$$
$$\ll \frac{\log R}{R}\sum_{s_{11}|j}\frac{\mu^2(s_{11})}{\phi(s_{11})}\sum_{s_{22}|j}\frac{\mu^2(s_{22})s_{22}}{\phi(s_{22})} \ll \frac{\log R}{R}\frac{j^*}{\phi(j^*)}\frac{j^*d(j^*)}{\phi(j^*)} \ll R^{-1+\epsilon}. \tag{5.28}$$

From (5.27) we see that the main term is now

$$\log R\prod_p\left(1-\frac{1}{p(p-1)^2}\right)\prod_{p|j}\frac{(p-1)^3}{p^3-2p^2+p-1}$$
$$\times\sum_{\substack{s_{11}|j \\ s_{22}|j}}\frac{\mu^2(s_{11})\mu^2(s_{22})}{\phi(s_{11})\phi(s_{22})}\sum_{\substack{t_1\leqslant R/s_{11} \\ (t_1,j)=1}}\mu(t_1)\prod_{p|t_1}\frac{p-1}{p^3-2p^2+p-1}. \tag{5.29}$$

As in (5.27) we have

$$\sum_{\substack{t_1 \leqslant R/s_{11} \\ (t_1,j)=1}} \mu(t_1) \prod_{p|t_1} \frac{p-1}{p^3-2p^2+p-1} = \prod_{p\nmid j}\left(1-\frac{p-1}{p^3-2p^2+p-1}\right) + O\left(\frac{s_{11}}{R}\right), \tag{5.30}$$

and this last error leads to a contribution of $O(R^{-1+\epsilon})$ just as in (5.28). The main term becomes

$$\log R \prod_p \left(1-\frac{1}{p(p-1)^2}\right) \prod_{p|j} \frac{(p-1)^3}{p^3-2p^2+p-1} \\ \times \prod_{p\nmid j}\left(1-\frac{p-1}{p^3-2p^2+p-1}\right) \sum_{\substack{s_{11}|j \\ s_{22}|j}} \frac{\mu^2(s_{11})\mu^2(s_{22})}{\phi(s_{11})\phi(s_{22})}. \tag{5.31}$$

Note that if $2 \nmid j$, then the third product is 0. We simplify (5.31) and obtain the final expression for the main term as

$$[2 \mid j]\, 2 \prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right) \prod_{\substack{p|j \\ p>2}} \left(\frac{p-1}{p-2}\right) \log R = \mathfrak{S}_2(j) \log R. \tag{5.32}$$

Since the largest error term all through this calculation (besides that of (5.11)) was $O(N(\log\log 3j^*)^2)$, this completes the proof of Theorem 2 in the case $k=3$, $r=2$ in view of (1.9).

We now calculate the expression (5.18), with $j_1 \neq j_2$ and $j_1 j_2 \neq 0$, for the case $k=r=3$ of Theorem 2. Some additional notation will prove to be convenient. Let

$$J = [j_1 j_2 (j_1-j_2)]^*, \qquad j_1^* = (j_1^*)'(j_1^*, j_2^*), \qquad j_2^* = (j_2^*)'(j_1^*, j_2^*), \qquad (j_1-j_2)^* = (j_1^*, j_2^*) j_3^*,$$

so that $J = (j_1^*)'(j_2^*)'(j_1^*, j_2^*) j_3^*$ as a product of relatively prime factors. The conditions on the variables are rewritten as

$$s_{11} \mid j_1^*, \qquad s_{22} \mid j_2^*, \qquad s_{12} \mid (j_2^*)', \qquad s_{21} \mid (j_1^*)', \qquad s_3 t_1 t_2 \mid j_3^*, \qquad (r s_{14} s_{24}, J) = 1,$$

with s_3, t_1, t_2 being pairwise coprime and the same for r, s_{14}, s_{24}.

We start by summing over r, observing that

$$\sum_{\substack{n \leqslant x \\ (n,k)=1}} \frac{\mu(r)\sigma(r)}{\phi^3(r)} = \sum_{\substack{n=1 \\ (n,k)=1}}^{\infty} \frac{\mu(r)\sigma(r)}{\phi^3(r)} + O\left(\frac{1}{x}\right) = \prod_{p\nmid k}\left(1-\frac{p+1}{(p-1)^3}\right) + O\left(\frac{1}{x}\right). \tag{5.33}$$

In our case

$$\sum_{\substack{r \leqslant \min(R/(s_{11}s_{12}t_1s_3s_{14}), R/(s_{21}s_{22}t_2s_3s_{24})) \\ (r, Js_{14}s_{24})=1}} \frac{\mu(r)\sigma(r)}{\phi^3(r)} \\ = \prod_p \frac{p^3-3p^2+2p-2}{(p-1)^3} \prod_{p|Js_{14}s_{24}} \frac{(p-1)^3}{p^3-3p^2+2p-2} \\ + O\left(\frac{1}{R}\max(s_{11}s_{12}t_1s_3s_{14}, s_{21}s_{22}t_2s_3s_{24})\right) \tag{5.34}$$

is fed into (5.18). The error term of (5.34) brings

$$\ll \frac{1}{R} \sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3s_{14}\leqslant R\\ s_{22}s_{21}t_2s_3s_{24}\leqslant R}} \frac{\mu^2(s_{11})\mu^2(s_{12})\mu^2(t_1)\mu^2(t_2)\mu^2(s_{22})\mu^2(s_{21})}{\phi(s_{11})\phi^2(s_{12})\phi^2(t_1)\phi^2(t_2)\phi(s_{22})\phi^2(s_{21})} \times \frac{\mu^2(s_3)s_3}{\phi^3(s_3)}\prod_{p|s_3}(p^2-p-1)\frac{\mu^2(s_{14})\mu^2(s_{24})}{\phi^2(s_{14})\phi^2(s_{24})}\max(s_{11}s_{12}t_1s_{14}, s_{21}s_{22}t_2s_{24}), \tag{5.35}$$

in which the next summation over s_{24} reads

$$\sum_{\substack{s_{24}\leqslant(s_{11}s_{12}t_1s_{14})/(s_{21}s_{22}t_2)\\ (s_{24},Js_{14})=1}} \frac{\mu^2(s_{24})}{\phi^2(s_{24})}s_{11}s_{12}t_1s_{14} + \sum_{\substack{(s_{11}s_{12}t_1s_{14})/(s_{21}s_{22}t_2)<s_{24}\leqslant R/(s_{21}s_{22}t_2s_3)\\ (s_{24},Js_{14})=1}} \frac{\mu^2(s_{24})}{\phi^2(s_{24})}s_{21}s_{22}t_2s_{24}. \tag{5.36}$$

This makes (5.35) majorized as

$$\ll \frac{1}{R}\Bigg\{ \sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3s_{14}\leqslant R\\ s_{22}s_{21}t_2s_3\leqslant R}} \frac{\mu^2(s_{11})s_{11}}{\phi(s_{11})}\frac{\mu^2(s_{12})s_{12}}{\phi^2(s_{12})}\frac{\mu^2(t_1)t_1}{\phi^2(t_1)}\frac{\mu^2(t_2)}{\phi^2(t_2)}\frac{\mu^2(s_{22})}{\phi(s_{22})}\frac{\mu^2(s_{21})}{\phi^2(s_{21})} \times \frac{\mu^2(s_3)s_3}{\phi^3(s_3)}\prod_{p|s_3}(p^2-p-1)\frac{\mu^2(s_{14})s_{14}}{\phi^2(s_{14})}$$
$$+ \sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3s_{14}\leqslant R\\ s_{22}s_{21}t_2s_3\leqslant R}} \frac{\mu^2(s_{11})}{\phi(s_{11})}\frac{\mu^2(s_{12})}{\phi^2(s_{12})}\frac{\mu^2(t_1)}{\phi^2(t_1)}\frac{\mu^2(t_2)t_2}{\phi^2(t_2)}\frac{\mu^2(s_{22})s_{22}}{\phi(s_{22})}\frac{\mu^2(s_{21})s_{21}}{\phi^2(s_{21})} \times\frac{\mu^2(s_3)s_3}{\phi^3(s_3)}\prod_{p|s_3}(p^2-p-1)\frac{\mu^2(s_{14})}{\phi^2(s_{14})}\log R\Bigg\}, \tag{5.37}$$

where we have made use of (2.8). Now we sum over s_{14} and majorize (5.37) as

$$\ll \frac{\log R}{R}\Bigg\{ \sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3\leqslant R\\ s_{22}s_{21}t_2s_3\leqslant R}} \frac{\mu^2(s_{11})s_{11}}{\phi(s_{11})}\frac{\mu^2(s_{12})s_{12}}{\phi^2(s_{12})}\frac{\mu^2(t_1)t_1}{\phi^2(t_1)}\frac{\mu^2(t_2)}{\phi^2(t_2)}\frac{\mu^2(s_{22})}{\phi(s_{22})}\frac{\mu^2(s_{21})}{\phi^2(s_{21})} \times \frac{\mu^2(s_3)s_3}{\phi^3(s_3)}\prod_{p|s_3}(p^2-p-1)$$
$$+ \sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3\leqslant R\\ s_{22}s_{21}t_2s_3\leqslant R}} \frac{\mu^2(s_{11})}{\phi(s_{11})}\frac{\mu^2(s_{12})}{\phi^2(s_{12})}\frac{\mu^2(t_1)}{\phi^2(t_1)}\frac{\mu^2(t_2)t_2}{\phi^2(t_2)}\frac{\mu^2(s_{22})s_{22}}{\phi(s_{22})}\frac{\mu^2(s_{21})s_{21}}{\phi^2(s_{21})} \times \frac{\mu^2(s_3)s_3}{\phi^3(s_3)}\prod_{p|s_3}(p^2-p-1)\Bigg\}$$
$$\ll R^{-1+\epsilon}, \tag{5.38}$$

where to see the last line it is enough to observe that all of the summations are over variables which divide j_1 or j_2 or $j_1 - j_2$, and a more precise (but more complicated looking) factor than R^ϵ could easily be given.

We revert to (5.18) with the sum over r already performed in (5.34) so that the main term has been turned into

$$\prod_p \frac{p^3-3p^2+2p-2}{(p-1)^3} \prod_{p|J} \frac{(p-1)^3}{p^3-3p^2+2p-2} \\ \times \sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3s_{14}\leqslant R\\ s_{22}s_{21}t_2s_3s_{24}\leqslant R}} \frac{\mu^2(s_{11})\mu(s_{12})\mu^2(s_{22})\mu(s_{21})\mu(t_1)\mu(t_2)}{\phi(s_{11})\phi^2(s_{12})\phi(s_{22})\phi^2(s_{21})\phi^2(t_1)\phi^2(t_2)} \\ \times \frac{\mu^2(s_3)}{\phi^3(s_3)} \prod_{p|s_3}(p^2-p-1)\,\mu(s_{14})\mu(s_{24}) \prod_{p|s_{14}s_{24}} \frac{p-1}{p^3-3p^2+2p-2}. \tag{5.39}$$

For the sum over s_{24} we have

$$\sum_{\substack{s_{24}\leqslant R/(s_{21}s_{22}t_2s_3)\\ (s_{24},Js_{14})=1}} \mu(s_{24}) \prod_{p|s_{24}} \frac{p-1}{p^3-3p^2+2p-2} \\ = \prod_{p\nmid Js_{14}} \left(1 - \frac{p-1}{p^3-3p^2+2p-2}\right) + O\left(\frac{s_{21}s_{22}t_2s_3}{R}\right). \tag{5.40}$$

The error terms, here and in what follows, can be considered in a similar way to above ending up with $O(R^{-1+\epsilon})$ as in (5.38). So from now on we shall just concentrate on the main term, which upon (5.39) has become

$$\prod_p \frac{p^3-3p^2+p-1}{(p-1)^3} \prod_{p|J} \frac{(p-1)^3}{p^3-3p^2+p-1} \\ \times \sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3s_{14}\leqslant R\\ s_{22}s_{21}t_2s_3\leqslant R}} \frac{\mu^2(s_{11})\mu(s_{12})\mu^2(s_{22})\mu(s_{21})\mu(t_1)\mu(t_2)}{\phi(s_{11})\phi^2(s_{12})\phi(s_{22})\phi^2(s_{21})\phi^2(t_1)\phi^2(t_2)} \\ \times \frac{\mu^2(s_3)}{\phi^3(s_3)} \prod_{p|s_3}(p^2-p-1)\,\mu(s_{14}) \prod_{p|s_{14}} \frac{p-1}{p^3-3p^2+p-1}. \tag{5.41}$$

Next in the row is the sum over s_{14},

$$\sum_{\substack{s_{14}\leqslant R/(s_{11}s_{12}t_1s_3)\\ (s_{14},J)=1}} \mu(s_{14}) \prod_{p|s_{14}} \frac{p-1}{p^3-3p^2+p-1} \\ = \prod_{p\nmid J} \left(\frac{p^2(p-3)}{p^3-3p^2+p-1}\right) + O\left(\frac{s_{11}s_{12}t_1s_3}{R}\right). \tag{5.42}$$

This shows that if $3 \nmid J$, then the main term is 0. Note that $2 \mid J$ always. So now we can express the main term as

$$[3 \mid J] \prod_{p>3} \frac{p^2(p-3)}{(p-1)^3} \prod_{\substack{p|J\\ p>3}} \frac{(p-1)^3}{p^2(p-3)} \\ \times \sideset{}{^*}\sum_{\substack{s_{11}s_{12}t_1s_3\leqslant R\\ s_{22}s_{21}t_2s_3\leqslant R}} \frac{\mu^2(s_{11})\mu(s_{12})\mu^2(s_{22})\mu(s_{21})\mu(t_1)\mu(t_2)}{\phi(s_{11})\phi^2(s_{12})\phi(s_{22})\phi^2(s_{21})\phi^2(t_1)\phi^2(t_2)} \frac{\mu^2(s_3)}{\phi^3(s_3)} \prod_{p|s_3}(p^2-p-1). \tag{5.43}$$

From now on the inequality conditions in $\sum^*$ become superfluous as all of the remaining variables to be summed over are divisors of J, which is $\ll R^\epsilon$, and therefore satisfy these inequalities anyway. The sum over s_3 is

$$\sum_{s_3|j_3^*/(t_1t_2)} \mu^2(s_3)\prod_{p|s_3}\frac{p^2-p-1}{(p-1)^3} = \prod_{p|j_3^*/(t_1t_2)}\frac{p^3-2p^2+2p-2}{(p-1)^3}, \tag{5.44}$$

turning the main term into

$$[3\mid J]\prod_{p>3}\frac{p^2(p-3)}{(p-1)^3}\prod_{\substack{p|J\\ p>3}}\frac{(p-1)^3}{p^2(p-3)}\prod_{p|j_3^*}\frac{p^3-2p^2+2p-2}{(p-1)^3}$$
$$\times\sum{}^{*}\frac{\mu^2(s_{11})\mu(s_{12})\mu^2(s_{22})\mu(s_{21})}{\phi(s_{11})\phi^2(s_{12})\phi(s_{22})\phi^2(s_{21})}\mu(t_1)\mu(t_2)\prod_{p|t_1t_2}\frac{p-1}{p^3-2p^2+2p-2}. \tag{5.45}$$

Now the sum over t_2 is

$$\sum_{t_2|(j_3^*/t_1)}\mu(t_2)\prod_{p|t_2}\frac{p-1}{p^3-2p^2+2p-2} = \prod_{p|(j_3^*/t_1)}\frac{p^3-2p^2+p-1}{p^3-2p^2+2p-2}, \tag{5.46}$$

and the main term becomes

$$[3\mid J]\prod_{p>3}\frac{p^2(p-3)}{(p-1)^3}\prod_{\substack{p|J\\ p>3}}\frac{(p-1)^3}{p^2(p-3)}\prod_{p|j_3^*}\frac{p^3-2p^2+p-1}{(p-1)^3}$$
$$\times\sum{}^{*}\frac{\mu^2(s_{11})\mu(s_{12})\mu^2(s_{22})\mu(s_{21})}{\phi(s_{11})\phi^2(s_{12})\phi(s_{22})\phi^2(s_{21})}\mu(t_1)\prod_{p|t_1}\frac{p-1}{p^3-2p^2+p-1}. \tag{5.47}$$

We continue by summing over t_1,

$$\sum_{t_1|j_3^*}\mu(t_1)\prod_{p|t_1}\frac{p-1}{p^3-2p^2+p-1} = \prod_{p|j_3^*}\frac{p^2(p-2)}{p^3-2p^2+p-1}, \tag{5.48}$$

where the very last product is 0 if $2 \mid j_3^*$ (which is equivalent to $2\nmid j_1j_2$). Now the main term has been simplified to

$$[3\mid J]\prod_{p>3}\frac{p^2(p-3)}{(p-1)^3}\prod_{\substack{p|J\\ p>3}}\frac{(p-1)^3}{p^2(p-3)}\prod_{p|j_3^*}\frac{p^2(p-2)}{(p-1)^3}\sum{}^{*}\frac{\mu^2(s_{11})\mu(s_{12})\mu^2(s_{22})\mu(s_{21})}{\phi(s_{11})\phi^2(s_{12})\phi(s_{22})\phi^2(s_{21})}. \tag{5.49}$$

For the final reduction of the main term we have

$$\begin{aligned}
\sum_{s_{11}|j_1^*}\frac{\mu^2(s_{11})}{\phi(s_{11})} &= \frac{j_1^*}{\phi(j_1^*)} = \frac{(j_1^*)'(j_1^*,j_2^*)}{\phi((j_1^*)')\phi((j_1^*,j_2^*))},\\
\sum_{s_{22}|j_2^*}\frac{\mu^2(s_{22})}{\phi(s_{22})} &= \frac{(j_2^*)'(j_1^*,j_2^*)}{\phi((j_2^*)')\phi((j_1^*,j_2^*))},\\
\sum_{s_{12}|(j_2^*)'}\frac{\mu(s_{12})}{\phi^2(s_{12})} &= \prod_{p|(j_2^*)'}\frac{p(p-2)}{(p-1)^2},\\
\sum_{s_{21}|(j_1^*)'}\frac{\mu(s_{21})}{\phi^2(s_{21})} &= \prod_{p|(j_1^*)'}\frac{p(p-2)}{(p-1)^2},
\end{aligned} \tag{5.50}$$

which shows that the main term is 0 if $2 \mid (j_1^*)'(j_2^*)'$. Hence, since $2\mid J$, in order to have a non-zero main term it must be that $2 \mid (j_1^*, j_2^*)$. Thus we find that the result of the summation

(5.18) is

$$[2 \mid (j_1^*, j_2^*)][3 \mid J] \prod_{p>3} \frac{p^2(p-3)}{(p-1)^3} \prod_{\substack{p|J \\ p>3}} \frac{(p-1)^3}{p^2(p-3)} \prod_{p|j_3^*} \frac{p^2(p-2)}{(p-1)^3} \times \frac{(j_1^*)'}{\phi((j_1^*)')} \frac{(j_2^*)'}{\phi((j_2^*)')} \left(\frac{(j_1^*, j_2^*)}{\phi((j_1^*, j_2^*))}\right)^2 \prod_{p|j_1^* j_2^*} \frac{p(p-2)}{(p-1)^2}. \quad (5.51)$$

To re-organize (5.51), note that the product over all $p > 3$ is $\frac{4}{3}C_2C_3$ (see (1.13)), and recall that for the main term to be non-zero 2 must divide (j_1^*, j_2^*) so that $2 \nmid (j_1^*)'(j_2^*)'j_3^*$, and then consider one by one the four possibilities arising from $3 \mid (j_1^*)'(j_2^*)'(j_1^*, j_2^*)j_3^*$ as to which factor 3 divides. In this way we find that the main term is

$$[2 \mid (j_1, j_2)][3 \mid j_1 j_2(j_1 - j_2)]\, 6C_2C_3 \prod_{\substack{p|(j_1,j_2) \\ p>2}} \left(\frac{p-1}{p-2}\right) \prod_{\substack{p|j_1j_2(j_1-j_2) \\ p>3}} \left(\frac{p-2}{p-3}\right) = \mathfrak{S}_2((j_1, j_2))\mathfrak{S}_3(j_1j_2(j_1 - j_2)), \quad (5.52)$$

and, by (1.10), this completes the proof of Theorem 2 for the case $k = r = 3$.

6. *Re-expression of the pure triple correlations* $\mathcal{S}_3(N, \boldsymbol{j}, \boldsymbol{a})$

In this section the sum

$$\mathcal{S}_3(N, \boldsymbol{j}, \boldsymbol{a}) = \sum_{n=1}^{N} \lambda_R(n + j_1)\lambda_R(n + j_2)\lambda_R(n) \quad (6.1)$$

will be reduced to a multiple sum over relatively prime variables. Later on three cases will be considered: if $j_1 = j_2 = 0$, then the sum is $\mathcal{S}_3(N, (0), (3))$; if $j_1 = j_2 = j \neq 0$, then the sum is $\mathcal{S}_3(N, (0, j), (1, 2))$; if $j_1 \neq j_2$ and $j_1 j_2 \neq 0$, then the sum is $\mathcal{S}_3(N, (0, j_1, j_2), (1, 1, 1))$. From the definition of the $\lambda_R(n)$ we have

$$\mathcal{S}_3(N, \boldsymbol{j}, \boldsymbol{a}) = \sum_{r_1,r_2,r_3 \leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)\mu^2(r_3)}{\phi(r_1)\phi(r_2)\phi(r_3)} \sum_{\substack{d|r_1 \\ e|r_2 \\ f|r_3}} d\mu(d)e\mu(e)f\mu(f) \sum_{\substack{n \leqslant N \\ n \equiv -j_1 (\mathrm{mod}\, d) \\ n \equiv -j_2 (\mathrm{mod}\, e) \\ n \equiv 0 (\mathrm{mod}\, f)}} 1. \quad (6.2)$$

The innermost sum is over the values of n in a unique residue class modulo $[d, e, f]$ whenever $(d, e) \mid j_1 - j_2$, $(d, f) \mid j_1$, and $(e, f) \mid j_2$ in which case its value is $N/[d, e, f] + O(1)$, otherwise the innermost sum is void. As in (4.5), by (4.2) the last $O(1)$ leads to a contribution of $O(R^3)$ in (6.2). Hence

$$\mathcal{S}_3(N, \boldsymbol{j}, \boldsymbol{a}) = N \sum_{r_1,r_2,r_3 \leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)\mu^2(r_3)}{\phi(r_1)\phi(r_2)\phi(r_3)} \sum_{\substack{d|r_1, e|r_2, f|r_3 \\ (d,e)|j_1-j_2 \\ (d,f)|j_1, (e,f)|j_2}} \frac{d\mu(d)e\mu(e)f\mu(f)}{[d, e, f]} + O(R^3). \quad (6.3)$$

We can express the summation variables as products of coprime factors (since the Möbius function restricts us to the square-free r_i)

$$r_1 = a_1 a_{12} a_{13} a_{123}, \qquad r_2 = a_2 a_{12} a_{23} a_{123}, \qquad r_3 = a_3 a_{13} a_{23} a_{123}, \quad (6.4)$$

with the understanding that for a subscript χ, a_χ is a divisor of those r_j where j occurs in χ. We can now write

$$d = d_1 d_{12} d_{13} d_{123}, \qquad e = e_2 e_{12} e_{23} e_{123}, \qquad f = f_3 f_{13} f_{23} f_{123}, \tag{6.5}$$

where a d or e or f with a certain subscript is a divisor of the a with the same subscript (for example, $d_{12} \mid a_{12}$). Then we have

$$[d, e, f] = d_1 e_2 f_3 [d_{12}, e_{12}][d_{13}, f_{13}][e_{23}, f_{23}][d_{123}, e_{123}, f_{123}]. \tag{6.6}$$

So the inner sum of (6.3) over d, e, f becomes

$$\sideset{}{^*}\sum_{\substack{d_1 d_{12} d_{13} d_{123} | a_1 a_{12} a_{13} a_{123} \\ e_2 e_{12} e_{23} e_{123} | a_2 a_{12} a_{23} a_{123} \\ f_3 f_{13} f_{23} f_{123} | a_3 a_{13} a_{23} a_{123}}} \frac{d_{12} e_{12}}{[d_{12}, e_{12}]} \frac{d_{13} f_{13}}{[d_{13}, f_{13}]} \frac{e_{23} f_{23}}{[e_{23}, f_{23}]} \frac{d_{123} e_{123} f_{123}}{[d_{123}, e_{123}, f_{123}]} \mu(d_1) \dots \mu(f_{123}), \tag{6.7}$$

where ... indicates that we have the Möbius functions of all of the twelve variables coming from (6.5), and the star in $\sum^*$ reminds us that the variables of summation also obey the conditions

$$\begin{aligned} (d, e) &= (d_{12}, e_{12})(d_{123}, e_{123}) \mid j_1 - j_2, \\ (d, f) &= (d_{13}, f_{13})(d_{123}, f_{123}) \mid j_1, \\ (e, f) &= (e_{23}, f_{23})(e_{123}, f_{123}) \mid j_2. \end{aligned} \tag{6.8}$$

Now (6.7) can be broken into simpler sums as

$$\begin{aligned} &\sum_{\substack{d_{123}, e_{123}, f_{123} | a_{123} \\ (d_{123}, e_{123}) | j_1 - j_2 \\ (d_{123}, f_{123}) | j_1 \\ (e_{123}, f_{123}) | j_2}} \frac{\mu(d_{123})\mu(e_{123})\mu(f_{123}) d_{123} e_{123} f_{123}}{[d_{123}, e_{123}, f_{123}]} \\ &\times \sum_{\substack{d_{12}, e_{12} | a_{12} \\ (d_{12}, e_{12}) | (j_1 - j_2)/(d_{123}, e_{123})}} \mu(d_{12})\mu(e_{12})(d_{12}, e_{12}) \sum_{\substack{d_{13}, f_{13} | a_{13} \\ (d_{13}, f_{13}) | (j_1/(d_{123}, f_{123}))}} \mu(d_{13})\mu(f_{13})(d_{13}, f_{13}) \\ &\times \sum_{\substack{e_{23}, f_{23} | a_{23} \\ (e_{23}, f_{23}) | (j_2/(e_{123}, f_{123}))}} \mu(e_{23})\mu(f_{23})(e_{23}, f_{23}) \sum_{d_1 | a_1} \mu(d_1) \sum_{e_2 | a_2} \mu(e_2) \sum_{f_3 | a_3} \mu(f_3). \end{aligned} \tag{6.9}$$

The last three sums yield a non-zero contribution only if $a_1 = a_2 = a_3 = 1$. As for the other sums, by multiplicativity, it suffices to evaluate them when the a_χ are prime. We have, for square-free a_χ,

$$\begin{aligned} \sum_{\substack{d_{12}, e_{12} | a_{12} \\ (d_{12}, e_{12}) | (j_1 - j_2)/(d_{123}, e_{123})}} \mu(d_{12})\mu(e_{12})(d_{12}, e_{12}) &= \mu(a_{12})\mu \cdot \phi(a_{12}, j_1 - j_2), \\ \sum_{\substack{d_{13}, f_{13} | a_{13} \\ (d_{13}, f_{13}) | (j_1/(d_{123}, f_{123}))}} \mu(d_{13})\mu(f_{13})(d_{13}, f_{13}) &= \mu(a_{13})\mu \cdot \phi(a_{13}, j_1), \\ \sum_{\substack{e_{23}, f_{23} | a_{23} \\ (e_{23}, f_{23}) | (j_2/(e_{123}, f_{123}))}} \mu(e_{23})\mu(f_{23})(e_{23}, f_{23}) &= \mu(a_{23})\mu \cdot \phi(a_{23}, j_2), \end{aligned} \tag{6.10}$$

and

$$\sum_{\substack{d_{123},\, e_{123},\, f_{123} | a_{123} \\ (d_{123}, e_{123}) | j_1 - j_2 \\ (d_{123}, f_{123}) | j_1 \\ (e_{123}, f_{123}) | j_2}} \frac{\mu(d_{123})\mu(e_{123})\mu(f_{123}) d_{123} e_{123} f_{123}}{[d_{123}, e_{123}, f_{123}]}$$
$$= \mu\cdot\phi((a_{123}, j_1, j_2))\phi_2((a_{123}, j_1)) \times \phi_2\left(\frac{(a_{123}, j_2)}{(a_{123}, j_1, j_2)}\right)\phi_2\left(\frac{(a_{123}, j_1 - j_2)}{(a_{123}, j_1 - j_2, j_1 j_2)}\right) \prod_{\substack{p | a_{123} \\ p \nmid j_1 j_2 (j_1 - j_2)}} (-2). \tag{6.11}$$

Thus the sum in the main term of (6.3) has been transformed into a sum over pairwise coprime variables a_χ,

$$\sum_{\substack{a_{12} a_{13} a_{123} \leqslant R \\ a_{12} a_{23} a_{123} \leqslant R \\ a_{13} a_{23} a_{123} \leqslant R}}{}' \frac{\mu(a_{12})\mu\cdot\phi((a_{12}, j_1 - j_2))\mu(a_{13})\mu\cdot\phi((a_{13}, j_1))\mu(a_{23})\mu\cdot\phi((a_{23}, j_2))}{\phi^2(a_{12})\phi^2(a_{13})\phi^2(a_{23})}$$
$$\times \frac{\mu^2(a_{123})\mu\cdot\phi((a_{123}, j_1, j_2))\phi_2((a_{123}, j_1))\phi_2\big(\frac{(a_{123}, j_2)}{(a_{123}, j_1, j_2)}\big)\phi_2\big(\frac{(a_{123}, j_1 - j_2)}{(a_{123}, j_1 - j_2, j_1 j_2)}\big)}{\phi^3(a_{123})}$$
$$\times \mu\cdot d\left(\frac{a_{123}}{(a_{123}, j_1 j_2 (j_1 - j_2))}\right). \tag{6.12}$$

7. *Pure triple correlations: the case $j_1 = j_2 = 0$*

In this case (6.12) reads

$$\sum_{\substack{uvy \leqslant R \\ uwy \leqslant R \\ vwy \leqslant R}}{}' \frac{\mu^2(u)\mu^2(v)\mu^2(w)\mu(y)\phi_2(y)}{\phi(u)\phi(v)\phi(w)\phi^2(y)}, \tag{7.1}$$

where u, v, w and y are pairwise coprime. We begin by summing over u as

$$\sum_{\substack{u \leqslant \min(R/(vy), R/(wy)) \\ (u, vwy) = 1}} \frac{\mu^2(u)}{\phi(u)}, \tag{7.2}$$

which is evaluated by (2.15), and plugging into (7.1) we have

$$\sum_{vwy \leqslant R}{}' \frac{\mu^2(v)\mu^2(w)\mu(y)\phi_2(y)}{vwy\phi(y)} \left[\log\min\left(\frac{R}{vy}, \frac{R}{wy}\right) + D_1 + \sum_{p | vwy} \frac{\log p}{p}\right]$$
$$+ O\left(\sum_{vwy \leqslant R}{}' \frac{\mu^2(v)\mu^2(w)\mu^2(y)\phi_2(y) m(vwy)}{\phi(v)\phi(w)\phi^2(y)\sqrt{\min(R/(vy), R/(wy))}}\right), \tag{7.3}$$

where we have written D_1 for the constant

$$\gamma + \sum_p \frac{\log p}{p(p-1)}.$$

We will regard the log min occurring in (7.3) as

$$\log\left(\frac{R}{vwy}\right) + \log\min(v,w).$$

First, upon letting $z = vwy$, we have

$$\begin{aligned}
\sideset{}{'}\sum_{vwy\leqslant R} \frac{\mu^2(v)\mu^2(w)\mu(y)\phi_2(y)}{vwy\phi(y)} \log\left(\frac{R}{vwy}\right) &= \sum_{z\leqslant R} \frac{\mu^2(z)}{z}\log\left(\frac{R}{z}\right)\sum_{y|z}\frac{\mu(y)\phi_2(y)}{\phi(y)}\sum_{w|(z/y)} 1 \\
&= \sum_{z\leqslant R} \frac{\mu^2(z)d(z)}{z}\log\left(\frac{R}{z}\right)\sum_{y|z}\frac{\mu(y)\phi_2(y)}{d(y)\phi(y)} \\
&= \sum_{z\leqslant R} \frac{\mu^2(z)d(z)}{z}\log\left(\frac{R}{z}\right)\prod_{p|z}\left(1-\frac{p-2}{2(p-1)}\right) \\
&= \sum_{z\leqslant R} \frac{\mu^2(z)}{\phi(z)}\log\left(\frac{R}{z}\right) \\
&= \tfrac{1}{2}\log^2 R + O(\log R), \qquad (7.4)
\end{aligned}$$

by (2.16). Similarly, we have

$$\sideset{}{'}\sum_{vwy\leqslant R} \frac{\mu^2(v)\mu^2(w)\mu(y)\phi_2(y)}{vwy\phi(y)} = \sum_{z\leqslant R}\frac{\mu^2(z)}{\phi(z)} = \log R + O(1), \qquad (7.5)$$

and

$$\begin{aligned}
\sideset{}{'}\sum_{vwy\leqslant R} \frac{\mu^2(v)\mu^2(w)\mu(y)\phi_2(y)}{vwy\phi(y)} \sum_{p|vwy}\frac{\log p}{p} &= \sum_{z\leqslant R}\frac{\mu^2(z)}{\phi(z)}\sum_{p|z}\frac{\log p}{p} \\
&= \sum_{p\leqslant R}\frac{\log p}{p}\sum_{\substack{z\leqslant R\\ p|z}}\frac{\mu^2(z)}{\phi(z)} \\
&= \sum_{p\leqslant R}\frac{\log p}{p\phi(p)}\sum_{\substack{m\leqslant R/p\\ (m,p)=1}}\frac{\mu^2(m)}{\phi(m)} \\
&\ll \sum_{p\leqslant R}\frac{\log(2R/p)\log p}{p^2} \ll \log R. \qquad (7.6)
\end{aligned}$$

We can take the $\log\min(v,w)$-term as twice the summands with $w<v$, that is,

$$2\sideset{}{'}\sum_{\substack{vwy\leqslant R\\ w<v}} \frac{\mu^2(v)\mu^2(w)\mu(y)\phi_2(y)}{vwy\phi(y)}\log w. \qquad (7.7)$$

Now observe that in the sum of (7.5) the terms with $v=w$, which can only be present if $v=w=1$, contribute

$$\sum_{y\leqslant R}\frac{\mu(y)\phi_2(y)}{y\phi(y)} \ll 1, \qquad (7.8)$$

by Lemma 2. So the sum on the left-hand side of (7.5) is

$$2\sideset{}{'}\sum_{\substack{vwy\leqslant R\\ w<v}} \frac{\mu^2(v)\mu^2(w)\mu(y)\phi_2(y)}{vwy\phi(y)} + O(1).$$

Writing

$$\sum_{\substack{y\leqslant R/(vw)\\(y,vw)=1}} \frac{\mu(y)\phi_2(y)}{y\phi(y)} = f_{vw}\left(\frac{R}{vw}\right), \qquad \sum_{\substack{w<v\leqslant R/w\\(v,w)=1}} \frac{\mu^2(v)}{v} f_{vw}\left(\frac{R}{vw}\right) = g_w(w) \tag{7.9}$$

(the subscripts of f and g refer to the coprimality conditions in the sums), we see that (7.5) says that

$$2\sum_{w\leqslant\sqrt{R}} \frac{\mu^2(w)}{w} g_w(w) = \log R + O(1). \tag{7.10}$$

With this notation (7.7) can be rewritten as

$$2\sum_{w\leqslant\sqrt{R}} \frac{\mu^2(w)}{w} g_w(w)\log w, \tag{7.11}$$

and by partial summation on (7.10) we find that

$$2\sum_{w\leqslant\sqrt{R}} \frac{\mu^2(w)}{w} g_w(w)\log w = \tfrac{1}{4}\log^2 R + O(\log R). \tag{7.12}$$

Thus the first line of (7.3) has been shown to be $(\frac{3}{4}\log^2 R + O(\log 2R))$, for $R\geqslant 1$.

It remains to consider the error term of (7.3), which can be rewritten as

$$\frac{2}{\sqrt{R}} \sideset{}{'}\sum_{\substack{vwy\leqslant R\\w<v}} \frac{\mu^2(v)\mu^2(w)\mu^2(y)m(vwy)\phi_2(y)\sqrt{vy}}{\phi(v)\phi(w)\phi^2(y)} + \frac{1}{\sqrt{R}}\sum_{y\leqslant R}\frac{\mu^2(y)m(y)\phi_2(y)\sqrt{y}}{\phi^2(y)}. \tag{7.13}$$

Let

$$h_{vw}\left(\frac{R}{vw}\right) = \sum_{\substack{y\leqslant R/(vw)\\(y,vw)=1}} \frac{\mu^2(y)m(y)\phi_2(y)\sqrt{y}}{\phi^2(y)}, \tag{7.14}$$

so that (7.13) is

$$\begin{aligned}
&\frac{2}{\sqrt{R}} \sideset{}{'}\sum_{\substack{vw\leqslant R\\w<v}} \frac{\mu^2(v)\mu^2(w)m(vw)\sqrt{v}}{\phi(v)\phi(w)} h_{vw}\left(\frac{R}{vw}\right) + h_1(R)\\
&\qquad = \frac{2}{\sqrt{R}}\sum_{w\leqslant\sqrt{R}}\frac{\mu^2(w)m(w)}{\phi(w)} \sum_{\substack{w<v\leqslant R/w\\(v,w)=1}} \frac{\mu^2(v)m(v)\sqrt{v}}{\phi(v)} h_{vw}\left(\frac{R}{vw}\right) + h_1(R)\\
&\qquad = \frac{2}{\sqrt{R}}\sum_{w\leqslant\sqrt{R}}\frac{\mu^2(w)m(w)}{\phi(w)} k_w(w) + h_1(R),
\end{aligned} \tag{7.15}$$

say. Note that $h_1(R)$ is the contribution of the terms with $v=w=1$ in the error term of (7.3), and we know by (2.11) that

$$h_1(R) \leqslant \frac{1}{\sqrt{R}}\sum_{y\leqslant R}\frac{\mu^2(y)m(y)}{\sqrt{y}} \ll 1. \tag{7.16}$$

Now consider the expression

$$\frac{1}{\sqrt{R}} \sideset{}{'}\sum_{vwy\leqslant R} \frac{\mu^2(v)\mu^2(w)\mu^2(y)m(vwy)\phi_2(y)\sqrt{vwy}}{\phi(v)\phi(w)\phi^2(y)}. \tag{7.17}$$

On one hand, with the notation defined in (7.14) and (7.15), (7.17) may be re-expressed as

$$\frac{2}{\sqrt{R}}\sum_{w\leqslant\sqrt{R}}\frac{\mu^2(w)m(w)\sqrt{w}}{\phi(w)}k_w(w)+h_1(R). \tag{7.18}$$

On the other hand, by putting $z=vwy$, we find that the sum of (7.17) becomes

$$\begin{aligned}\sum_{z\leqslant R}\frac{\mu^2(z)m(z)\sqrt{z}}{\phi(z)}\sum_{y|z}\frac{\phi_2(y)}{\phi(y)}\sum_{w|(z/y)}1 &= \sum_{z\leqslant R}\frac{\mu^2(z)m(z)\sqrt{z}d(z)}{\phi(z)}\sum_{y|z}\frac{\phi_2(y)}{d(y)\phi(y)}\\ &= \sum_{z\leqslant R}\frac{\mu^2(z)m(z)\sqrt{z}d(z)}{\phi(z)}\prod_{p|z}\left(1+\frac{p-2}{2(p-1)}\right)\\ &= \sum_{z\leqslant R}\mu^2(z)\prod_{p|z}\frac{3p-4}{(p-1)(\sqrt{p}-1)}. \end{aligned}\tag{7.19}$$

The last sum has been calculated in Lemma 3, giving

$$\sum_{w\leqslant\sqrt{R}}\frac{\mu^2(w)m(w)\sqrt{w}}{\phi(w)}k_w(w)=2P(1)R^{1/2}\log^2\sqrt{R}+O(R^{1/2}\log R). \tag{7.20}$$

From here we obtain by partial summation

$$\sum_{w\leqslant\sqrt{R}}\frac{\mu^2(w)m(w)}{\phi(w)}k_w(w)=P(1)R^{1/4}\log^2 R+O(R^{1/4}\log R). \tag{7.21}$$

By (7.21) and (7.16), the quantity in (7.15), that is, the error term of (7.3), is $O(1)$. This completes the evaluation.

8. *Pure triple correlations: the case* $j_1=j_2=j\neq 0$

In this case (6.12) assumes the form

$$\sideset{}{'}\sum_{\substack{uvy\leqslant R\\ uwy\leqslant R\\ vwy\leqslant R}}\frac{\mu^2(u)\mu(v)\mu\cdot\phi((v,j))\mu(w)\mu\cdot\phi((w,j))\mu^2(y)\mu\cdot\phi((y,j))\phi_2(y)}{\phi(u)\phi^2(v)\phi^2(w)\phi^3(y)}, \tag{8.1}$$

and calculating first the sum over u as in (7.2), we get

$$\begin{aligned}&\sideset{}{'}\sum_{vwy\leqslant R}\frac{\mu(v)\mu\cdot\phi((v,j))\mu(w)\mu\cdot\phi((w,j))\mu^2(y)\mu\cdot\phi((y,j))\phi_2(y)}{v\phi(v)w\phi(w)y\phi^2(y)}\\ &\qquad\times\left[\log\min\left(\frac{R}{vy},\frac{R}{wy}\right)+D_1+\sum_{p|vwy}\frac{\log p}{p}\right]\\ &\qquad+O\left(\sideset{}{'}\sum_{vwy\leqslant R}\frac{\mu^2(v)\mu^2(w)\mu^2(y)\phi_2(y)\phi((vwy,j))m(vwy)}{\phi^2(v)\phi^2(w)\phi^3(y)\sqrt{\min(R/(vy),R/(wy))}}\right).\end{aligned}\tag{8.2}$$

The last error term is at most as large as the error term of (7.3) which was shown to be $O(1)$. As in (7.4), letting $z=vwy$, the log term in the brackets being $\log(R/z)+\log\min(v,w)$, we view the main part of (8.2) except for the contribution of the last log min term in the form

$$\sum_{z\leqslant R}\frac{\mu(z)\mu\cdot\phi((z,j))a(z)}{z\phi(z)}\sum_{y|z}\frac{\mu(y)\phi_2(y)}{\phi(y)}\sum_{w|(z/y)}1=\sum_{z\leqslant R}\frac{\mu(z)\mu\cdot\phi((z,j))a(z)}{\phi^2(z)}. \tag{8.3}$$

In the present situation

$$a(z) = \log\left(\frac{R}{z}\right) + D_1 + \sum_{p|z} \frac{\log p}{p}.$$

By Lemma 4 and (2.21), we have

$$\begin{aligned}\sum_{z\leqslant R} \frac{\mu(z)\mu\cdot\phi((z,j))}{\phi^2(z)}\left(\log\frac{R}{z} + D_1\right)\\ &= \mathfrak{S}_2(j)(\log R + D_1) + O\left(\frac{j^* d(j^*)\log R}{\phi(j^*)R}\right)\\ &\quad + [2\nmid j]\mathfrak{S}_2(2j)\frac{\log 2}{2} + [2\mid j]\mathfrak{S}_2(j)\left[\sum_{p\nmid j}\frac{\log p}{p(p-2)} - \sum_{p|j}\frac{\log p}{p}\right]. \qquad (8.4)\end{aligned}$$

Next we have, by Lemma 4,

$$\begin{aligned}\sum_{z\leqslant R}\frac{\mu(z)\mu\cdot\phi((z,j))}{\phi^2(z)}\sum_{p|z}\frac{\log p}{p}\\ &= \sum_{p\leqslant R}\frac{\log p}{p}\sum_{\substack{z\leqslant R\\ p|z}}\frac{\mu(z)\mu\cdot\phi((z,j))}{\phi^2(z)}\\ &= \sum_{p\leqslant R}\frac{\log p}{p}\frac{\mu(p)\mu\cdot\phi((p,j))}{\phi^2(p)}\sum_{\substack{u\leqslant R/p\\ (u,p)=1}}\frac{\mu(u)\mu\cdot\phi((u,j))}{\phi^2(u)}\\ &= -\frac{\mu((2,j))\log 2}{2}\sum_{\substack{u\leqslant R/2\\ (u,2)=1}}\frac{\mu(u)\mu\cdot\phi((u,j))}{\phi^2(u)}\\ &\quad - \sum_{2<p\leqslant R}\frac{\mu\cdot\phi((p,j))\log p}{p\phi^2(p)}\sum_{\substack{u\leqslant R/p\\ (u,p)=1}}\frac{\mu(u)\mu\cdot\phi((u,j))}{\phi^2(u)}\\ &= -\frac{\mu((2,j))\log 2}{4}\mathfrak{S}_2(2j) - [2\mid j]\,2C_2\sum_{2<p\leqslant R}\frac{\mu\cdot\phi((p,j))\log p}{p^2(p-2)}\prod_{\substack{q:\text{prime}\\ q|j\\ q\neq 2,\,p}}\left(\frac{q-1}{q-2}\right)\\ &\quad + O\left(\frac{j^*d(j^*)}{R\phi(j^*)}\sum_{p\leqslant R}\frac{\phi((p,j))\log p}{(p-1)^2}\right)\\ &= -\frac{\mu((2,j))\log 2}{4}\mathfrak{S}_2(2j) + \mathfrak{S}_2(j)\left[\sum_{\substack{2<p\leqslant R\\ p|j}}\frac{\log p}{p^2} - \sum_{\substack{2<p\leqslant R\\ p\nmid j}}\frac{\log p}{p^2(p-2)}\right]\\ &\quad + O\left(\frac{j^*d(j^*)}{R\phi(j^*)}\left[\sum_{\substack{p\leqslant R\\ p|j}}\frac{\log p}{p-1} + \sum_{\substack{p\leqslant R\\ p\nmid j}}\frac{\log p}{(p-1)^2}\right]\right), \qquad (8.5)\end{aligned}$$

where, by (2.1), the sums in the very last O-term are $\ll \log\log 3j^*$.

To complete this case we deal with the log min term for which we have

$$\begin{aligned}\Bigg|\sum_{z\leqslant R}\frac{\mu(z)\mu\cdot\phi((z,j))}{z\phi(z)}\sum_{y|z}\frac{\mu(y)\phi_2(y)}{\phi(y)}\sum_{w|(z/y)}\log\min\left(\frac{z/y}{w},w\right)\Bigg|\\ &\leqslant \sum_{z\leqslant R}\frac{\mu^2(z)\phi((z,j))}{z\phi(z)}\sum_{y|z}\frac{\mu^2(y)\phi_2(y)}{\phi(y)}d\left(\frac{z}{y}\right)\log z\\ &= \sum_{z\leqslant R}\frac{\mu^2(z)\phi((z,j))d(z)\log z}{z\phi(z)}\prod_{p|z}\left(1+\frac{p-2}{2(p-1)}\right)\\ &\leqslant \sum_{z\leqslant R}\frac{\mu^2(z)\phi((z,j))3^{\omega(z)}\log z}{z\phi(z)}\\ &= \sum_{d|j^*}\phi(d)\sum_{\substack{z\leqslant R\\(z,j)=d}}\frac{\mu^2(z)3^{\omega(z)}\log z}{z\phi(z)}\\ &= \sum_{d|j^*}\frac{3^{\omega(d)}}{d}\sum_{\substack{t\leqslant R/d\\(t,j)=1}}\frac{\mu^2(t)3^{\omega(t)}}{t\phi(t)}\log(dt).\end{aligned}\tag{8.6}$$

Since $\omega(t)=o(\log t)$ by the prime number theorem, the last line of (8.6) is bounded as

$$\ll 1+\sum_{d|j^*}\frac{3^{\omega(d)}\log d}{d}=1+\sum_{p|j^*}\log p\sum_{\substack{k|j^*\\p|k}}\frac{3^{\omega(k)}}{k},\tag{8.7}$$

in which the inner sum over k is equal to

$$\frac{3}{p}\sum_{m|(j^*/p)}\frac{3^{\omega(m)}}{m}=\frac{3}{p}\prod_{\substack{q|(j^*/p)\\q:\text{prime}}}\left(1+\frac{3}{q}\right)=\frac{3}{p+3}\prod_{\substack{q|j^*\\q:\text{prime}}}\left(1+\frac{3}{q}\right).$$

Hence (8.7) is equal to

$$1+3\prod_{p|j^*}\left(1+\frac{3}{p}\right)\sum_{p|j^*}\frac{\log p}{p+3}\ll(\log\log 3j^*)^4.\tag{8.8}$$

The results of this section are combined to give the evaluation of (8.1) when $R\to\infty$ as

$$\mathfrak{S}_2(j)\log R+O((\log\log 3j^*)^4).\tag{8.9}$$

9. *Pure triple correlations: the case $j_1\neq j_2$ and $j_1j_2\neq 0$*

We begin the evaluation of (6.12) by employing Lemma 4 to sum over a_{12} as

$$\begin{aligned}&\sum_{\substack{a_{12}\leqslant\min(R/(a_{13}a_{123}),R/(a_{23}a_{123}))\\(a_{12},a_{13}a_{23}a_{123})=1}}\frac{\mu(a_{12})\mu\cdot\phi((a_{12},j_1-j_2))}{\phi^2(a_{12})}\\ &\quad=\{1-[2\nmid a_{13}a_{23}a_{123}]\mu((2,j_1-j_2))\}C_2\prod_{\substack{p|a_{13}a_{23}a_{123}\\p>2}}\frac{(p-1)^2}{p(p-2)}\prod_{\substack{p|j_1-j_2\\p\nmid a_{13}a_{23}a_{123}\\p>2}}\left(\frac{p-1}{p-2}\right)\\ &\qquad+O\left(\frac{d(j')j'}{\phi(j')\min(R/(a_{13}a_{123}),R/(a_{23}a_{123}))}\right),\end{aligned}\tag{9.1}$$

where

$$j' = \frac{(j_1 - j_2)^*}{((j_1 - j_2)^*, a_{13}a_{23}a_{123})}. \tag{9.2}$$

Note that in (6.12), when a_{123} is even the ϕ_2-factors are 0 no matter what the parities of j_1 and j_2 are. Let us first consider the main term, which after (9.1) takes the form

$$\begin{aligned}
C_2 &\sideset{}{'}\sum_{\substack{a_{13}a_{23}a_{123} \leqslant R \\ a_{123}\colon \mathrm{odd}}} \{1 - [2 \nmid a_{13}a_{23}a_{123}]\mu((2, j_1 - j_2))\} \frac{\mu(a_{13})\mu \cdot \phi((a_{13}, j_1))}{\phi^2(a_{13})} \\
&\times \frac{\mu(a_{23})\mu \cdot \phi((a_{23}, j_2))}{\phi^2(a_{23})} \frac{\mu^2(a_{123})\mu \cdot \phi((a_{123}, j_1, j_2))}{\phi^3(a_{123})} \\
&\times \phi_2((a_{123}, j_1))\phi_2\left(\frac{(a_{123}, j_2)}{(a_{123}, j_1, j_2)}\right)\phi_2\left(\frac{(a_{123}, j_1 - j_2)}{(a_{123}, j_1 - j_2, j_1j_2)}\right) \\
&\times \mu \cdot d\left(\frac{a_{123}}{(a_{123}, j_1j_2(j_1 - j_2))}\right) \prod_{\substack{p \mid a_{13}a_{23}a_{123} \\ p > 2}} \frac{(p-1)^2}{p(p-2)} \prod_{\substack{p \mid j_1 - j_2 \\ p \nmid a_{13}a_{23}a_{123} \\ p > 2}} \frac{p-1}{p-2}.
\end{aligned} \tag{9.3}$$

Writing $a_{13} = v$, $a_{23} = w$, $a_{123} = y$, and $z = vwy$ turns this expression into

$$\begin{aligned}
C_2 &\prod_{\substack{p \mid j_1 - j_2 \\ p > 2}} \left(\frac{p-1}{p-2}\right) \sum_{z \leqslant R} \{1 - [2 \nmid z]\mu((2, j_1 - j_2))\} \frac{\mu(z)}{\frac{z}{(z,2)}\phi_2\left(\frac{z}{(z,2)}\right)} \frac{\phi_2}{\phi}\left(\left(\frac{z}{(z,2)}, j_1 - j_2\right)\right) \\
&\times \sum_{\substack{y \mid z \\ y\colon \mathrm{odd}}} \frac{\mu(y)\mu \cdot \phi((y, j_1, j_2))\phi_2((y, j_1))\phi_2\left(\frac{(y, j_2)}{(y, j_1, j_2)}\right)\phi_2\left(\frac{(y, j_1 - j_2)}{(y, j_1 - j_2, j_1j_2)}\right)}{\phi(y)} \\
&\times \mu \cdot d\left(\frac{y}{(y, j_1j_2(j_1 - j_2))}\right) \sum_{w \mid (z/y)} \mu \cdot \phi((w, j_2))\mu \cdot \phi\left(\left(\frac{z}{yw}, j_1\right)\right).
\end{aligned} \tag{9.4}$$

The innermost sum, over w, is equal to

$$\begin{aligned}
\mu \cdot \phi&\left(\left(\frac{z}{y}, j_1\right)\right) \sum_{w \mid (z/y)} \frac{\mu \cdot \phi((w, j_2))}{\mu \cdot \phi((w, j_1))} \\
&= \mu \cdot \phi\left(\left(\frac{z}{y}, j_1\right)\right) \prod_{p \mid (z/y, j_1, j_2)} 2 \prod_{\substack{p \mid (z/y, j_1) \\ p \nmid j_2}} \left(1 - \frac{1}{p-1}\right) \prod_{\substack{p \mid (z/y, j_2) \\ p \nmid j_1}} (1 - (p-1)) \prod_{\substack{p \mid (z/y) \\ p \nmid j_1j_2}} 2 \\
&= \mu \cdot \phi\left(\left(\frac{z}{y}, j_1\right)\right) d\left(\left(\frac{z}{y}, j_1, j_2\right)\right) d\left(\frac{z/y}{(z/y, j_1j_2)}\right) \\
&\quad \times \frac{\phi_2}{\phi}\left(\frac{(z/y, j_1)}{(z/y, j_1, j_2)}\right) \mu \cdot \phi_2\left(\frac{(z/y, j_2)}{(z/y, j_1, j_2)}\right).
\end{aligned} \tag{9.5}$$

If $j_1 - j_2$ is odd, then

$$\{1 - [2 \nmid z]\mu((2, j_1 - j_2))\} = \begin{cases} 1 & \text{if } z \text{ is even,} \\ 0 & \text{if } z \text{ is odd;} \end{cases}$$

hence only the even z contribute to the expression in (9.4). In this case z/y is even, and exactly one of j_1, j_2 is even, so that

$$\phi_2\left(\frac{(z/y, j_1)}{(z/y, j_1, j_2)}\right) \phi_2\left(\frac{(z/y, j_2)}{(z/y, j_1, j_2)}\right) = 0.$$

That is, the expression (9.4) is 0 because of the ϕ_2-factors coming from (9.5). On the other hand, if $j_1 - j_2$ is even, then $(z,2)\{1 - [2 \nmid z]\mu((2, j_1 - j_2))\} = 2$. Plugging (9.5) into (9.4), upon effecting simplifications, we find that (9.4) becomes

$$\begin{aligned}
&\mathfrak{S}_2(j_1 - j_2)\\
&\quad\times \sum_{z \leqslant R} \frac{\mu \cdot d(z) \mu \cdot \phi_2\big(\big(\frac{z}{(z,2)}, j_1\big)\big) \mu \cdot \phi_2\big(\big(\frac{z}{(z,2)}, j_2\big)\big) \mu \cdot \phi \cdot d((z, j_1, j_2)) \phi_2\big(\big(\frac{z}{(z,2)}, j_1 - j_2\big)\big)}{z \phi_2\big(\frac{z}{(z,2)}\big) \phi((z, j_1 - j_2)) d((z, j_1 j_2)) \phi_2^2\big(\big(\frac{z}{(z,2)}, j_1, j_2\big)\big)}\\
&\quad\times \sum_{y | z/(z,2)} \frac{\mu((y, j_1))\mu((y, j_2))\mu((y, j_1 j_2 (j_1 - j_2)))\phi_2((y, j_1, j_2))\phi_2((y, j_1 - j_2))d((y, j_1 j_2))}{\phi(y)\phi_2((y, j_1 - j_2, j_1 j_2))d((y, j_1, j_2))d((y, j_1 j_2 (j_1 - j_2)))}.
\end{aligned} \tag{9.6}$$

Recalling the notation used in § 5, $J = [j_1 j_2 (j_1 - j_2)]^*$, $j_1^* = (j_1^*)'(j_1^*, j_2^*)$, $j_2^* = (j_2^*)'(j_1^*, j_2^*)$, and $(j_1 - j_2)^* = (j_1^*, j_2^*) j_3^*$, so that $J = (j_1^*)'(j_2^*)'(j_1^*, j_2^*) j_3^*$ is a product of relatively prime factors, we see that the inner sum over y is expressed as

$$\begin{aligned}
\sum_{y | z/(z,2)} \frac{\mu \cdot \phi_2((y, j_1^*, j_2^*)) \mu \cdot \phi_2((y, j_3^*))}{\phi(y) d((y, j_1^*, j_2^*)) d((y, j_3^*))} &= \sum_{y | (z/(z,2))} \frac{\mu^2(y)}{\phi(y)} \frac{\mu \cdot \phi_2}{d}((y, (j_1 - j_2)))\\
&= \prod_{\substack{p | (z/(z,2))\\ p \nmid j_1 - j_2}} \left(1 + \frac{1}{p-1}\right) \prod_{p | (z/(z,2), j_1 - j_2)} \left(1 - \frac{p-2}{2(p-1)}\right)\\
&= \frac{z}{\phi(z) d((z, j_1 - j_2))},
\end{aligned} \tag{9.7}$$

since $j_1 - j_2$ must be even for non-vanishing (9.6). Hence (9.6) becomes, upon further simplifications,

$$\begin{aligned}
&\mathfrak{S}_2(j_1 - j_2)\\
&\quad\times \sum_{z \leqslant R} \frac{\mu \cdot d(z) \mu \cdot \phi_2\big(\big(\frac{z}{(z,2)}, j_1\big)\big) \mu \cdot \phi_2\big(\big(\frac{z}{(z,2)}, j_2\big)\big) \mu \cdot \phi \cdot d((z, j_1, j_2)) \phi_2\big(\big(\frac{z}{(z,2)}, j_1 - j_2\big)\big)}{\phi(z) \phi_2\big(\frac{z}{(z,2)}\big) \phi \cdot d((z, j_1 - j_2)) d((z, j_1 j_2)) \phi_2^2\big(\big(\frac{z}{(z,2)}, j_1, j_2\big)\big)}\\
&\qquad = \mathfrak{S}_2(j_1 - j_2) \sum_{z \leqslant R} \frac{\mu(z) d(z)}{\phi(z) \phi_2\big(\frac{z}{(z,2)}\big)} \frac{\mu}{d}((z, J)) \frac{\mu}{\phi}((z, j_3^*)) \phi_2\left(\left(\frac{z}{(z,2)}, J\right)\right).
\end{aligned} \tag{9.8}$$

The last sum has been settled in Lemma 5, so that (9.4), that is the main term, is evaluated as

$$\begin{aligned}
&2\, \mathfrak{S}_2(j_1 - j_2)\, [2 \nmid j_3^*] \prod_{p \nmid J} \frac{p(p-3)}{(p-1)(p-2)} \prod_{\substack{p | j_1 j_2\\ p > 2}} \frac{p}{p-1} \prod_{\substack{p | j_3^*\\ p > 2}} \frac{p(p-2)}{(p-1)^2}\\
&\qquad = [3 \mid J][2 \mid j_1 - j_2][2 \nmid j_3^*]\, 4C_2 \prod_{p \nmid J} \frac{p(p-3)}{(p-1)(p-2)} \prod_{\substack{p | J\\ p > 2}} \frac{p}{p-1} \prod_{\substack{p | (j_1^*, j_2^*)\\ p > 2}} \frac{p-1}{p-2}\\
&\qquad = [3 \mid J][2 \mid j_1 - j_2][2 \nmid j_3^*] \left(2C_2 \prod_{\substack{p | (j_1^*, j_2^*)\\ p > 2}} \frac{p-1}{p-2} \right) \left(3C_3 \prod_{\substack{p | J\\ p > 3}} \frac{p-2}{p-3} \right)\\
&\qquad = \mathfrak{S}_3(J) \mathfrak{S}_2((j_1^*, j_2^*)) = \mathfrak{S}((0, j_1, j_2)),
\end{aligned} \tag{9.9}$$

and there is an error of $O(R^{-1+\epsilon})$.

It remains to carry out the follow-up of the error term (9.1) in (6.12), which is

$$\begin{aligned}
&\ll \frac{(j_1-j_2)^* d((j_1-j_2)^*)}{\phi((j_1-j_2)^*)} \sideset{}{'}\sum_{z=vwy\leqslant R} \frac{\mu^2(z)\phi((v,j_1))\phi((w,j_2))\phi((y,j_1,j_2))}{\phi^2(z)\phi(y)\min(R/(vy),R/(wy))} \\
&\quad\times \phi_2((y,j_1))\phi_2\left(\frac{(y,j_2)}{(y,j_1,j_2)}\right)\phi_2\left(\frac{(y,j_1-j_2)}{(y,j_1-j_2,j_1j_2)}\right) d\left(\frac{y}{(y,(j_1-j_2)j_1j_2)}\right) \\
&\ll R^{-1+\epsilon}\sum_{z\leqslant R}\frac{\mu^2(z)z\phi((z,J))\phi_2(y)}{\phi^2(z)\phi(y)\min(w,v)} \\
&= R^{-1+\epsilon}\sum_{z\leqslant R}\frac{\mu^2(z)z\phi((z,J))}{\phi^2(z)}\sum_{y|z}\frac{\phi_2(y)}{\phi(y)}\sum_{w|(z/y)}\frac{1}{\min(w,z/(yw))} \\
&\leqslant R^{-1+\epsilon}\sum_{z\leqslant R}\frac{\mu^2(z)z\phi((z,J))d(z)}{\phi^2(z)}\sum_{y|z}\frac{\phi_2(y)}{\phi(y)d(y)} \\
&= R^{-1+\epsilon}\sum_{z\leqslant R}\mu^2(z)\phi((z,J))\prod_{p|z}\frac{p(3p-4)}{(p-1)^3} \\
&= R^{-1+\epsilon}\sum_{j|J}\phi(j)\sum_{\substack{z\leqslant R\\(z,J)=j}}\mu^2(z)\prod_{p|z}\frac{p(3p-4)}{(p-1)^3} \\
&= R^{-1+\epsilon}\sum_{j|J}\prod_{p|j}\left(\frac{p(3p-4)}{(p-1)^2}\right)\sum_{\substack{z'\leqslant R/j\\(z',j)=1}}\mu^2(z')\prod_{p|z'}\frac{p(3p-4)}{(p-1)^3} \\
&= R^{-1+\epsilon}\prod_{p|J}\frac{4p^2-6p+1}{(p-1)^2} \\
&\ll R^{-1+\epsilon}. \qquad (9.10)
\end{aligned}$$

This completes the proof of Theorem 1 in the case considered in this section.

10. *Preliminaries for proof of Theorem* 3

Consider the quantities $\mathcal{M}_k'(N,h,\psi_R,\rho,C,A)$, call them $\mathcal{M}_k'$ for brevity, defined as

$$\mathcal{M}_k' = \sum_{n=N+1}^{2N}(\psi_R(n+h)-\psi_R(n)-h-CA)^{k-1}(\psi(n+h)-\psi(n)-h-\rho A), \qquad (10.1)$$

where ρ and C are to be constants, and $A=o(h)$ will be chosen appropriately. The $'$ on $\mathcal{M}_k'$ (and on M_k', and $\mathcal{S}_k'$) signifies that the sum over n runs from $N+1$ to $2N$. Our main interest here is the case $k=3$. The Generalized Riemann Hypothesis will be assumed in this section.

For $k=1$ we have by (1.37),

$$\begin{aligned}
\mathcal{M}_1' &= \sum_{n=N+1}^{2N}(\psi(n+h)-\psi(n)-h-\rho A) \\
&= M_1'(N,h,\psi)-(h+\rho A)N \\
&= -\rho AN + O(N^{1/2}h\log^2 N), \qquad (10.2)
\end{aligned}$$

valid for $h \leqslant N$ and $N \to \infty$. The $k = 2$ case is more illuminating in that it helps us see the appropriate choice for A. We have

$$\begin{aligned}\mathcal{M}_2' &= \sum_{n=N+1}^{2N} (\psi_R(n+h) - \psi_R(n) - h - CA)(\psi(n+h) - \psi(n) - h - \rho A) \\ &= \tilde{M}_2'(N, h, \psi_R) - (h + \rho A) M_1'(N, h, \psi_R) \\ &\quad - (h + CA) M_1'(N, h, \psi) + (h + \rho A)(h + CA)N. \end{aligned} \tag{10.3}$$

Equations (1.28) and (4.1) give

$$M_1'(N, h, \psi_R) = \sum_{1 \leqslant j_1 \leqslant h} \mathcal{S}_1'(N, (j_1), (1)) = Nh + O(Rh). \tag{10.4}$$

By (1.38) we have

$$\tilde{M}_2'(N, h, \psi_R) = \mathcal{L}_1(R) \sum_{1 \leqslant j_1 \leqslant h} \tilde{\mathcal{S}}_1'(N, (j_1), (1)) + \sum_{\substack{1 \leqslant j_1, j_2 \leqslant h \\ \text{distinct}}} \tilde{\mathcal{S}}_2'(N, (j_1, j_2), (1, 1)) + O(RN^{\epsilon}). \tag{10.5}$$

Here the first sum is $M_1'(N, h, \psi)$ which we know. In the second sum we cannot use the evaluation (5.6) since the contribution of the error term coming from the Bombieri–Vinogradov theorem will be greater than the main term of $\mathcal{M}_2'$ when h is a power of N. We begin by proceeding similarly to (5.1)–(5.6), and we write $j = j_2 - j_1$, so that

$$\begin{aligned}
&\sum_{\substack{1 \leqslant j_1, j_2 \leqslant h \\ \text{distinct}}} \tilde{\mathcal{S}}_2'(N, (j_1, j_2), (1, 1)) \\
&\quad = N \sum_{\substack{1 \leqslant j_1, j_2 \leqslant h \\ \text{distinct}}} \sum_{r \leqslant R} \frac{\mu^2(r)}{\phi(r)} \sum_{\substack{d \mid r \\ (d, j_1 - j_2) = 1}} \frac{d\mu(d)}{\phi(d)} \\
&\qquad + O\left(\sum_{\substack{1 \leqslant j_1, j_2 \leqslant h \\ \text{distinct}}} \sum_{r \leqslant R} \frac{\mu^2(r)}{\phi(r)} \sum_{d \mid r} d \max_{x \leqslant 2N+h} |E(x; d, j_2 - j_1)| \right) \\
&\quad = N \sum_{\substack{1 \leqslant j_1, j_2 \leqslant h \\ \text{distinct}}} \left[\mathfrak{S}_2(j_2 - j_1) + O\left(\frac{j^* d(j^*)}{R\phi(j^*)} \right) \right] \\
&\qquad + O\left(\sum_{1 \leqslant j_1 \leqslant h} \sum_{r \leqslant R} \frac{\mu^2(r)}{\phi(r)} \sum_{d \mid r} d \sum_{1 \leqslant j \leqslant h-1} \max_{x \leqslant 2N+h} |E(x; d, j)| \right) \\
&\quad = 2N \sum_{1 \leqslant j \leqslant h-1} \sum_{1 \leqslant j_1 \leqslant h-j} \left[\mathfrak{S}_2(j) + O\left(\frac{jd(j)}{R\phi(j)} \right) \right] \\
&\qquad + O\left(\sum_{1 \leqslant j_1 \leqslant h} \sum_{r \leqslant R} \frac{\mu^2(r)}{\phi(r)} \sum_{d \mid r} d \left(\sum_{1 \leqslant j \leqslant h-1} 1 \right)^{1/2} \left(\sum_{1 \leqslant j \leqslant h-1} \max_{x \leqslant 2N+h} |E(x; d, j)|^2 \right)^{1/2} \right) \\
&\quad = 2N \sum_{1 \leqslant j \leqslant h} (h - j)\mathfrak{S}_2(j) + O\left(\frac{Nh^2 \log h}{R} \right) \\
&\qquad + O\left(h^{3/2} \sum_{r \leqslant R} \frac{\mu^2(r)}{\phi(r)} \sum_{d \mid r} d \left(\sum_{1 \leqslant j \leqslant h-1} \max_{x \leqslant 2N+h} |E(x; d, j)|^2 \right)^{1/2} \right).
\end{aligned} \tag{10.6}$$

For the calculation of the main term we know from [3] that

$$\sum_{1\leqslant j\leqslant h}(h-j)\mathfrak{S}_2(j)=\frac{h^2}{2}-\frac{h\log h}{2}+\frac{1-\gamma-\log 2\pi}{2}h+O(h^{1/2+\epsilon}). \tag{10.7}$$

The sum involving the $E(x;d,j)$ is $O((1+h/d)N\log^4 N)$ by Hooley's estimate (1.47) which depends on GRH, so that

$$\begin{aligned}\sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}}&\tilde{\mathcal{S}}_2'(N,(j_1,j_2),(1,1))\\ &=Nh^2+Nh(-\log h+1-\gamma-\log 2\pi)+O(Nh^{1/2+\epsilon})\\ &\quad+O\left(\frac{Nh^2\log h}{R}\right)+O(N^{1/2}h^{3/2}R\log^2 N)+O(N^{1/2}h^2R^{1/2}\log^2 N).\end{aligned} \tag{10.8}$$

Using (1.37) and (10.8) in (10.5) we have

$$\begin{aligned}\tilde{M}_2'(N,h,\psi_R)&=Nh^2+Nh(\mathcal{L}_1(R)-\log h+1-\gamma-\log 2\pi)+O(Nh^{1/2+\epsilon})+O\left(\frac{Nh^2\log h}{R}\right)\\ &\quad+O(N^{1/2}h\log^3 N)+O((N^{1/2}h^{3/2}R^{1/2}\log^2 N)(h^{1/2}+R^{1/2})),\end{aligned} \tag{10.9}$$

and plugging (10.9), (10.4), and (1.37) in (10.3) we obtain

$$\begin{aligned}\mathcal{M}_2'&=Nh(\mathcal{L}_1(R)-\log h+1-\gamma-\log 2\pi)+\rho CNA^2+O(Rh^2)+O(Nh^{1/2+\epsilon})\\ &\quad+O((N^{1/2}h^{3/2}R^{1/2}\log^2 N)(h^{1/2}+R^{1/2}))+O\left(\frac{Nh^2\log h}{R}\right)\\ &\quad+O(N^{1/2}h\log^2 N(h+\log N))+O(CAN^{1/2}h\log^2 N)+O(\rho ARh).\end{aligned} \tag{10.10}$$

Here the term ρCNA^2 will be of the same order of magnitude with the main term if we choose

$$A=(h\log N)^{1/2}. \tag{10.11}$$

Since (10.1) would be meaningful for $A=o(h)$, we must have $\log N=o(h)$. We need to have $R=o(N)$ so that (4.1) has an asymptotic interpretation, and $R\gg N^\epsilon$ so that $\mathcal{L}_1(R)$ is of the same order of magnitude with $\log N$. We also require that the error term $O((Nh^2\log h)/R)$ is smaller than Nh, and this imposes $h\log h=o(R)$. With these in mind (10.10) reduces to

$$\begin{aligned}\mathcal{M}_2'&=Nh(\rho C\log N+\mathcal{L}_1(R)-\log h+1-\gamma-\log 2\pi)+O(Nh^{1/2+\epsilon})\\ &\quad+O\left(\frac{Nh^2\log h}{R}\right)+O(N^{1/2}h^{3/2}R\log^2 N).\end{aligned} \tag{10.12}$$

This will have asymptotic significance when the very last error term is smaller than Nh, that is, for

$$h^{1/2}R=o(N^{1/2}/\log^2 N), \tag{10.13}$$

which restricts us to

$$h=o(N^{1/3}/\log^2 N). \tag{10.14}$$

When all these conditions are met we have

$$\mathcal{M}_2'\sim Nh(\rho C\log N+\log(R/h)). \tag{10.15}$$

If we had $\psi(n+h)-\psi(n)-h=o((h\log n)^{1/2})$, then (10.15) with $C=0$ would imply that

$$\sum_{n=N+1}^{2N}(\psi_R(n+h)-\psi_R(n)-h)\sim-\frac{Nh^{1/2}}{\rho}\frac{\log(R/h)}{(\log N)^{1/2}}$$

for any fixed $\rho\neq 0$, which is absurd since the left-hand side does not depend on ρ. Thus we obtain

$$\max_{x\leqslant y\leqslant 2x}|\psi(y+h)-\psi(y)-h|\gg_\epsilon (h\log x)^{1/2} \tag{10.16}$$

for $1\leqslant h\leqslant x^{1/3-\epsilon}$, a result which is already implicit in previous works (for example [6]) involving the lower bound method depending on the second order correlations of the $\lambda_R(n)$.

11. *The third mixed moment for the proof of Theorem* 3

We now turn our attention to

$$\begin{aligned}\mathcal{M}_3'&=\sum_{n=N+1}^{2N}(\psi_R(n+h)-\psi_R(n)-h-CA)^2(\psi(n+h)-\psi(n)-h-\rho A)\\&=\tilde{M}_3'(N,h,\psi_R)-(h+\rho A)M_2'(N,h,\psi_R)-2(h+CA)\tilde{M}_2'(N,h,\psi_R)\\&\quad+2(h+CA)(h+\rho A)M_1'(N,h,\psi_R)+(h+CA)^2M_1'(N,h,\psi)\\&\quad-(h+\rho A)(h+CA)^2N.\end{aligned} \tag{11.1}$$

We now need to evaluate $M_2'(N,h,\psi_R)$ and $\tilde{M}_3'(N,h,\psi_R)$. By (1.28) we have

$$M_2'(N,h,\psi_R)=\sum_{1\leqslant j_1\leqslant h}\mathcal{S}_2'(N,(j_1),(2))+2\sum_{1\leqslant j_1<j_2\leqslant h}\mathcal{S}_2'(N,(j_1,j_2),(1,1)). \tag{11.2}$$

The first sum on the right is evaluated by (4.7) as

$$\sum_{1\leqslant j_1\leqslant h}\mathcal{S}_2'(N,(j_1),(2))=Nh\mathcal{L}_1(R)+O(hR^2).) \tag{11.3}$$

For the second sum, letting $j=j_2-j_1$, we have, by (4.9),

$$2\sum_{1\leqslant j_1<j_2\leqslant h}\mathcal{S}_2'(N,(j_1,j_2),(1,1))=2N\sum_{1\leqslant j\leqslant h}(h-j)\mathfrak{S}_2(j)+O\left(\frac{Nh^2\log h}{R}\right)+O(h^2R^2) \tag{11.4}$$

(it is more convenient to keep the sum of the $\mathfrak{S}_2$ as it is, by not using (10.7) until the end, and also to view the first line of (10.9), except for the term $Nh\mathcal{L}_1(R)$, in this manner). Hence we obtain

$$M_2'(N,h,\psi_R)=2N\sum_{1\leqslant j\leqslant h}(h-j)\mathfrak{S}_2(j)+Nh\mathcal{L}_1(R)+O\left(\frac{Nh^2\log h}{R}\right)+O(R^2h^2). \tag{11.5}$$

By (1.38) we have

$$\begin{aligned}\tilde{M}_3'(N,h,\psi_R)=(\mathcal{L}_1(R))^2&\sum_{1\leqslant j_1\leqslant h}\tilde{\mathcal{S}}_1'(N,(j_1),(1))+\sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}}\tilde{\mathcal{S}}_3'(N,(j_1,j_2),(2,1))\\&+2\mathcal{L}_1(R)\sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}}\tilde{\mathcal{S}}_2'(N,(j_1,j_2),(1,1))\\&+\sum_{\substack{1\leqslant j_1,j_2,j_3\leqslant h\\ \text{distinct}}}\tilde{\mathcal{S}}_3'(N,(j_1,j_2,j_3),(1,1,1))+O(RN^\epsilon).\end{aligned} \tag{11.6}$$

In calculating the sum of the $\tilde{\mathcal{S}}_3'(N,(j_1,j_2),(2,1))$ over j_1 and j_2 it is not suitable to use Theorem 2 because we will pick up an error term $O(Nh^2 \log\log^2 3j^*)$ arising from the log min term of (5.20), and this will be greater than the eventual main term due to cancellations of larger terms (note that j_1 and j_2 are not the same as in § 5). Therefore we start anew as

$$\begin{aligned}
&\sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}} \tilde{\mathcal{S}}_3'(N,(j_1,j_2),(2,1))\\
&\quad= \sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}} \sum_{n=N+1}^{2N} \lambda_R^2(n+j_1)\Lambda(n+j_2)\\
&\quad= \sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}} \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{d|r_1\\ e|r_2}} d\mu(d)e\mu(e) \sum_{\substack{n=N+1\\ n\equiv -j_1 (\mathrm{mod}\,[d,e])}}^{2N} \Lambda(n+j_2)\\
&\quad= \sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}} \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)}\\
&\qquad\times \left\{ \sum_{\substack{d|r_1\\ e|r_2\\ ([d,e],j_2-j_1)=1}} \frac{N}{\phi([d,e])} d\mu(d)e\mu(e) + O\left(\sum_{\substack{d|r_1\\ e|r_2}} de \max_{u\leqslant 2N+h} |E(u;[d,e],j_2-j_1)| \right) \right\}\\
&\quad= N \sum_{1\leqslant |j|\leqslant h} (h-|j|) \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{d|r_1\\ e|r_2\\ ([d,e],j)=1}} \frac{d\mu(d)e\mu(e)}{\phi([d,e])}\\
&\qquad+ O\left(h \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{d|r_1\\ e|r_2}} de \sum_{1\leqslant |j|\leqslant h} \max_{u\leqslant 2N+h} |E(u;[d,e],j)| \right).
\end{aligned} \tag{11.7}$$

The last error term is dealt with in the same way as in (10.6), by first applying the Cauchy–Schwarz inequality to the sum over j_1 and then using Hooley's GRH-dependent estimate (1.47), which makes it

$$\begin{aligned}
&\ll N^{1/2}h^{3/2}\log^2 N \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{d|r_1\\ e|r_2}} de \left(1 + \left(\frac{h}{[d,e]}\right)^{1/2}\right)\\
&\ll N^{1/2}h^{3/2}R^2\log^2 N + N^{1/2}h^2\log^2 N \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{d|r_1\\ e|r_2}} \frac{de}{([d,e])^{1/2}}.
\end{aligned}$$

The last sum is treated similarly to the calculations following (4.5), and with the notation used there the inner sums over d and e become

$$\begin{aligned}\sum_{\delta|(r_1,r_2)} \delta^{3/2} \sum_{d'|(r_1/\delta)} (d')^{1/2} \sum_{\substack{e'|(r_2/\delta)\\(e',d')=1}} (e')^{1/2} \\ &= \prod_{p|r_2}(1+\sqrt{p}) \sum_{\delta|(r_1,r_2)} \prod_{p|\delta} \frac{p^{3/2}}{1+\sqrt{p}} \sum_{d'|(r_1/\delta)} (d')^{1/2} \prod_{p|(r_2,d')} \frac{1}{1+\sqrt{p}} \\ &= \prod_{p|r_2}(1+\sqrt{p}) \sum_{\delta|(r_1,r_2)} \prod_{p|\delta} \frac{p^{3/2}}{1+\sqrt{p}} \prod_{p|((r_1,r_2)/\delta)} \frac{1+2\sqrt{p}}{1+\sqrt{p}} \prod_{p|(r_1/(r_1,r_2))} (1+\sqrt{p}) \\ &= \prod_{p|r_1}(1+\sqrt{p}) \prod_{p|r_2}(1+\sqrt{p}) \prod_{p|(r_1,r_2)} (1+2\sqrt{p}+p^{3/2}).\end{aligned}$$

Hence the sum we are concerned with is now expressed as

$$\sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \prod_{p|r_1}(1+\sqrt{p}) \prod_{p|r_2}(1+\sqrt{p}) \prod_{p|(r_1,r_2)} (1+2\sqrt{p}+p^{3/2}),$$

and writing $r_1 = a_1a_{12}$ and $r_2 = a_2a_{12}$ with $a_{12} = (r_1, r_2)$ we see that this becomes

$$\begin{aligned}\sideset{}{'}\sum_{\substack{a_1a_{12}\leqslant R\\ a_2a_{12}\leqslant R}} \mu^2(a_1)\mu^2(a_2)\mu^2(a_{12}) \prod_{p|a_1a_2} \frac{1}{\sqrt{p}-1} \prod_{p|a_{12}} \frac{1+2\sqrt{p}+p^{3/2}}{(p-1)^2} \\ &\ll \sum_{a_{12}\leqslant R} \prod_{p|a_{12}} \frac{1+2\sqrt{p}+p^{3/2}}{(p-1)^2} \left(\sum_{a\leqslant R/a_{12}} \prod_{p|a} \frac{1}{\sqrt{p}-1}\right)^2 \\ &\ll R \sum_{a_{12}\leqslant R} \prod_{p|a_{12}} \frac{1+2\sqrt{p}+p^{3/2}}{p(p-1)^2} \ll R.\end{aligned}$$

Thus the error term of (11.7) has been shown to be

$$\ll N^{1/2}h^{3/2}R^2\log^2 N + N^{1/2}h^2R\log^2 N. \tag{11.8}$$

We now treat the main term of (11.7), which is equal to

$$2N \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{b|r_1\\ e|r_2}} \frac{b\mu(b)e\mu(e)}{\phi([b,e])} \sum_{\substack{1\leqslant j\leqslant h\\ (j,[b,e])=1}} (h-j) \tag{11.9}$$

(we have re-named the dummy variable d in (11.7) as b to avoid confusion with the divisor function). Here the innermost sum is

$$\begin{aligned}\sum_{\substack{1\leqslant j\leqslant h\\ (j,[b,e])=1}} (h-j) &= \sum_{1\leqslant j\leqslant h} (h-j) \sum_{k|(j,[b,e])} \mu(k) = \sum_{k|[b,e]} \mu(k) \sum_{\substack{1\leqslant j\leqslant h\\ k|j}} (h-j) \\ &= \sum_{k|[b,e]} \mu(k) \left\{ h\left(\frac{h}{k}+O(1)\right) - k \sum_{1\leqslant j'\leqslant h/k} j' \right\} \\ &= \sum_{k|[b,e]} \mu(k) \left(\frac{h^2}{2k}+O(h)\right) \\ &= \frac{h^2}{2}\frac{\phi([b,e])}{[b,e]} + O(hd([b,e])).\end{aligned} \tag{11.10}$$

The main term of (11.10) plugged into (11.9) gives

$$Nh^2 \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{b|r_1\\ e|r_2}} \frac{b\mu(b)e\mu(e)}{[b,e]} = Nh^2\mathcal{L}_1(R) \tag{11.11}$$

where the last evaluation follows by virtue of (4.5)–(4.7). It remains to consider the contribution of the error term of (11.10) plugged into (11.9),

$$Nh \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{b|r_1\\ e|r_2}} \frac{bed([b,e])}{\phi([b,e])}. \tag{11.12}$$

Writing $\delta = (b,e)$, $b = b'\delta$, and $e = e'\delta$, we can express the inner sums over b and e as

$$\begin{aligned}
&\sum_{\delta|(r_1,r_2)} \frac{\delta^2 d(\delta)}{\phi(\delta)} \sum_{b'|(r_1/\delta)} \frac{b'd(b')}{\phi(b')} \sum_{\substack{e'|(r_2/\delta)\\ (e',b')=1}} \frac{e'd(e')}{\phi(e')}\\
&= \prod_{p|r_2} \frac{3p-1}{p-1} \sum_{\delta|(r_1,r_2)} \prod_{p|\delta} \frac{2p^2}{3p-1} \sum_{b'|(r_1/\delta)} \frac{b'd(b')}{\phi(b')} \prod_{p|(r_2,b')} \frac{p-1}{3p-1}\\
&= \prod_{p|r_2} \frac{3p-1}{p-1} \sum_{\delta|(r_1,r_2)} \prod_{p|\delta} \frac{2p^2}{3p-1} \prod_{p|((r_1,r_2)/\delta)} \frac{5p-1}{3p-1} \prod_{p|(r_1/(r_1,r_2))} \frac{3p-1}{p-1}\\
&= \prod_{p|(r_1/(r_1,r_2))} \frac{3p-1}{p-1} \prod_{p|(r_2/(r_1,r_2))} \frac{3p-1}{p-1} \prod_{p|(r_1,r_2)} \frac{2p^2+5p-1}{p-1}.
\end{aligned}$$

Hence (11.12) becomes

$$\begin{aligned}
&Nh \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \prod_{p|([r_1,r_2]/(r_1,r_2))} \frac{3p-1}{p-1} \prod_{p|(r_1,r_2)} \frac{2p^2+5p-1}{p-1}\\
&= Nh \sideset{}{'}\sum_{\substack{a_1a_{12}\leqslant R\\ a_2a_{12}\leqslant R}} \mu^2(a_1)\mu^2(a_2)\mu^2(a_{12}) \prod_{p|a_1a_2} \frac{3p-1}{(p-1)^2} \prod_{p|a_{12}} \frac{2p^2+5p-1}{(p-1)^3}\\
&\ll Nh \sum_{a_{12}\leqslant R} \prod_{p|a_{12}} \frac{2p^2+5p-1}{(p-1)^3} \left(\sum_{a\leqslant R/a_{12}} \prod_{p|a} \frac{3p-1}{(p-1)^2} \right)^2\\
&\ll Nh\log^6 N \sum_{a_{12}\leqslant R} \prod_{p|a_{12}} \frac{2p^2+5p-1}{(p-1)^3} \ll Nh\log^8 N.
\end{aligned} \tag{11.13}$$

This completes the evaluation begun in (11.7), giving

$$\begin{aligned}
\sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}} \tilde{\mathcal{S}}_3'(N,(j_1,j_2),(2,1)) &= Nh^2\mathcal{L}_1(R) + O(Nh\log^8 N)\\
&\quad + O(N^{1/2}h^{3/2}R^2\log^2 N) + O(N^{1/2}h^2R\log^2 N).
\end{aligned} \tag{11.14}$$

12. *Completion of the proof of Theorem* 3

The last quantity involving the correlations that remains to be considered for the mixed third moment is

$$\sum_{\substack{1\leqslant j_1,j_2,j_3\leqslant h\\ \text{distinct}}} \tilde{\mathcal{S}}_3'(N,(j_1,j_2,j_3),(1,1,1)) = \sum_{\substack{1\leqslant j_1,j_2,j_3\leqslant h\\ \text{distinct}}} \sum_{n=N+1}^{2N} \lambda_R(n+j_1)\lambda_R(n+j_2)\Lambda(n+j_3). \tag{12.1}$$

The inner sum here is the same as (5.7) except for a shift, and the innermost sum of (5.8) adapted to the present situation is

$$\sum_{\substack{n=N+1+j_3\\ n\equiv j_3-j_1 (\mathrm{mod}\, d)\\ n\equiv j_3-j_2 (\mathrm{mod}\, e)}}^{2N+j_3} \Lambda(n), \tag{12.2}$$

where the divisibility conditions, compatible only when $(d,e) \mid j_1-j_2$, can be combined as $n \equiv j (\mathrm{mod}\, [d,e])$ say. The last sum is equal to

$$[([d,e],j)=1]\frac{N}{\phi([d,e])} + E(2N+j_3;[d,e],j) - E(N+1+j_3;[d,e],j). \tag{12.3}$$

We get the same main term as that of (5.11) except that the conditions now read $(d,e) \mid j_1-j_2$, $(d,j_3-j_1)=1$, and $(e,j_3-j_2)=1$, and the calculation carried out in § 5 evaluates it as

$$N\mathfrak{S}((j_1,j_2,j_3)) + O(NR^{-1+\epsilon}). \tag{12.4}$$

We shall take up the contribution of (12.4) in (12.1) after considering the contribution of the error terms of (12.3) to (12.1) which can be majorized as

$$\ll \sum_{\substack{1\leqslant j_1,j_2,j_3\leqslant h\\ \text{distinct}}} \sum_{r_1,r_2\leqslant R} \frac{\mu^2(r_1)\mu^2(r_2)}{\phi(r_1)\phi(r_2)} \sum_{\substack{d\mid r_1\\ e\mid r_2\\ (d,e)\mid j_1-j_2}} de \max_{u\leqslant 2N+h} |E(u;[d,e],j)|. \tag{12.5}$$

To simplify the expressions we may interchange the order of the two double sums in (12.5) as in the first line of (5.10). This switching costs a factor of $\log^2 R$, which is unimportant in our application. (On two occasions in § 11 we did not resort to this interchange.) Since j is a function of $k_1 = j_3 - j_1$ and $k_2 = j_3 - j_2$ (and also of d and e), we can rearrange the summations so as to see that (12.5) is

$$\ll \log^2 R \sum_{1\leqslant j_3\leqslant h} \sum_{d,\, e\leqslant R} \frac{\mu^2(d)\mu^2(e)de}{\phi(d)\phi(e)} \sum_{\substack{j_3-h\leqslant k_1,\, k_2\leqslant j_3-1\\ (k_2-k_1)k_1k_2\neq 0\\ (d,e)\mid k_2-k_1}} \max_{u\leqslant 2N+h} |E(u;[d,e],j)|. \tag{12.6}$$

Note that if $(d,e) \geqslant h$, then the innermost sum is void. The number of permissible pairs k_1, k_2 is

$$\ll h\left(\left\lceil \frac{h}{(d,e)}\right\rceil - 1\right).$$

Upon applying the Cauchy–Schwarz inequality and Hooley's GRH-dependent estimate (1.47) we find that (12.6) is

$$\ll h\log^2 R \sum_{\substack{d,\,e\leqslant R\\ (d,e)<h}} \frac{\mu^2(d)\mu^2(e)de}{\phi(d)\phi(e)} \frac{h}{\sqrt{(d,e)}} \left[\left(1+\frac{h^2}{de}\right) \sum_{j\,(\mathrm{mod}\,[d,e])} \max_{u\leqslant 2N+h} |E(u;[d,e],j)|^2\right]^{1/2}$$
$$\ll N^{1/2}h^2\log^4 N \sum_{\substack{d,\,e\leqslant R\\ (d,e)<h}} \frac{\mu^2(d)\mu^2(e)de}{\phi(d)\phi(e)\sqrt{(d,e)}} + N^{1/2}h^3\log^4 N \sum_{\substack{d,\,e\leqslant R\\ (d,e)<h}} \frac{\mu^2(d)\mu^2(e)\sqrt{de}}{\phi(d)\phi(e)\sqrt{(d,e)}}.$$

We have

$$\sum_{\substack{d,\,e\leqslant R\\ (d,e)<h}} \frac{\mu^2(d)\mu^2(e)de}{\phi(d)\phi(e)\sqrt{(d,e)}} \ll \sum_{\delta<h} \frac{\delta^{3/2}}{\phi^2(\delta)} \left(\sum_{d'\leqslant R/\delta} \frac{d'}{\phi(d')}\right)^2 \ll R^2 \sum_{\delta<h} \frac{1}{\delta^{1/2}\phi^2(\delta)} \ll R^2,$$
$$\sum_{\substack{d,\,e\leqslant R\\ (d,e)<h}} \frac{\mu^2(d)\mu^2(e)\sqrt{de}}{\phi(d)\phi(e)\sqrt{(d,e)}} \ll \sum_{\delta<h} \frac{\delta^{1/2}}{\phi^2(\delta)} \left(\sum_{d'\leqslant R/\delta} \frac{\sqrt{d'}}{\phi(d')}\right)^2 \ll R \sum_{\delta<h} \frac{1}{\delta^{1/2}\phi^2(\delta)} \ll R.$$

Hence the contribution of the error terms in (12.3) to (12.1) is

$$\ll N^{1/2}h^2R^2\log^4 N + N^{1/2}h^3R\log^4 N. \tag{12.7}$$

Now we calculate the contribution of (12.4) to (12.1), that of the error term being $O(Nh^3R^{-1+\epsilon})$. Inverting (1.24) we have

$$\mathfrak{S}((j_1,j_2,j_3)) = \sum_{\mathcal{J}\subset\{j_1,j_2,j_3\}} \mathfrak{U}(\mathcal{J}), \tag{12.8}$$

which implies by (1.23) that (since $\mathfrak{U}(\emptyset)=1$, $\mathfrak{U}((k))=0$, and $\mathfrak{U}((k,l))=\mathfrak{S}((k,l))-1$ for $k\neq l$)

$$\begin{aligned}\sum_{\substack{1\leqslant j_1,j_2,j_3\leqslant h\\ \text{distinct}}} \mathfrak{S}((j_1,j_2,j_3)) &= -2h(h-1)(h-2) + 3(h-2) \sum_{\substack{1\leqslant j_1,j_2\leqslant h\\ \text{distinct}}} \mathfrak{S}((j_1,j_2)) + R_3(h)\\ &= -2h(h-1)(h-2) + 6(h-2) \sum_{1\leqslant j\leqslant h} (h-j)\mathfrak{S}_2(j) + R_3(h). \end{aligned} \tag{12.9}$$

From (10.7) and (1.26) we obtain

$$\sum_{\substack{1\leqslant j_1,j_2,j_3\leqslant h\\ \text{distinct}}} \mathfrak{S}((j_1,j_2,j_3)) = h^3 - 3h^2\log h + 3(1-\gamma-\log 2\pi)h^2 + O(h^{3/2+\epsilon}) \tag{12.10}$$

(note that the full force of (1.26) has not been used; here we only need to know that $R_3(h) \ll h^{3/2+\epsilon}$). Hence we have

$$\begin{aligned}\sum_{\substack{1\leqslant j_1,j_2,j_3\leqslant h\\ \text{distinct}}} &\tilde{\mathcal{S}}_3'(N,(j_1,j_2,j_3),(1,1,1))\\ &= Nh^3 - 3Nh^2\log h + 3(1-\gamma-\log 2\pi)Nh^2 + O(Nh^{3/2+\epsilon})\\ &\quad + O(Nh^3R^{-1+\epsilon}) + O(N^{1/2}h^2R\log^4 N(R+h)). \end{aligned} \tag{12.11}$$

Combining (1.37), (10.8), (11.14) and (12.11) in (11.6) we obtain

$$\begin{aligned}\tilde{M}_3'(N,h,\psi_R) = Nh^3 &+ 3Nh^2(\mathcal{L}_1(R) - \log h) + 3Nh^2(1-\gamma-\log 2\pi)\\ &+ O(Nh^{3/2+\epsilon}) + O(Nh^3R^{-1+\epsilon}) + O(Nh\log^8 N)\\ &+ O(N^{1/2}h^2R^2\log^4 N) + O(N^{1/2}h^3R\log^4 N). \end{aligned} \tag{12.12}$$

Now as we compile our findings in (11.1), keeping the sums $\sum_{1\leqslant j\leqslant h}(h-j)\mathfrak{S}_2(j)$ (which appear in several terms) unevaluated until the end not only facilitates the calculation, but also reveals the complete cancellation of the terms which contain $Nh\sum_{1\leqslant j\leqslant h}(h-j)\mathfrak{S}_2(j)$. We get

$$\begin{aligned}\mathcal{M}_3' = (2C+\rho)ANh&[\log h - \mathcal{L}_1(R) - (1-\gamma-\log 2\pi)] - \rho C^2A^3N\\ &+ O(ANh^{1/2+\epsilon}) + O(N^{1/2}h^2R^2\log^4 N) + O(N^{1/2}h^3R\log^4 N)\\ &+ O(Nh\log^8 N) + O(h^3R^2) + O(Nh^3R^{-1+\epsilon}) + NR_3(h). \end{aligned} \tag{12.13}$$

The main terms are the same order of magnitude if $A=(h\log N)^{1/2}$ as before in (10.11) for $\mathcal{M}_2'$. This choice of A makes (12.13) read as

$$\begin{aligned}\mathcal{M}_3' = Nh^{3/2}\log^{1/2} N&[-\rho C^2\log N + (2C+\rho)(\log h - \mathcal{L}_1(R) - (1-\gamma-\log 2\pi)]\\ &+ \text{ error terms of (12.13)}. \end{aligned} \tag{12.14}$$

We are assuming that $R \gg N^{\epsilon}$, and the requirement that the error terms are smaller than the main term, that is, $o(Nh^{3/2}\log^{3/2} N)$, brings the restrictions

$$h \ll R^{2/3-\epsilon}, \qquad \log^{13} N = o(h), \qquad h^{1/2}R^2 = o(N^{1/2}\log^{-5/2} N). \tag{12.15}$$

Note that the cancellation mentioned before (12.13) is essential in reaching a result, for if (12.12) and (10.9) which depend on the evaluation (10.7) had been used, then we would have acquired an error term $O(Nh^{3/2+\epsilon})$ that is larger than the main term. Upon this we need the estimate (1.26) of Montgomery and Soundararajan for $R_3(h)$. Thus for

$$\log^{14} N \ll h \ll N^{1/7-\epsilon} \tag{12.16}$$

we have the asymptotic result

$$\mathcal{M}_3' \sim -Nh^{3/2}\log^{1/2} N(\rho C^2\log N + (2C+\rho)\log(R/h)). \tag{12.17}$$

(The factor N^{ϵ} in (12.16) can be replaced by a small power of $\log N$ if one bounds (5.38) more precisely as was remarked). From (12.17) we can get the result (10.16), only this time for the smaller range (12.16). The significance of (12.17) is that it allows us also to get a result of the type (10.16) without the absolute value. It is convenient to write $R = N^{\theta}$ and $h = N^{\alpha}$. For a fixed ρ satisfying $0 < |\rho| < \sqrt{\theta-\alpha}$, with the choice $C = -(\theta-\alpha)/\rho$, (12.17) reads

$$\mathcal{M}_3' \sim -\rho(\theta-\alpha)\left(1 - \frac{\theta-\alpha}{\rho^2}\right)Nh^{3/2}\log^{3/2} N. \tag{12.18}$$

We see that $\mathcal{M}_3'$ is positive for $0<\rho<\sqrt{\theta-\alpha}$, and negative for $-\sqrt{\theta-\alpha}<\rho<0$, and $\gg N(h\log N)^{3/2}$ in either case. This means that given an arbitrarily small but fixed $\eta>0$, for all sufficiently large N and h subject to (12.16), there exist $n_1, n_2 \in [N+1, 2N]$ such that

$$\begin{aligned}\psi(n_1+h) - \psi(n_1) - h &> \left(\tfrac{1}{2}\sqrt{1-5\alpha} - \eta\right)(h\log N)^{1/2},\\ \psi(n_2+h) - \psi(n_2) - h &< -\left(\tfrac{1}{2}\sqrt{1-5\alpha} - \eta\right)(h\log N)^{1/2}. \end{aligned} \tag{12.19}$$

References

1. H. Davenport, *Multiplicative number theory*, 3rd edn, revised and with a preface by H. L. Montgomery (Springer, New York, 2000).

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

3. D. A. Goldston, 'Linnik's theorem on Goldbach numbers in short intervals', *Glasgow Math. J.* 32 (1990) 285–297.

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

5. D. A. Goldston and H. L. Montgomery, 'Pair correlation of zeros and primes in short intervals', *Analytic number theory and Diophantine problems* (ed. A. C. Adolphson, J. B. Conrey, A. Ghosh and R. I. Yager; Birkhäuser, Boston, MA, 1987) 183–203.

6. D. A. Goldston and C. Y. Yildirim, 'On the second moment for primes in an arithmetic progression', *Acta Arith.* C.1 (2001) 85–104.

7. D. A. Goldston and C. Y. Yildirim, 'Higher correlations of divisor sums related to primes I: triple correlations', *Integers* 3 (2003) A5, 66 pp. (electronic).

8. 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.

9. A. Hildebrand, 'Über die punktweise Konvergenz von Ramanujan-Entwicklungen zahlentheoretischer Funktionen', *Acta Arith.* 44 (1984) 109–140.

10. H. Halberstam and H. -E. Richert, 'On a result of R. R. Hall', *J. Number Theory* 11 (1979) 76–89.

11. D. R. Heath-Brown, 'The ternary Goldbach problem', *Rev. Mat. Iberoamericana* 1 (1985) 45–59.

12. C. Hooley, 'On the Barban–Davenport–Halberstam theorem VI', *J. London Math. Soc.* (2) 13 (1976) 57–64.

13. C. Hooley, 'On the Barban–Davenport–Halberstam theorem XII', *Number theory in progress*, Zakopane, 1997 (ed. K. Győry, H. Iwaniec and J. Urbanowicz), Vol. II (Walter de Gruyter, Berlin, 1999) 893–910.

14. C. Hooley, 'On the Barban–Davenport–Halberstam theorem XIII', *Acta Arith.* 94 (2000) 53–86.

15. A. A. Karatsuba and S. M. Voronin, *The Riemann zeta-function* (translated from the Russian by N. Koblitz; Walter de Gruyter, Berlin, 1992).

16. H. L. Montgomery and K. Soundararajan, 'Primes in short intervals', *Comm. Math. Phys.* 252 (2004) 589–617.

17. A. Selberg, 'On an elementary method in the theory of primes', *Norske Vid. Selsk. Forh. Trondhjem* 19 (1947) 64–67.

18. E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd edn, edited and with a preface by D. R. Heath-Brown (The Clarendon Press, Oxford University Press, New York, 1986).

D. A. Goldston
Department of Mathematics
San Jose State University
San Jose, CA 95192
USA

goldston@math.sjsu.edu

C. Y. Yıldırım
Department of Mathematics
Boğaziçi University
Bebek
Istanbul, 34342
Turkey
and
Feza Gürsey Enstitüsü
Çengelköy
Istanbul, P.K. 6, 81220
Turkey

yalciny@boun.edu.tr