SMALL GAPS BETWEEN PRIMES II (PRELIMINARY)

D. A. GOLDSTON, J. PINTZ, AND C. Y. YILDIRIM

Abstract. We examine an idea for approximating prime tuples.

1. Statement of results (Preliminary)

In the present work we will prove the following result. Let p_n denote the nth prime. Then

$$\liminf_{n\to\infty} \frac{(p_{n+1}-p_n)}{\log p_n (\log\log p_n)^{-1} \log\log\log\log p_n} < \infty. \tag{1.1}$$

Further we show that supposing the validity of the Bombieri–Vinogradov theorem up to $Q \le X^\vartheta$ with any level $\vartheta > 1/2$ we have bounded differences between consecutive primes infinitely often:

$$\liminf_{n\to\infty} (p_{n+1}-p_n) \le C(\vartheta) \tag{1.2}$$

with a constant $C(\vartheta)$ depending only on ϑ. If the Bombieri–Vinogradov theorem holds with a level $\vartheta > 20/21$, in particular if the Elliott–Halberstam conjecture holds, then we obtain

$$\liminf_{n\to\infty} (p_{n+1}-p_n) \le 20, \tag{1.3}$$

that is $p_{n+1}-p_n \le 20$ for infinitely many n.

Inequalities (1.2)–(1.3) will follow from the even stronger following result

Theorem A. *Suppose the Bombieri–Vinogradov theorem is true for $Q \le X^\vartheta$ with some $\vartheta > 1/2$. Then there exists a constant $C'(\vartheta)$ such that any admissible k-tuple contains at least two primes for any*

$$k \ge C'(\vartheta) \quad \text{if} \quad \vartheta > 1/2, \tag{1.4}$$

where $C'(\vartheta)$ is an explicitly calculable constant depending only on ϑ. Further we have at least two primes for

$$k = 7 \quad \text{if} \quad \vartheta > 20/21. \tag{1.5}$$

Remark. For the definition of admissibility see (2.2) below.

We will show some more general results for the quantity (ν is a given positive integer)

$$E_\nu = \liminf_{n\to\infty} \frac{p_{n+\nu}-p_n}{\log p_n}. \tag{1.6}$$

Date: February 8, 2005.

1991 *Mathematics Subject Classification.* Primary 11N05; Secondary 11P32.

Key words and phrases. prime number.

The first author was supported in part by an NSF Grant, the third author by TÜBITAK.

Bombieri and Davenport [1] showed $E_\nu \le \nu - 1/2$, which was later improved by Huxley [10, 11], to $E_\nu \le \nu - 5/8 + o(1/\nu)$, by the first and third author [5] to $E_\nu \le (\sqrt{\nu} - 1/2)^2$, and by H. Maier to [12] $E_\nu \le e^{-\gamma}\left(\nu - \frac{5}{8} + o\left(\frac{1}{\nu}\right)\right)$. We can show in a relatively simple way, using our basic Theorems 1 and 2 from Section 2 the following

Theorem B. *Suppose the Bombieri–Vinogradov theorem is true for $Q \le X^\vartheta$ with all $\vartheta < \vartheta_0 \in [1/2, 1]$. Then we have for all $\nu \ge 1$*

$$E_\nu \le \max(\nu - 2\vartheta_0, 0). \tag{1.7}$$

Corollary 1. *We have unconditionally (with $\vartheta_0 = 1/2$)*

$$E_\nu \le \nu - 1, \tag{1.8}$$

in particular,

$$E_1 = \liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\log p_n} = 0. \tag{1.9}$$

Corollary 2. *If the Elliott–Halberstam conjecture is true, that is, we can choose $\vartheta_0 = 1$ in Theorem B, then for $\nu \ge 2$ we have*

$$E_\nu \le \nu - 2, \tag{1.10}$$

in particular,

$$E_2 = \liminf_{n\to\infty} \frac{p_{n+2} - p_n}{\log p_n} = 0. \tag{1.11}$$

In Section 3 we will show (following an idea of Granville and Soundararajan) how Theorems A and B follow from Theorems 1 and 2 in a simple way. (Theorems 1 and 2 will be proved in Sections 6 and 7, respectively.) However, a more complicated argument (see Section 12) will show that Theorems 1 and 2 in fact imply the following stronger results as well.

Theorem C. *Supposing the condition of Theorem B we have for $\nu \ge 2$*

$$E_\nu \le \left(\sqrt{\nu} - \sqrt{2\vartheta_0}\right)^2. \tag{1.12}$$

In particular we have unconditionally for $\nu \ge 1$

$$E_\nu \le \left(\sqrt{\nu} - 1\right)^2 \tag{1.13}$$

and under the Elliott–Halberstam conjecture for $\nu \ge 2$

$$E_\nu \le \left(\sqrt{\nu} - \sqrt{2}\right)^2. \tag{1.14}$$

We note that if we couple the ideas of the present work with H. Maier's well-known matrix method [12], then we can prove instead of (1.12) the stronger inequality

$$E_\nu \le e^{-\gamma}\left(\sqrt{\nu} - \sqrt{2\vartheta_0}\right)^2, \tag{1.15}$$

where γ is Euler's constant, with obvious improvements in (1.13)–(1.14). This will be the subject of a further paper in this series.

The work [1] of Bombieri and Davenport was generalized by Huxley [10] for primes in arithmetic progressions with a fixed modulus. A large part of the difficulties was due to the evaluation of the singular series. We mention that our present method, including the new treatment of the singular series in Section 10, allows for a far-reaching generalization of the situation of primes in an interval $[n+1, n+h]$ of length $h = [\lambda \log N]$. In fact we can prove that if $a_1, a_2, \ldots, a_h$ are arbitrary distinct integers in $[1, N]$, then there exists $n \in [N, 2N]$ such that at least two of the numbers $n+a_i$, $n+a_j$ are primes, if $h = [\lambda \log N]$, λ arbitrary positive constant, $N > N_0(\lambda)$ (or, alternatively, we obtain ν primes if $\lambda > (\sqrt{\nu} - 1)^2 + \varepsilon$).

We think that none of the previous methods (that is, the methods of Erdős, Bombieri–Davenport or Maier) would yield the above result with any fixed $\lambda < 1$.

The proof of the mentioned generalization will be also subject of another part of this series.

Finally we mention that the simpler diagonal method, used in the proofs of Theorems A and B may be refined to yield

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{(\log p_n)^{5/6} (\log\log p_n)^{1/2}} < \infty. \tag{1.16}$$

The proof of the above relation will be the subject of the next paper of this series.

This work is preliminary and requires further checking. This manuscript is intended for limited distribution.

2. Approximating prime tuples

Let

$$\mathcal{H} = \{h_1, h_2, \ldots, h_k\}, \quad \text{with } 1 \le h_1, h_2, \cdots, h_k \le h \ \text{ distinct integers}, \tag{2.1}$$

and let $\nu_p(\mathcal{H})$ denote the number of distinct residue classes modulo p occupied by the elements of $\mathcal{H}$. For squarefree integers d we extend this definition to $\nu_d(\mathcal{H})$ by multiplicativity. We next define the singular series

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right) \tag{2.2}$$

If $\mathfrak{S}(\mathcal{H}) \neq 0$ then $\mathcal{H}$ is called *admissible*. Thus $\mathcal{H}$ is admissible if and only if $\nu_p(\mathcal{H}) < p$ for all p.

A major unsolved problem in prime number theory is to prove that, assuming $\mathcal{H}$ is admissible, the tuple

$$(n + h_1, n + h_2, \ldots, n + h_k) \tag{2.3}$$

will have primes in every component for infinitely many n. Hardy and Littlewood made the quantitative conjecture that there is an asymptotic formula for the number of such tuples with $1 \le n \le N$. Let $\Lambda(n)$ denote the von Mangoldt function, and define

$$\Lambda(n; \mathcal{H}) = \Lambda(n + h_1)\Lambda(n + h_2) \cdots \Lambda(n + h_k). \tag{2.4}$$

This function detects prime tuples (including prime powers which can later be removed), and the Hardy–Littlewood prime tuple conjecture [7] states that for $\mathcal{H}$ admissible,

$$\sum_{n \le N} \Lambda(n; \mathcal{H}) = N\big(\mathfrak{S}(\mathcal{H}) + o(1)\big), \quad \text{as} \ \ N \to \infty. \tag{2.5}$$

(This is trivially true if $\mathcal{H}$ is not admissible.) Except for the prime number theorem (1-tuples), this conjecture is unproved and is very likely to remain so for a long time.

The program the first and last authors have been working on since 1999 is to compute approximations for (2.5) using short divisor sums and apply the results to problems on primes. The simplest approximation of $\Lambda(n)$ is based on the elementary formula

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d} \tag{2.6}$$

which we approximate with the smoothly truncated divisor sum

$$\Lambda_R(n) = \sum_{\substack{d|n \\ d \le R}} \mu(d) \log \frac{R}{d}. \tag{2.7}$$

Then our approximation for $\Lambda(n;\mathcal{H})$ is

$$\Lambda_R(n+h_1)\Lambda_R(n+h_2)\cdots\Lambda_R(n+h_k). \tag{2.8}$$

In the first paper in this series we applied this approximation to find small gaps between primes. During the course of that work we realized that for some applications there might be much better approximations for prime tuples than (2.8), but the approximation we devised was ultimately unsuccessful. Recently, however, we were able to obtain such an approximation, and in this paper we apply this to the problem of small gaps between primes.

The idea for our new approximation came from a paper of Heath-Brown [8] concerned with almost prime tuples. Heath-Brown's result is itself a generalization of Selberg's proof that the polynomial $n(n+2)$ will infinitely often have at most 5 prime factors, and thus the same is true for the pair $(n, n+2)$. We consider in connection with the tuple in (2.3) the polynomial

$$P_{\mathcal{H}}(n) = (n+h_1)(n+h_2)\cdots(n+h_k) \tag{2.9}$$

and note that the tuple (2.3) will be a prime tuple if and only if $P_{\mathcal{H}}(n)$ has exactly k prime factors. However, instead of detecting prime-tuples we will consider more generally almost prime tuples where the number of distinct prime factors of $P_{\mathcal{H}}(n)$ will be $k+\ell$ with $\ell = o(k)$. This means most of $n+h_i$'s are prime, some of them may be almost primes in the above sense. We detect this condition by using the $(k+\ell)^{\text{th}}$ generalized von Mangoldt function

$$\Lambda_{k+\ell}(n) = \sum_{d|n} \mu(d) \left(\log \frac{n}{d}\right)^{k+\ell} \tag{2.10}$$

which is zero if n has more than $k+\ell$ distinct prime factors. In analogy with (2.7), which is the case $k=1$, $\ell=0$ here, we approximate this by the smoothed truncated divisor sum

$$\sum_{\substack{d|n \\ d \le R}} \mu(d) \left(\log \frac{R}{d}\right)^{k+\ell}.$$

Our almost prime tuple detecting function is

$$\frac{1}{(k+\ell)!}\Lambda_{k+\ell}(P_{\mathcal{H}}(n)), \tag{2.11}$$

where the normalization by $\frac{1}{(k+\ell)!}$ simplifies the statement of our results. We next define the almost prime tuple approximation for a set $\mathcal{H}$ of size k

$$\Lambda_R(n;\mathcal{H},\ell) = \frac{1}{(k+\ell)!} \sum_{\substack{d|P_{\mathcal{H}}(n) \\ d\le R}} \mu(d) \left(\log \frac{R}{d}\right)^{k+\ell}. \tag{2.12}$$

(In [5] we had $\Lambda_R(n;\mathcal{H})$ given by (2.8)). As we will see in the next section, in the special case $\ell = 0$ this approximation suggests the Hardy–Littlewood type conjecture

$$\sum_{n\le N} \Lambda_k(P_{\mathcal{H}}(n)) = N\left(\mathfrak{S}(\mathcal{H}) + o(1)\right), \tag{2.13}$$

and a similar one if $\ell > 0$. While this conjecture and the Hardy–Littlewood conjecture have about the same content, the approximations for each of them are different. If in the sum in (2.12) we restrict ourselves to d's with all prime factors larger than h, then the condition $d|P_{\mathcal{H}}(n)$ implies that we can write $d = d_1d_2\cdots d_k$ uniquely with $d_i|n+h_i$, $1\le i\le k$, the d_i's pairwise relatively prime, and $d_1d_2\cdots d_k \le R$. In our application to prime gaps we require $R \le N^{\frac14-\varepsilon}$. On the other hand, the previous approximation (2.8) when multiplied out gives a sum over $d_i|n+h_i$, $1\le i\le k$, with $d_1 \le R$, $d_2 \le R$, …, $d_k \le R$. The application to prime gaps here requires that $R^k \le N^{\frac14-\varepsilon}$, so that $R \le N^{\frac{1}{4k}-\frac{\varepsilon}{k}}$. Thus we see that the earlier approximation has a more severe restriction on the range of the divisors. An additional technical advantage is that having one truncation rather than k truncations makes our calculations much easier.

Our main results on $\Lambda_R(n;\mathcal{H},\ell)$ are contained in the following two theorems. Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are both sets of k_1 and k_2 distinct positive integers, respectively, that are $\le h$. We always assume that at least one of these sets is non-empty. Let $M = k_1 + k_2 + \ell_1 + \ell_2$.

Theorem 1. *Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, $|\mathcal{H}_i| = k_i$, and $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$. If $R \le \frac{N^{\frac12}}{(\log N)^{4M}}$ and $h \le R^C$ for any given constant $C > 0$, then we have for $R, N \to \infty$,*

$$\sum_{n\le N} \Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2) = \binom{\ell_1+\ell_2}{\ell_1} \frac{(\log R)^{r+\ell_1+\ell_2}}{(r+\ell_1+\ell_2)!}(\mathfrak{S}(\mathcal{H}) + o_M(1))N. \tag{2.14}$$

In the following we will use the notation

$$\vartheta(n) = \begin{cases} \log n; & \text{if } n \text{ prime} \\ 0, & \text{otherwise} \end{cases}. \tag{2.15}$$

Theorem 2. *Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$, $|\mathcal{H}_i| = k_i$, $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$, $1 \le h_0 \le h$, and $\mathcal{H}^0 = \mathcal{H} \cup \{h_0\}$. If $R \ll_M N^{\frac14}(\log N)^{-B(M)}$ for a sufficiently large positive constant*

$B(M)$, and $h \le R$, then we have for $R, N \to \infty$,

(2.16)
$$\sum_{n\le N} \Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\vartheta(n+h_0)$$
$$= \begin{cases} \binom{\ell_1+\ell_2}{\ell_1}\dfrac{(\log R)^{r+\ell_1+\ell_2}}{(r+\ell_1+\ell_2)!}(\mathfrak{S}(\mathcal{H}^0)+o_M(1))N, & \text{if } h_0 \notin \mathcal{H}; \\ \binom{\ell_1+\ell_2+1}{\ell_1+1}\dfrac{(\log R)^{r+\ell_1+\ell_2+1}}{(r+\ell_1+\ell_2+1)!}(\mathfrak{S}(\mathcal{H})+o_M(1))N, & \text{if } h_0 \in \mathcal{H}_1 \text{ and } h_0 \notin \mathcal{H}_2; \\ \binom{\ell_1+\ell_2+2}{\ell_1+1}\dfrac{(\log R)^{r+\ell_1+\ell_2+1}}{(r+\ell_1+\ell_2+1)!}(\mathfrak{S}(\mathcal{H})+o_M(1))N, & \text{if } h_0 \in \mathcal{H}_1 \cap \mathcal{H}_2. \end{cases}$$

Assuming the Elliott–Halberstam conjecture, then equation (2.16) *holds for $R \ll_M N^{\frac12-\varepsilon}$ and $h \le R^\varepsilon$, with any $\varepsilon > 0$.*

Remark. By relabeling the variables we obtain the corresponding form if $h_0 \in \mathcal{H}_2, h_0 \notin \mathcal{H}_1$.

For applications, we apply Theorems 1 and 2 to evaluate the weighted average

(2.17)
$$S_R(N,K,\ell,\nu) = \frac{1}{Nh^{2K+1}}\sum_{n=N+1}^{2N}\left(\sum_{p=n+1}^{n+h}\log p - \nu\log 3N\right)(\psi_R(K,\ell,n,h))^2,$$

where
$$\psi_R(K,\ell,n,h) = \sum_{\substack{|\mathcal{H}|=K\\ h_i\in[1,h]}} \Lambda_R(n;\mathcal{H},\ell).$$

The positivity of S_R clearly implies $p_{j+\nu} - p_j \le h-1$ for some primes $p_j, p_{j+\nu}$ in the interval $[N, 2N+h]$.

Remark. Theorems 1 and 2 suffice to prove (even in the special case $\ell_1 = \ell_2$) Theorems A, B and C, including the relation

$$\liminf_{n\to\infty}\frac{p_{n+1}-p_n}{\log p_n} = 0, \qquad (2.18)$$

in an ineffective way. In order to show (1.1) (in an effective way) we need:

(i) a careful analysis of the error terms implicit in $o_M(1)$ with respect to their dependence on the variables k_i, ℓ_i, if $k_i = k_i(N) \to \infty$, $\ell_i = \ell_i(N) \to \infty$ (Sections 6 and 7);

(ii) Heath-Brown's well-known theorem that the existence of Siegel zeros implies that there are infinitely many twin primes (which is completely effective), or another alternative treatment of Siegel zeros (Sections 8 and 9);

(iii) modifications of the Bombieri–Vinogradov theorem (Sections 8 and 9);

(iv) new results about the behaviour of the singular series $\mathfrak{S}(\mathcal{H})$ in (2.2) if $k = k(h) \to \infty$ (Section 10).

The results including (i)–(iii) will appear as Theorems $1''$ and $2''$ at the end of Section 8. They can be considered as refinements of Theorems 1 and 2.

We will actually prove instead of Theorems 1 and 2 more precise forms as Theorems $1'$ and $2'$ (see Section 6) but they will not be applied to small gaps between primes.

In the following $c, C, c_i, C_i, c^*, \widetilde{c}$ will denote positive absolute constants whose value in case of c and C is not necessarily the same at each appearances. In general c and C will denote (sufficiently) small and (sufficiently) large constants, respectively. Constants implied by pure o, O, $\ll$ symbols will be absolute. We will use $\log_\nu x$ to denote the ν-fold iterated logarithm.

Although ϑ is used both for $\vartheta(n) = \log n$ if n is prime and for the level of the validity of the Bombieri–Vinogradov theorem, no confusion is possible.

Unless stated otherwise, we will consider sets $\mathcal{H}$ and $\mathcal{H}'$ different if the permutation of the same elements is different.

3. Proofs of Theorems A and B

In this section we will use a simple argument due to Granville and Soundararajan in a somewhat more general setting to show that Theorems A and B, including the crucial result (2.18) follow easily from Theorems 1 and 2. We will formulate the following

Hypothesis $BV(\vartheta, c)$**.** *Let $c > 0$ be a fixed positive constant, $Q = X^\vartheta \exp(-c\sqrt{\log X})$. Then for any $A > 0$ we have*

$$\sum_{q \le Q} E'(X, q) \ll_A \frac{X}{\log^A X}, \tag{3.1}$$

where (cf. (7.1)*)*

$$E'(X, q) = \max_{(a,q)=1} \left| \sum_{\substack{p \le X \\ p \equiv a(q)}} \log p - \frac{X}{\varphi(q)} \right|. \tag{3.2}$$

First we will show Theorem A of Section 1 supposing Theorems 1 and 2.

Proof of Theorem A. For $\ell \ge 0$, $\mathcal{H}_k = \{h_1, h_2, \dots, h_k\}$ we have for $R = N^{\vartheta/2} \exp(-c\sqrt{\log X}/2)$, from Theorem 1

$$\sum_{n \le N} \Lambda_R(n; \mathcal{H}_k, \ell)^2 \sim \frac{1}{(k+2\ell)!} \binom{2\ell}{\ell} \mathfrak{S}(\mathcal{H}_k) N (\log R)^{k+2\ell} \tag{3.3}$$

and with arbitrary $h_i \in \mathcal{H}_k$, from Theorem 2 by notation (2.15)

$$\sum_{n \le N} \Lambda_R(n; \mathcal{H}_k, \ell)^2 \vartheta(n + h_i) \sim \frac{2}{(k+2\ell+1)!} \binom{2\ell+1}{\ell} \mathfrak{S}(\mathcal{H}_k) N (\log R)^{k+2\ell+1}. \tag{3.4}$$

Now consider

(3.5)

$$\begin{aligned}\mathcal{S} := \sum_{n=N+1}^{2N}\left(\sum_{i=1}^{k}\vartheta(n+h_i) - \log 3N\right)\Lambda_R(n;\mathcal{H}_k,\ell)^2 \\ \sim k\frac{2}{(k+2\ell+1)!}\binom{2\ell+1}{\ell}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell+1} \\ - \log 3N\frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell} \\ \sim \left(\frac{2k}{k+2\ell+1}\frac{2\ell+1}{\ell+1}\log R - \log 3N\right)\frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell}.\end{aligned}$$

The tuple $\mathcal{H}_k$ will contain at least two primes if $\mathcal{S} > 0$ and this is true when

$$\frac{k}{k+2\ell+1}\frac{2\ell+1}{\ell+1}\vartheta > 1. \tag{3.6}$$

The original thoerem of Bombieri–Vinogradov, which has $\vartheta = \frac{1}{2}$, just misses if k and ℓ are taken large with $\ell < \varepsilon k$. But it is clear, this will be true for any $\vartheta > \frac{1}{2}$. Assuming the Elliott–Halberstam conjecture or even $\vartheta > 20/21$ we see this is true with $\ell = 1$ and $k = 7$. The admissible 7-tuple $\{1, 3, 7, 9, 13, 19, 21\}$ shows the truth of (1.3).

Proof of Theorem B. We modify the previous proof by now considering

$$\tilde{\mathcal{S}} := \sum_{n=N+1}^{2N}\left(\sum_{1\le h_0\le h}\vartheta(n+h_0) - \nu\log 3N\right)\sum_{\substack{1\le h_1,h_2,\ldots,h_k\le h\\ \text{distinct}}}\Lambda_R(n;\mathcal{H}_k,\ell)^2, \tag{3.7}$$

where our tuples $\mathcal{H}_k$ satisfy $1 \le h_1, h_2, \ldots, h_k \le h$. Since by Gallagher's theorem [3]

$$\sum_{\substack{1\le h_1,h_2,\ldots h_k\le h\\ \text{distinct}}}\mathfrak{S}(\mathcal{H}_k) \sim h^k, \tag{3.8}$$

we have by Theorems 1 and 2 for $R \le N^{\vartheta/2-\varepsilon}$,

(3.9)

$$\begin{aligned}\tilde{\mathcal{S}} \sim \sum_{\substack{1\le h_1,h_2,\ldots,h_k\le h\\ \text{distinct}}}\Bigg(k\frac{2}{(k+2\ell+1)!}\binom{2\ell+1}{\ell}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell+1} \\ + \sum_{\substack{1\le h_0\le h\\ h_0\ne h_i, 1\le i\le k}}\frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k\cup\{h_0\})N(\log R)^{k+2\ell} \\ - \nu\log 3N\frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H}_k)N(\log R)^{k+2\ell}\Bigg) \\ \sim \left(\frac{2k}{k+2\ell+1}\frac{2\ell+1}{\ell+1}\log R + h - \nu\log 3N\right)\frac{1}{(k+2\ell)!}\binom{2\ell}{\ell}Nh^k(\log R)^{k+2\ell}.\end{aligned}$$

Choosing $R = N^{\frac{\vartheta_0}{2}-\varepsilon}$, this implies there are at least ν primes in some interval $(n, n+h]$, $N < n \le 2N$ provided

$$h > \left(\nu - \frac{2k}{k+2\ell+1}\frac{2\ell+1}{\ell+1}\Big(\frac{\vartheta_0}{2} - \varepsilon\Big)\right)\log N \tag{3.10}$$

which on letting $\ell = [\sqrt{k}/2]$ and k sufficiently large proves the theorem, since the product of the two fractions above is $4 + O(1/\sqrt{k})$. Q.E.D.

Remark 3.1. As we can choose $\vartheta_0 = 1/2$ and the ε in the definition of R as small as $B\log_2 N/\log N$ in the unconditional case with the constant B appearing in Bombieri–Vinogradov's theorem, we immediately see that this proof leads to prime pairs in some intervals of length $h = \frac{c\log N}{\sqrt{k}}$ for any $c > 2$ (fixed absolute constant) if $k = k(N)$ tends to infinity in such a way that Theorems 1 and 2 still hold.

Remark 3.2. If we are looking for the minimal $k \ge 1$ such that there exists $\ell \ge 0$ with

$$\frac{k}{k+2\ell+1}\cdot\frac{2\ell+1}{\ell+1} > 1 \tag{3.11}$$

$(k, \ell \in \mathbb{Z})$ the answer is $k = 7$, $\ell = 1$.

Proof.

$$\begin{aligned}
\ell = 0:&\quad k > k+1 \quad \text{is impossible}\\
\ell = 1:&\quad 3k > 2(k+3) \Leftrightarrow k > 6\\
\ell = 2:&\quad 5k > 3(k+5) \Leftrightarrow k > 7.5\\
\ell \ge 3:&\quad 1 < \frac{k}{k+2\ell+1}\cdot\frac{2\ell+1}{\ell+1} < \frac{k}{k+7}\cdot 2 \Rightarrow k > 7.
\end{aligned}$$

Remark 3.3. If $h_i \in \mathbb{Z}$, $h_1 < h_2 < \cdots < h_7$, $\mathcal{H} = \{h_i\}_{i=1}^7$ is admissible, then $h_7 - h_1 \ge 20$. The inequality is sharp as shown by $\mathcal{H}_0 = \{11, 13, 17, 19, 23, 29, 31\}$, which is clearly "equivalent" with $\mathcal{H}_1 = \{1, 3, 7, 9, 13, 19, 21\}$.

Proof. (i) $\mathcal{H}_0$ is really admissible since for any prime $p \le 7$ none of the elements of $\mathcal{H}$ are divisible by p. For any $p > 7$, the 7 elements of $\mathcal{H}$ can clearly not cover all residue classes mod p.

(ii) Let $\mathcal{H}$ be any admissible system. Since translation does not change admissibility we can suppose $h_1 = 1$. This implies $2 \nmid h_i$.

Considering any block of 6 consecutive integers, it contains 3 odd numbers, each in different residue classes mod 3, so we can have at most 2 elements in any interval $I_k = [6k+1, 6k+6]$ for $k = 0, 1, 2$. We have two cases.

Case (ii)/1. At least one of these 3 intervals contains at most one h_i. Then we must have $h_6 \ge 19$, therefore as $2 \nmid h_i$, $h_7 \ge 21$, $h_7 - h_1 \ge 20$.

Case (ii)/2. All three intervals I_k contain exactly 2 elements. In this case we have either

Case A. $h_i \not\equiv 0 \pmod 3$ for all $i \le 7$ or

Case B. $h_i \not\equiv 2 \pmod 3$ for all $i \le 7$.

In either cases the 6 elements from $[1, 18]$ are completely determined, as $2 \nmid h_i$.

Case A. $\{h_i\}_{i=1}^6 = 1, 5, 7, 11, 13, 17$.

Case B. $\{h_i\}_{i=1}^6 = 1, 3, 7, 9, 13, 15$.

Here Case B is not admissible since it covers all residue classes mod 5.

Case A covers all residue classes except for 4 (mod 5), so $h_7 \neq 19$. Since $2 \nmid h_7$, $h_7 \geq 21$, $h_7 - h_1 \geq 20$. Q.E.D.

Remark 3.4. This argument implicitly shows that if $\mathcal{H} = \{h_i\}_{i=1}^6$ is admissible, then $h_6 - h_1 \geq 16$, as seen by Case A.

4. Lemmas

The Riemann zeta-function has the Euler product representation, with $s = \sigma + it$,

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}, \quad \sigma > 1. \tag{4.1}$$

The zeta-function is analytic except for a simple pole at $s = 1$ where we have as $s \to 1$,

$$\zeta(s) = \frac{1}{s-1} + \gamma + O(|s-1|). \tag{4.2}$$

We need standard information concerning the classical zero-free region of the Riemann zeta-function. By Theorem 3.11 and (3.11.8) of [16] there exists a small positive constant $\overline{c}$, for which we assume $\overline{c} \leq 10^{-3}$ such that $\zeta(\sigma + it) \neq 0$ in the region

$$\sigma \geq 1 - \frac{4\overline{c}}{\log(|t| + 3)} \tag{4.3}$$

for all t, and further

$$\begin{aligned} \zeta(\sigma + it) - \frac{1}{\sigma - 1 + it} \ll \log(|t| + 3), \quad \frac{1}{\zeta(\sigma + it)} \ll \log(|t| + 3), \\ \frac{\zeta'}{\zeta}(\sigma + it) + \frac{1}{\sigma - 1 + it} \ll \log(|t| + 3), \end{aligned} \tag{4.4}$$

in this region. We will fix this $\overline{c}$ for the rest of the paper (we could take $\overline{c} = 10^{-3}$), and let $\omega = e^{\sqrt{\overline{c}\log R}}$. Let $\mathcal{L}$ denote the contour given by

$$s = -\frac{\overline{c}}{\log(|t| + 3)} + it. \tag{4.5}$$

Lemma 1. *We have, for $R \geq C$, $k \geq 2$, $B \leq Ck$*

$$\int_{\mathcal{L}} (\log(|s| + 3))^B \left| \frac{R^s ds}{s^k} \right| \ll C_2^k R^{-c_3} + e^{-\sqrt{\overline{c}\log R}/2}, \tag{4.6}$$

where C_2, c_3 and the implied constant in $\ll$ depends only on the constant C in the formulation of the lemma.

Proof. The left-hand side of (4.6) is

$$\begin{aligned} &\ll \int_0^\infty R^{\sigma(t)} \frac{(\log(|t| + 3))^B}{(|t| + \overline{c})^k} \, dt \\ &\ll \int_0^{C_1} C_2^k R^{-c_3} dt + \int_{C_1}^{\omega - 3} \frac{R^{-\frac{\overline{c}}{\log(|t|+3)}}}{t^{3/2}} dt + \int_{\omega-3}^\infty t^{-3/2} dt \\ &\ll C_2^k R^{-c_3} + e^{-\frac{\overline{c}\log R}{\log \omega}} + \omega^{-\frac{1}{2}} \end{aligned} \tag{4.7}$$

on choosing $\log \omega = \sqrt{\overline{c}\log R}$. Q.E.D.

Corollary 3. *We have for $R \geq C$, $k \geq 2$, $B \leq Ck$, $k \leq c_5 \log R$ with a sufficiently small c_5 (depending on C, $\overline{c}$)*

$$\int_{\mathcal{L}} (\log(|s|+3))^B \left| \frac{R^s ds}{s^k} \right| \ll e^{-\sqrt{\overline{c}\log R}/2}. \tag{4.8}$$

Next we will show some explicit estimates for the sum of the generalized divisor function.

Let $\omega(q)$ denote the number of prime factors of a squarefree integer q, let m be first an integer. Then

$$d_m(q) = m^{\omega(q)}. \tag{4.9}$$

Let us denote the set of squarefree integers by $\mathbb{Z}_0$, the set of real numbers by $\mathbb{R}$. Using (4.9) as a definition, we can extend the generalized divisor function for positive non-integer values of m as well. The resulting function $d_m(q) = m^{\omega(q)}$ will be monotonically increasing as a function of m for fixed $q \in \mathbb{Z}_0$. In the following $\sum^\flat$ will always mean summation over squarefree integers.

Lemma 2. *For $m \in \mathbb{Z}$ we have*

$$D'(x,m) := \sum_{q \leq x}{}^{\flat} \frac{d_m(q)}{q} \leq (m + \log x)^m \quad \textit{for } x \geq 1. \tag{4.10}$$

Proof. We will show the result by induction. The assertion is true for $m = 1$, when $d_1(q) = 1$ by definition. Suppose (4.10) is proved for $m-1$. Let us denote the smallest term in a given product representation of q by $j = j(q) \leq x^{1/m}$. Then this factor can stand at m places, therefore we have for $q = q'j(q) = q'j$

$$\sum_{q \leq x}{}^{\flat} \frac{d_m(q)}{q} \leq m \sum_{j=1}^{x^{1/m}}{}^{\flat} \frac{1}{j} \sum_{q' \leq x/j}{}^{\flat} \frac{d_{m-1}(q')}{q'} \leq m \cdot (1 + \log x^{1/m})\,(m - 1 + \log x)^{m-1} \leq (m + \log x)(m + \log x)^{m-1}. \qquad \text{Q.E.D.} \tag{4.11}$$

Let us denote $\lceil y \rceil = \min\{n \in \mathbb{Z}; y \leq n\}$. Then Lemma 2 clearly implies

Lemma 3. *For any real m we have*

$$D'(x,m) \leq (\lceil m \rceil + \log x)^{\lceil m \rceil} \leq (m + 1 + \log x)^{m+1}.$$

Corollary 4. *We have for any real $m > 0$, $x \geq 1$*

$$D^*(x,m) := \sum_{q \leq x}{}^{\flat} d_m(q) \leq x(\lceil m \rceil + \log x)^{\lceil m \rceil} \leq x(m + 1 + \log x)^{m+1}.$$

Lemma 4. *For $q \in \mathbb{Z}_0$, $y, m_1, m_2 \in \mathbb{R}$*

$$d_{m_1}(q) d_{m_2}(q) = d_{m_1 m_2}(q), \qquad (d_m(q))^y = d_{m^y}(q). \tag{4.12}$$

Proof. Trivial by (4.9).

We will use Hölder's inequality later. For this reason let

$$\nu \geq c' \log(K+1), \quad \nu \geq 1. \tag{4.13}$$

The following lemma is valid with an absolute constant C' depending on c' in (4.13).

Lemma 5. $\sum_{q\le x}^{\flat} \dfrac{(d_{3K}(q))^{1+1/\nu}}{q} \le (C'K+\log x)^{C'K}$ *for* $x\ge 1,\ K\ge 1.$

Proof. By (4.9) we have

$$(d_{3K}(q))^{1+1/\nu} = d_j(q) \tag{4.14}$$

with

$$j=(3K)^{1+1/\nu} \le 9e^{1/c'}K. \tag{4.15}$$

Now Lemma 5 is true with $C' = 9e^{1/c'}+1$ by Lemma 3.

We shall prove below a combinatorial identity needed later in evaluating the residue in Section 6.

Let us define for non-negative integers d,u,v the quantity

$$S(d,u,v) := \frac{1}{u!}\sum_{i=0}^{u}\binom{u}{i}(-1)^i\frac{d(d+1)\dots(d+i-1)}{(v+d+i)!}, \tag{4.16}$$

where, as usual, the empty product (for $i=0$) means $1, 0!=1, \binom{0}{0}=1$.

Lemma 6. $S(d,u,v)=\binom{v+u}{u}\frac{1}{(d+u+v)!}$.

Proof. We will prove Lemma 6 for arbitrary values of d and v by induction on u. For $u=0$ the statement is trivially true for every d and v, so let us suppose $u\ge 1$ and that it is true for all d and v with $u-1$ in place of u. In this case we have by the identity $\binom{u}{i}=\binom{u-1}{i}+\binom{u-1}{i-1}$ (where we define for $i=0$ and u, $\binom{u-1}{u}=\binom{u-1}{-1}=0$):
(4.17)

$$\begin{aligned}
S(d,u,v) &= \frac{1}{u!}\sum_{i=0}^{u}\left(\binom{u-1}{i}+\binom{u-1}{i-1}\right)(-1)^i\frac{d(d+1)\dots(d+i-1)}{(v+d+i)!}\\
&= \frac{1}{u!}\Bigg\{\sum_{i=0}^{u-1}\binom{u-1}{i}(-1)^i\frac{d(d+1)\dots(d+i-1)}{(v+d+i)!}\\
&\qquad -\sum_{j=0}^{u-1}\binom{u-1}{j}(-1)^j\frac{d(d+1)\dots(d+j)}{(v+d+j+1)!}\Bigg\}\\
&= \frac{1}{u!}\sum_{i=0}^{u-1}\binom{u-1}{i}(-1)^i\frac{d(d+1)\dots(d+i-1)}{(v+d+i)!}\left(1-\frac{d+i}{v+d+i+1}\right)\\
&= \frac{v+1}{u}\cdot\frac{1}{(u-1)!}\sum_{i=0}^{u-1}\binom{u-1}{i}(-1)^i\frac{d(d+1)\dots(d+i-1)}{(v+1+d+i)!}\\
&= \frac{v+1}{u}S(d,u-1,v+1) = \frac{v+1}{u}\cdot\binom{v+u}{u-1}\cdot\frac{1}{(d+u+v)!} \qquad \text{Q.E.D.}
\end{aligned}$$

5. A special case of Theorem 1

We will first prove a special case of Theorem 1, which illustrates the method used to prove our results. We assume $\mathcal{H}$ is non-empty and thus $k\ge 1$. Further we take the simple case $\ell=0$. For $h\le R^C$, C any fixed positive number, we have now with $\Lambda_R(n;\mathcal{H},0)=\Lambda_R(n;\mathcal{H})$ for any

$$k \ll_{\eta_0} (\log R)^{1/2-\eta_0} \text{ with an arbitrary fixed } \eta_0>0 \tag{5.1}$$

the relation

$$\sum_{n=1}^{N}\Lambda_R(n;\mathcal{H}) = \mathfrak{S}(\mathcal{H})N + O(Ne^{-c\sqrt{\log R}}) + O\big(R(2\log R)^{2k}\big), \tag{5.2}$$

which is the basis for making the conjecture (2.13).

Proof. We have

$$\mathcal{S}_R(N;\mathcal{H}) := \sum_{n=1}^{N}\Lambda_R(n;\mathcal{H}) = \frac{1}{k!}\sum_{d\le R}\mu(d)\left(\log\frac{R}{d}\right)^k \sum_{\substack{1\le n\le N\\ d|P_{\mathcal{H}}(n)}} 1. \tag{5.3}$$

If for a prime p we have $p|P_{\mathcal{H}}(n)$ then among the solutions $n\equiv -h_i (\bmod\, p)$, $1\le i\le k$, there will be $\nu_p(\mathcal{H})$ distinct solutions modulo p. For d squarefree we then have by multiplicativity $\nu_d(\mathcal{H})$ distinct solutions for n modulo d which satisfy $d|P_{\mathcal{H}}(n)$, and for each solution one has n running through a residue class modulo d. Hence we see

$$\sum_{\substack{1\le n\le N\\ d|P_{\mathcal{H}}(n)}} 1 = \nu_d(\mathcal{H})\left(\frac{N}{d}+O(1)\right).$$

Since trivially for squarefree q, $\nu_q(\mathcal{H})\le k^{\omega(q)} = d_k(q)$, we conclude

(5.4)

$$\begin{aligned}\mathcal{S}_R(N;\mathcal{H}) &= N\left(\frac{1}{k!}\sum_{d\le R}\frac{\mu(d)\nu_d(\mathcal{H})}{d}\left(\log\frac{R}{d}\right)^k\right) + O\left(\frac{(\log R)^k}{k!}\sum_{d\le R}{}^{\flat}\,\nu_d(\mathcal{H})\right)\\ &= N\mathcal{T}_R(N;\mathcal{H}) + O\big(R(k+\log R)^{2k}\big),\end{aligned}$$

where we have made use of Corollary 2.

Let (a) denote the contour $s = a+it$, $-\infty<t<\infty$. We now apply the formula, for $c>0$,

$$\frac{1}{2\pi i}\int_{(c)}\frac{x^s}{s^{k+1}}\,ds = \begin{cases} 0, & \text{if } 0<x\le 1,\\ \frac{1}{k!}(\log x)^k, & \text{if } x\ge 1,\end{cases} \tag{5.5}$$

and have that

$$\mathcal{T}_R(N;\mathcal{H}) = \frac{1}{2\pi i}\int_{(1)} F(s)\frac{R^s}{s^{k+1}}ds, \tag{5.6}$$

where, letting $s=\sigma+it$ and assuming $\sigma>0$,

$$F(s) = \sum_{d=1}^{\infty}\frac{\mu(d)\nu_d(\mathcal{H})}{d^{1+s}} = \prod_p\Big(1-\frac{\nu_p(\mathcal{H})}{p^{1+s}}\Big). \tag{5.7}$$

Since here $\nu_p(\mathcal{H}) = k$ for all $p>h$, we see that we can write

$$F(s) = \frac{G_{\mathcal{H}}(s)}{\zeta(1+s)^k}, \tag{5.8}$$

where by (4.1)

$$G_{\mathcal{H}}(s) = \prod_p\left(1-\frac{\nu_p(\mathcal{H})}{p^{1+s}}\right)\left(1-\frac{1}{p^{1+s}}\right)^{-k} \tag{5.9}$$

is analytic and uniformly bounded for $\sigma > -\frac{1}{2}+\delta$ for any $\delta>0$. Also we see immediately from (2.2) that

$$G_{\mathcal{H}}(0) = \mathfrak{S}(\mathcal{H}). \tag{5.10}$$

From (5.8) and (4.4), $F(s)$ satisfies the bound in the region to right of $\mathcal{L}$

$$F(s) \ll |G_{\mathcal{H}}(s)|\big(C\log(2+|t|)\big)^k. \tag{5.11}$$

Here $G_{\mathcal{H}}(s)$ while analytic and bounded in this region has not only a dependence on k but also the size h of the components of $\mathcal{H}$. We see from (5.9) that

$$G_{\mathcal{H}}(s) = \prod_p \left(1 + \frac{k-\nu_p(\mathcal{H})}{p^{1+s}} + O\Big(\frac{k^2}{p^{2+2\sigma}}\Big)\right).$$

Here $\nu_p(\mathcal{H}) = k$ not only if $p > h$ but whenever $p \nmid \Delta$, where

$$\Delta := \prod_{1\le i<j\le k} |h_j - h_i| \le h^{k^2}, \qquad \log\Delta \le U := Ck^2\log(2h), \tag{5.12}$$

since then all k of the h_i's are distinct modulo p. Hence, for $-\frac{1}{100} < \sigma \le 1$, we have with $\delta = \max(-\sigma, 0)$,

(5.13)

$$\begin{aligned}
|G_{\mathcal{H}}(s)| &\le \prod_{p|\Delta \text{ or } p<2k^2} \left(1+\frac{k}{p^{1-\delta}}\right)\left(1+\frac{2}{p^{1-\delta}}\right)^k \prod_{p\nmid\Delta, p\ge 2k^2} \left|\left(1-\frac{k}{p^{1+s}}\right)\left(1-\frac{1}{p^{1+s}}\right)^{-k}\right| \\
&\le \exp\left(3k\sum_{p|\Delta \text{ or } p<2k^2}\frac{1}{p^{1-\delta}} + \sum_{p\nmid\Delta, p\ge 2k^2}\sum_{\nu\ge2}\frac{k^\nu}{p^{(1-\delta)\nu}}\right) \\
&\le \exp\left(3k\sum_{p\le U}\frac{1}{p^{1-\delta}} + \sum_{p>U}\sum_{\nu\ge2}\frac{k^\nu}{p^{(1-\delta)\nu}}\right) \\
&\le \exp\left(3kU^\delta\sum_{p\le U}\frac{1}{p} + 2k^2\sum_{n>U}\frac{1}{n^{2-2\delta}}\right) \\
&\ll \exp(4kU^\delta\log\log U),
\end{aligned}$$

where in the second line the expression has been majorized in such a way that we changed the primes between the first and the second sum by using the smallest possible set of primes that could divide Δ for the first sum, and put the other primes into the second sum; in this way we got the expression in the third line. This is allowed since for $p > U$ we have clearly

$$\frac{3k}{p^{1-\delta}} > \sum_{\nu=2}^{\infty}\left(\frac{k}{p^{1-\delta}}\right)^\nu.$$

The same argument will be used later several times without mentioning details. The conclusion is that, for $h \ll R^C$ where C is any fixed positive number as large as we wish, and for s on $\mathcal{L}$ or to the right of $\mathcal{L}$,

$$F(s) \ll (C\log(|t|+2))^k\exp(4kU^\delta\log\log U). \tag{5.14}$$

Returning to (5.6), we see the integrand vanishes as $|t| \to \infty$, $-1/100 < \sigma \le 1$. Moving the contour from (1) to the left to $\mathcal{L}$ by (5.8) we pass a simple pole at $s = 0$

and have by (4.2), (5.10), the argument used in Corollary 3 and (5.14), for any k satisfying (5.1)

$$\begin{aligned} \mathcal{T}_R(N;\mathcal{H}) &= G_{\mathcal{H}}(0) + \frac{1}{2\pi i}\int\limits_{\mathcal{L}} F(s)\frac{R^s}{s^{k+1}}ds \\ &= \mathfrak{S}(\mathcal{H}) + O(e^{-c\sqrt{\log R}}). \end{aligned} \tag{5.15}$$

Thus (5.2) follows from this and (5.4).

Remark. The exponent 1/2 in the restriction $k \ll (\log R)^{1/2-\eta_0}$ is not critical in any sense, since at other places we have stronger restrictions for k; further, using Vinogradov's zero-free region for $\zeta(s)$ we can replace 1/2 by 3/5 in the exponent.

6. Proof of Theorem 1

The proof here is similar to the corresponding result proved in [8]. We let

(6.1)
$$\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2,\ |\mathcal{H}_1| = k_1, \quad |\mathcal{H}_2| = k_2,\ k = k_1 + k_2,\ r = |\mathcal{H}_1 \cap \mathcal{H}_2|,\ M = k_1 + k_2 + \ell_1 + \ell_2.$$

Thus $|\mathcal{H}| = k - r$. We will show beyond Theorem 1 the following sharper form of it.

Theorem 1′. *We have for $h \ll R^C$, with any given fixed positive C, as $R, N \to \infty$,*

(6.2)
$$\begin{aligned} \sum_{n\le N} \Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2) = &\binom{\ell_1+\ell_2}{\ell_1}\frac{(\log R)^{r+\ell_1+\ell_2}}{(r+\ell_1+\ell_2)!}\mathfrak{S}(\mathcal{H})N \\ &+ N\sum_{j=1}^{r+\ell_1+\ell_2} \mathcal{D}_j(\mathcal{H}_1,\mathcal{H}_2)(\log R)^{r+\ell_1+\ell_2-j} \\ &+ O\left(Ne^{-c\sqrt{\log R}}\right) + O(R^2(3\log R)^{3k+M}), \end{aligned}$$

where the $\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2)$'s are functions independent of R and N which satisfy the bound

$$\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2) \ll_M (\log U)^{C_j} \ll_M (\log\log 10h)^{C'_j} \tag{6.3}$$

for some positive constants C_j, C'_j depending on M.

Remark. This result has no meaning if $M \to \infty$, due to the implicit dependence of the estimate (6.3) on M. Therefore we omitted the condition (5.1) (with M in place of k) in the formulation of Theorem 1′, which is actually necessary to obtain the error term $O(Ne^{-c\sqrt{\log R}})$ above.

As mentioned earlier in Section 2, Theorems 1 (and therefore the above refinement of Theorem 1) and 2 already suffice to show not only (2.18), but also Theorems A, B and C. On the other hand, to prove (1.1) we will need a stronger form of this, where the dependence on k_i, ℓ_i are made explicit. This result will be Theorem 1″ (respectively Theorem 2″) at the end of Section 8.

Proof. We can assume that both $\mathcal{H}_1$ and $\mathcal{H}_2$ are non-empty since the case where one of these sets is empty can be covered in the same way as we did in case of $\ell = 0$

in the last section. Thus $k \geq 2$ and we have

(6.4)
$$\begin{aligned}
&\mathcal{S}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2)\\
&:= \sum_{n=1}^{N} \Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\\
&= \frac{1}{(k_1+\ell_1)!(k_2+\ell_2)!}\sum_{d,e\leq R}\mu(d)\mu(e)\left(\log\frac{R}{d}\right)^{k_1+\ell_1}\left(\log\frac{R}{e}\right)^{k_2+\ell_2}\sum_{\substack{1\leq n\leq N\\ d|P_{\mathcal{H}_1}(n)\\ e|P_{\mathcal{H}_2}(n)}} 1.
\end{aligned}$$

For the inner sum, we let $d = a_1a_{12}$, $e = a_2a_{12}$ where $(d,e) = a_{12}$. Thus a_1, a_2, and a_{12} are pairwise relatively prime, and the divisibility conditions $d|P_{\mathcal{H}_1}(n)$ and $e|P_{\mathcal{H}_2}(n)$ become $a_1|P_{\mathcal{H}_1}(n)$, $a_2|P_{\mathcal{H}_2}(n)$, $a_{12}|P_{\mathcal{H}_1}(n)$, and $a_{12}|P_{\mathcal{H}_2}(n)$. As in the last section we get $\nu_{a_1}(\mathcal{H}_1)$ solutions for n modulo a_1, and $\nu_{a_2}(\mathcal{H}_2)$ solutions for n modulo a_2. If $p|a_{12}$, then from the two divisibility conditions we get $\nu_p(\mathcal{H}_1(p)\cap\mathcal{H}_2(p))$ solutions for n modulo p, where

$$\mathcal{H}(p) = \{{h'}_1,\ldots,{h'}_{\nu_p(\mathcal{H})} : {h'}_j \equiv h_i \in \mathcal{H} \text{ for some } i, 1\leq {h'}_j \leq p\}$$

Here if $p > h$ then $\mathcal{H}(p) = \mathcal{H}$. Alternatively, we can avoid this definition which is necessary only for small primes by defining

$$\overline{\nu}_p(\mathcal{H}_1\overline{\cap}\mathcal{H}_2) := \nu_p(\mathcal{H}_1(p)\cap\mathcal{H}_2(p)) := \nu_p(\mathcal{H}_1)+\nu_p(\mathcal{H}_2)-\nu_p(\mathcal{H}), \tag{6.5}$$

and then extend this definition to squarefree numbers by multiplicativity.[1] Thus we see that

$$\sum_{\substack{1\leq n\leq N\\ d|P_{\mathcal{H}_1}(n)\\ e|P_{\mathcal{H}_2}(n)}} 1 = \nu_{a_1}(\mathcal{H}_1)\nu_{a_2}(\mathcal{H}_2)\overline{\nu}_{a_{12}}(\mathcal{H}_1\overline{\cap}\mathcal{H}_2)\left(\frac{N}{a_1a_2a_{12}}+O(1)\right),$$

and have

(6.6)
$$\begin{aligned}
&\mathcal{S}_R(N;\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2)\\
&\quad= \frac{N}{(k_1+\ell_1)!(k_2+\ell_2)!}\sideset{}{'}\sum_{\substack{a_1a_{12}\leq R\\ a_2a_{12}\leq R}}\frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2\nu_{a_1}(\mathcal{H}_1)\nu_{a_2}(\mathcal{H}_2)\overline{\nu}_{a_{12}}(\mathcal{H}_1\overline{\cap}\mathcal{H}_2)}{a_1a_2a_{12}}\\
&\qquad\qquad\times\left(\log\frac{R}{a_1a_{12}}\right)^{k_1+\ell_1}\left(\log\frac{R}{a_2a_{12}}\right)^{k_2+\ell_2}\\
&\qquad + O\Big((\log R)^M\sideset{}{'}\sum_{\substack{a_1a_{12}\leq R\\ a_2a_{12}\leq R}}\mu(a_1)^2\mu(a_2)^2\mu(a_{12})^2\nu_{a_1}(\mathcal{H}_1)\nu_{a_2}(\mathcal{H}_2)\overline{\nu}_{a_{12}}(\mathcal{H}_1\overline{\cap}\mathcal{H}_2\Big)\\
&\quad= N\mathcal{T}_R(\ell_1,\ell_2;\mathcal{H}_1,\mathcal{H}_2)+O\big(R^2(3\log R)^{3k+M}\big),
\end{aligned}$$

[1] We are making a convention here that for $\overline{\nu}_p$ we take intersections modulo p.

where the prime on the sum indicates the summands are pairwise relatively prime, and we have estimated the error term using Lemma 4 and Corollary 4 by

$$
\begin{aligned}
&\ll (\log R)^M \sum_{q\le R^2}{}^{\flat} \sum_{q=a_1a_2a_{12}} d_k(q) \\
&\ll (\log R)^M \sum_{q\le R^2}{}^{\flat} d_3(q)d_k(q) \\
&\ll (\log R)^M \sum_{q\le R^2}{}^{\flat} d_{3k}(q) \\
&\ll R^2(3\log R)^{3k+M}.
\end{aligned} \tag{6.7}
$$

By (5.5) we have for $c_1, c_2 > 0$

$$
\mathcal{T}_R(\ell_1,\ell_2;\mathcal{H}_1,\mathcal{H}_2) = \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)} F(s_1,s_2)\frac{R^{s_1}}{{s_1}^{k_1+\ell_1+1}}\frac{R^{s_2}}{{s_2}^{k_2+\ell_2+1}}ds_1ds_2, \tag{6.8}
$$

where, letting $s_j = \sigma_j + it_j$ and assuming $\sigma_1, \sigma_2 > 0$,

$$
\begin{aligned}
F(s_1,s_2) &= \sum_{1\le a_1,a_2,a_{12}<\infty}{}' \frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2\nu_{a_1}(\mathcal{H}_1)\nu_{a_2}(\mathcal{H}_2)\overline{\nu}_{a_{12}}(\mathcal{H}_1\overline{\cap}\mathcal{H}_2)}{{a_1}^{1+s_1}{a_2}^{1+s_2}{a_{12}}^{1+s_1+s_2}} \\
&= \prod_p \Big(1 - \frac{\nu_p(\mathcal{H}_1)}{p^{1+s_1}} - \frac{\nu_p(\mathcal{H}_2)}{p^{1+s_2}} + \frac{\overline{\nu}_p(\mathcal{H}_1\overline{\cap}\mathcal{H}_2)}{p^{1+s_1+s_2}}\Big).
\end{aligned} \tag{6.9}
$$

Since for all $p > h$ we have $\nu_p(\mathcal{H}_1) = k_1$, $\nu_p(\mathcal{H}_2) = k_2$, and $\nu_p(\mathcal{H}_1\cap\mathcal{H}_2) = r$, we factor out the dominant zeta-factors and write

$$
F(s_1,s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)\frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1}\zeta(1+s_2)^{k_2}}, \tag{6.10}
$$

where by (4.1)

$$
G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2) = \prod_p \left(\frac{\Big(1-\frac{\nu_p(\mathcal{H}_1)}{p^{1+s_1}} - \frac{\nu_p(\mathcal{H}_2)}{p^{1+s_2}} + \frac{\overline{\nu}_p(\mathcal{H}_1\overline{\cap}\mathcal{H}_2)}{p^{1+s_1+s_2}}\Big)\Big(1-\frac{1}{p^{1+s_1+s_2}}\Big)^r}{\Big(1-\frac{1}{p^{1+s_1}}\Big)^{k_1}\Big(1-\frac{1}{p^{1+s_2}}\Big)^{k_2}}\right) \tag{6.11}
$$

is analytic and uniformly bounded for $\sigma_1, \sigma_2 > -1/4+\delta$ for any fixed $\delta > 0$. Also we see immediately from (2.2), (6.1) and (6.5) that

$$
G_{\mathcal{H}_1,\mathcal{H}_2}(0,0) = \mathfrak{S}(\mathcal{H}). \tag{6.12}
$$

Further, the same argument leading to (5.13) shows that with $\sigma_1, \sigma_2 > -1/100$, $\delta_i = -\min(\sigma_i, 0)$ we have for s_1, s_2 on $\mathcal{L}$ or to the right of $\mathcal{L}$ with U defined in (5.12) the estimate

$$
G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2) \ll \exp(CkU^{\delta_1+\delta_2}\log\log U). \tag{6.13}
$$

Thus we have for s_1 and s_2 on $\mathcal{L}$ or to the right of $\mathcal{L}$ that

$$
F(s_1,s_2) \ll \exp(CkU^{\delta_1+\delta_2}\log\log U)\big(\log(2+|t_1|)\log(2+|t_2|)\big)^{2k}\max\left(1,\frac{1}{|s_1+s_2|^r}\right). \tag{6.14}
$$

The integrand of (6.8) vanishes as either $|t_1|\to\infty$ or $|t_2|\to\infty$, $\sigma_1,\sigma_2 \in [-1/100, 1]$.

In what follows we will examine a more general situation. Although full generality is not needed here, a slightly different situation in the proof of Theorem 2 will be covered by the general form below.

Let us examine the integral

(6.15)
$$\mathcal{T}'_R(d,a,b,k_1,k_2,\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2) := \frac{1}{(2\pi i)^2}\int\limits_{(1)}\int\limits_{(1)} \frac{G(s_1,s_2)\zeta^d(1+s_1+s_2)R^{s_1+s_2}ds_1ds_2}{\zeta^a(1+s_1)\zeta^b(1+s_2)s_1^{k_1+\ell_1+1}s_2^{k_2+\ell_2+1}}$$

where

(6.16)
$$u := k_1 - a + \ell_1 \geq 0,\ v := k_2 - b + \ell_2 \geq 0,\ a \geq 0,\ b \geq 0,\ d \geq 0,$$
$$\min(a,b) \geq \max(d, cK),\ \max(a,b,d) \leq CK,\ \max(u,v) \leq CK\log_2 K/\log K,$$
$$k_1 \geq 1,\ k_2 \geq 1.$$

The only restriction we impose on K now is the one mentioned in (5.1), that is,

(6.17) $$K \ll_{\eta_0} (\log R)^{1/2-\eta_0} \text{ with an arbitrary fixed } \eta_0 > 0.$$

The necessary further restrictions for K will appear at the relevant places of our present examination. The formulation of Theorems $1''$ and $2''$ will be left for the end of Section 8.

The only property of $G(s_1,s_2)$ used in this analysis will be that $G(s_1,s_2)$ is regular on $\mathcal{L}$ and to the right of $\mathcal{L}$, and satisfies the estimate (6.13). (At the end we will substitute the evaluation (6.12) at $s_1 = s_2 = 0$, but the value does not play any role in our examination.) Using the notation $\zeta(1+s)s = W(s)$ we can write this as

(6.18) $$I = \mathcal{T}^*_R(d,a,b,u,v,\mathcal{H}_1,\mathcal{H}_2) := \frac{1}{(2\pi i)^2}\int\limits_{(1)}\int\limits_{(1)} \frac{D(s_1,s_2)R^{s_1+s_2}ds_1ds_2}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}.$$

where

(6.19) $$D(s_1,s_2) = \frac{G(s_1,s_2)W^d(s_1+s_2)}{W^a(s_1)W^b(s_2)}$$

is regular on $\mathcal{L}$ and to the right of $\mathcal{L}$. Further, similarly to the deduction from (6.14) the integrand vanishes as $|t_1| \to \infty$ or $|t_2| \to \infty$.

We will concentrate first on the main term which will be the integral I_1 below. The analysis of the error terms is relatively simple if M is an arbitrarily large fixed constant, which is sufficient to prove (2.18).

First step. Move the contour (1) for the integral over s_1 to $\mathcal{L}$. We pass a pole of order $u+1$ at $s_1 = 0$. We obtain

(6.20) $$I = I_1 + \frac{1}{(2\pi i)^2}\int_{(1)}\int_{\mathcal{L}} \frac{D(s_1,s_2)R^{s_1+s_2}ds_1ds_2}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d} = I_1 + I_2,$$

where

(6.21)
$$\begin{aligned} I_1 :&= \frac{1}{2\pi i}\int_{(1)} \operatorname{Res}_{s_1=0}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\right)ds_2 \\ &= \frac{1}{2\pi i}\int\limits_{(1)} \frac{1}{u!}\left\{\sum_{i=0}^{u}\binom{u}{i}(\log R)^{u-i}\frac{\partial^i}{\partial s_1^i}\left(\frac{D(s_1,s_2)}{(s_1+s_2)^d}\right)\Bigg|_{s_1=0}\right\}\frac{R^{s_2}}{s_2^{v+1}}ds_2. \end{aligned}$$

We denote the complete integrand above by $Z(s_2)$ and express

$$\begin{aligned}\frac{\partial^i}{\partial s_1^i}\left(\frac{D(s_1,s_2)}{(s_1+s_2)^d}\right)\Bigg|_{s_1=0} &= (-1)^i\frac{D(0,s_2)d(d+1)\dots(d+i-1)}{s_2^{d+i}}\\ &+\sum_{j=1}^{i}\binom{i}{j}\frac{\partial^j}{\partial s_1^j}D(s_1,s_2)\Bigg|_{s_1=0}\cdot(-1)^{i-j}\frac{d(d+1)\dots(d+i-j-1)}{s_2^{d+i-j}}\end{aligned}\tag{6.22}$$

where in case of $i=j$ (including also the case when $i=j=0$ and $d\ge 0$ arbitrary) the empty product is 1 in the numerator.

Second step. Let us denote the contribution of the first term in (6.22) to (6.21) by $I_1(i,0)$ and the others by $I_1(i,j)$ $(1\le j\le i)$. $I_1(i,0)$ will belong to the main term, all $I_1(i,j)$ with $j\ge 1$ will just contribute to the secondary terms. Let us move now the contour (1) for the integral over s_2 to $\mathcal{L}$ in (6.21). We pass a pole of order $v+1+d+i-j$ in case of $I_1(i,j)$ and we obtain in this way

$$\begin{aligned}I_1 &= \frac{1}{u!}\sum_{i=0}^{u}\binom{u}{i}(\log R)^{u-i}\sum_{j=0}^{i}(-1)^{i-j}\binom{i}{j}\frac{d(d+1)\dots(d+i-j-1)}{(v+d+i-j)!}\times\\ &\times\sum_{\nu=0}^{v+d+i-j}\binom{v+d+i-j}{\nu}(\log R)^{v+d+i-j-\nu}\cdot\frac{\partial^\nu}{\partial s_2^\nu}\frac{\partial^j}{\partial s_1^j}D(s_1,s_2)\Bigg|_{s_1=s_2=0}\\ &+\frac{1}{2\pi i}\int_{\mathcal{L}}Z(s_2)ds_2 =: I_{1,1}+I_{1,2}.\end{aligned}\tag{6.23}$$

The rather complicated formula (6.23) yields the main term and all secondary terms of the form $(\log R)^m$ exclusively for $m\in[d,d+u+v-1]$ and will additionally contribute to other secondary terms for $m\in[0,d-1]$. However, from the terms $I_{1,1}(i,j,\nu)$ belonging to (i,j,ν) in the triple summation only those with $\nu=0$, $j=0$ contribute to the main term of order $(\log R)^{d+u+v}$, since in all other terms the exponent of $\log R$ is $d+u+v-j-\nu$.

We will often use Cauchy's estimate for the derivatives of functions, basically for that of ζ, G or $D(s_1,s_2)$ defined above in (6.19), where all a,b,d will satisfy (6.16) with a parameter K satisfying (6.17).

Further $G(s_1,s_2)$ will satisfy the estimate (6.13), which therefore will hold for $D(s_1,s_2)$, too, on $\mathcal{L}$ and to the right of $\mathcal{L}$. Cauchy's estimate yields ($z=\sigma+it, z_0=\sigma_0+it_0$)

$$|f^{(j)}(z_0)|\le j!\max_{|z-z_0|=\eta}|f(z)|\cdot\eta^{-j},\tag{6.24}$$

if $f(z)$ is analytic for $|z-z_0|\le\eta$. We will actually choose

$$\eta=(C\log U\log T)^{-1}\tag{6.25}$$

with

$$T=T(t)=|t|+3,\qquad U=CK^2\log(2h).\tag{6.26}$$

(We remind the reader that the generic constants c, C etc. may take different values at different appearances.) We list a few consequences applied to the function $D(s_1,s_2)=\frac{G(s_1,s_2)W^d(s_1+s_2)}{W^a(s_1)W^b(s_2)}$ with G defined in (6.11). By (6.25)–(6.26), if z_0 is on $\mathcal{L}$ or to the right of $\mathcal{L}$ then still the whole circle $|z-z_0|=\eta$ will remain in the "good" region (4.3) and therefore we will have in case of (6.16) by the estimate (6.13), for s_1, s_2 on $\mathcal{L}$ or to the right of $\mathcal{L}$,

$$\begin{aligned}
&\frac{\partial^j}{\partial s_1^j}\frac{\partial^\nu}{\partial s_2^\nu}D(s_1,s_2)\\
&\ll j!\nu!(C\log U)^{j+\nu}(\log T_1)^j(\log T_2)^\nu \max_{|s_1^*-s_1|\le\eta,|s_2^*-s_2|\le\eta}|D(s_1^*,s_2^*)| \\
&\ll (CK)^{j+\nu}\frac{e^{CKU^{\delta_1+\delta_2}\log_2 U}(\log T_1)^{CK}(\log T_2)^{CK}}{\max(c,|s_1|)^a\max(c,|s_2|)^b}\max(1,|s_1+s_2|)^d,
\end{aligned} \tag{6.27}$$

supposing here and later always

$$\max(j,\nu)\le CK. \tag{6.28}$$

In case of $\max(|s_1|,|s_2|)\le C$ this reduces to

$$\frac{\partial^j}{\partial s_1^j}\frac{\partial^\nu}{\partial s_2^\nu}D(s_1,s_2)\ll (CK)^{j+\nu}e^{CK\cdot U^{\delta_1+\delta_2}\log_2 U}. \tag{6.29}$$

So we obtain for the contribution of a term $I_{1,1}(i,j,\nu)$ to I_1 in (6.23) (using (6.27) with $s_1=s_2=0$, that is $\delta_1=\delta_2=0$ and $\eta=1/C\log U$ by (6.25)) the upper estimate

(6.30)
$$\begin{aligned}
|I_{1,1}(i,j,\nu)| \le \frac{1}{u!}\binom{u}{i}\binom{i}{j}(\log R)^{u+v+d}2^{CK}(CK)^{j+\nu}e^{CK\log_2 U}\frac{(\log R)^{-j-\nu}}{(v+d)!}\times\\
\times\prod_{\mu=0}^{i-j-1}\frac{d+\mu}{v+d+1+\mu}.
\end{aligned}$$

On the other hand, concerning the main term $I_{1,0}$ corresponding to $\nu=0$, $j=0$, the total contribution of all i with $0\le i\le u$ will be by (4.16) and Lemma 6

(6.31)
$$\begin{aligned}
I_{1,0} :&= \sum_{i=0}^{u} I_1(i,0,0) := \frac{(\log R)^{d+u+v}}{u!}\sum_{i=0}^{u}\binom{u}{i}(-1)^i\frac{d(d+1)\cdots(d+i-1)}{(v+d+i)!}D(0,0)\\
&= G(0,0)(\log R)^{d+u+v}\cdot S(d,u,v)\\
&= G(0,0)\cdot\binom{v+u}{u}\cdot\frac{(\log R)^{d+u+v}}{(d+u+v)!}\\
&= \mathfrak{S}(\mathcal{H})\binom{v+u}{u}\frac{(\log R)^{d+v+u}}{(d+v+u)!}.
\end{aligned}$$

Since in (6.30) the last product is less than 1, $\binom{i}{j}\le 2^i\le 2^u$, the summation of $\binom{u}{i}$ over i contributes a factor 2^u, which both can be included together with 2^{CK} in (6.30) into $e^{CK\log_2 U}$. Thus, in order to show that the main term is dominant in the appropriate sense expressed in Theorems 1″ and 2″ we have to prove

(6.32)
$$B := e^{CK\log_2 U}\sum_{j=0}^{i}\sum_{\substack{\nu=0\\ \nu+j\ge1}}^{v+d+i-j}\left(\frac{CK}{\log R}\right)^j\left(\frac{CK}{\log R}\right)^\nu = O\left(\frac{(v+d)!}{(v+d+u)!}\right)(\log R)^{-\frac{2}{3}},$$

taking into account the quantities $(v+d)!$ and $(d+v+u)!$ in the denominators in (6.30) and (6.31), respectively.

Since the left-hand side is a geometric series, supposing the weak restriction

$$K \le (\log R)^{1/4} \asymp (\log N)^{1/4} \tag{6.33}$$

we obtain already

$$B \le C(\log R)^{-3/4} e^{CK \log_2 U}. \tag{6.34}$$

On the other hand, for the term before $(\log R)^{-2/3}$ on the right-hand side of (6.32) we have by (6.16) the lower estimate

$$e^{-Cu \log K} \ge e^{-CK \log_2 K} \ge e^{-CK \log_2 U}. \tag{6.35}$$

(In practice we will have $u, v \asymp \sqrt{K}$.) Thus the main term will dominate in the sense of (6.32), if

$$K \log_2 U \le c \log_2 R. \tag{6.36}$$

This will be satisfied if

$$K \le c \log_2 R / \log_4 R, \quad h \ll (\log R)^C \tag{6.37}$$

or

$$K \le c \log_2 R / \log_3 R, \quad h \ll R^C. \tag{6.38}$$

Remark. If we use a weaker restriction $u, v \le CK$ and not the slightly stronger (6.16), then instead of (6.37) we need

$$K(\log K + \log_2 U) = K(\log K + \log_2 K + \log_3 h + O(1)) \le c \log_2 R$$

which clearly holds if (6.38) is satisfied for K and h, which is much more general in h and only slightly weaker in K than (6.37).

Now we turn to the estimation of $I_{1,2}$ in (6.23) which will be analogous to (5.13)–(5.15). Taking into account (6.21)–(6.22) we have by (6.16)–(6.17), (6.27) analogously to the proof of Corollary 3

$$\begin{aligned} &\int_{\mathcal{L}} Z(s_2) ds_2 \\ &\ll \int_{\mathcal{L}} \frac{d^u \max(1, |s_2|)^d (\log(|t_2| + 3))^{CK} e^{CKU^{\delta_2} \log_2 U} R^{-\delta_2}}{|s_2|^{a+d}} |ds_2| \\ &\ll \int_{\mathcal{L}} \frac{e^{CKU^{\delta_2} \log_2 U} (\log(|t_2| + 3))^{CK} R^{-\delta_2}}{|s_2|^a} |ds_2| \\ &\ll e^{-c\sqrt{\log R}}. \end{aligned} \tag{6.39}$$

The above analysis actually reveals that by (6.31)–(6.38) we have in our original case (6.8) ($a = k_1$, $b = k_2$, $d = r$, i.e. $u = \ell_1$, $v = \ell_2$, $d = r$)

$$\begin{aligned} I_1 = \binom{\ell_1 + \ell_2}{\ell_1} \frac{(\log R)^{r+\ell_1+\ell_2}}{(r + \ell_1 + \ell_2)!} \mathfrak{S}(\mathcal{H}) + \sum_{\kappa=1}^{r+\ell_1+\ell_2} \mathcal{D}_\kappa(\ell_1, \ell_2, \mathcal{H}_1, \mathcal{H}_2)(\log R)^{r+\ell_1+\ell_2-\kappa} \\ + O(e^{-c\sqrt{\log R}}) \end{aligned} \tag{6.40}$$

where $\kappa = j + \nu$ in (6.23) and

$$\begin{aligned} &\mathcal{D}_\kappa(\ell_1, \ell_2, \mathcal{H}_1, \mathcal{H}_2) \ll e^{CK \log_2 U} \cdot (CK)^\kappa \Leftrightarrow \\ &\Leftrightarrow \mathcal{D}_\kappa(\ell_1, \ell_2, \mathcal{H}_1, \mathcal{H}_2) \ll e^{CK(\log_2 K + \log_3 h)} \cdot (CK)^\kappa \end{aligned} \tag{6.41}$$

which is really the required form (6.2)–(6.3). Now in view of (6.36)–(6.38) this would prove Theorems $1'$ and $1''$ if we can show a suitable estimate for the other integral I_2 defined in (6.20), which will represent a secondary term.

For I_2, after interchange of the two integrations we move the contour (1) to the left to $\mathcal{L}$ passing a pole of order d at $s_2 = -s_1$ and a pole of order $v+1$ at $s_2 = 0$ and obtain

$$
\begin{aligned}
I_2 &= \frac{1}{2\pi i}\int_{\mathcal{L}} \operatorname{Res}_{s_2=-s_1}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\Big)\,ds_1 \\
&\quad + \frac{1}{2\pi i}\int_{\mathcal{L}} \operatorname{Res}_{s_2=0}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\Big)\,ds_1 \\
&\quad + \frac{1}{(2\pi i)^2}\int_{\mathcal{L}}\int_{\mathcal{L}} F(s_1,s_2)\frac{R^{s_1}}{{s_1}^{k_1+1}}\frac{R^{s_2}}{{s_2}^{k_2+1}}ds_1ds_2 := I_{2,1}+I_{2,2}+I_{2,3}.
\end{aligned}
\tag{6.42}
$$

(Here if $d = 0$ the first term is zero.) By the argument of Corollary 3 and (6.14) the third integral $I_{2,3}$ is $\ll e^{-c\sqrt{\log R}}$. The second integral $I_{2,2}$ is completely analogous to $I_{1,2}$ in (6.23), which was estimated by $e^{-c\sqrt{\log R}}$ in (6.39), the only change being that the role of s_1 and s_2 is interchanged.

Finally, for the first term we have $d \geq 1$ and

$$
\begin{aligned}
&\operatorname{Res}_{s_2=-s_1}\Big(\frac{D(s_1,s_2)R^{s_1+s_2}}{s_1^{u+1}s_2^{v+1}(s_1+s_2)^d}\Big) \\
&= \lim_{s_2\to -s_1}\frac{1}{(d-1)!}\frac{\partial^{d-1}}{\partial {s_2}^{d-1}}\left(\frac{D(s_1,s_2)R^{s_1+s_2}}{{s_1}^{u+1}{s_2}^{v+1}}\right) \\
&= \frac{1}{(d-1)!}\sum_{j=0}^{d-1}\mathcal{B}_j(s_1,\mathcal{H}_1,\mathcal{H}_2)(\log R)^{d-1-j},
\end{aligned}
\tag{6.43}
$$

where

$$
\mathcal{B}_j(s_1,\mathcal{H}_1,\mathcal{H}_2) = \binom{d-1}{j}\sum_{\nu=0}^{j}\binom{j}{\nu}\frac{\partial^{j-\nu}}{\partial s_2^{j-\nu}}D(s_1,s_2)\Big|_{s_2=-s_1}\cdot\frac{(-1)^\nu(v+1)\dots(v+\nu)}{(-1)^{\nu+v+1}s_1^{u+v+\nu+2}}.
\tag{6.44}
$$

We thus obtain

$$
I_2 = \frac{1}{(d-1)!}\sum_{j=0}^{d-1}\mathcal{C}_j(\mathcal{H}_1,\mathcal{H}_2)(\log R)^{d-1-j} + O_(e^{-c\sqrt{\log R}}),
\tag{6.45}
$$

where

$$
\mathcal{C}_j(\mathcal{H}_1,\mathcal{H}_2) = \frac{1}{2\pi i}\int_{\mathcal{L}}\mathcal{B}_j(s_1,\mathcal{H}_1,\mathcal{H}_2)\,ds_1 \quad (j=0,1,2,\dots,d-1).
\tag{6.46}
$$

It remains to estimate these quantities, which are independent of R. By (6.44) we see that the functions $\mathcal{B}_j(s_1)$ tend to zero as $|t_1| \to \infty$, $-1/100 \leq \sigma \leq 1$, further by (6.27)

$$
\mathcal{B}_j(s_1,\mathcal{H}_1,\mathcal{H}_2) \ll \sum_{\nu=0}^{j}(2d)^j(CK)^{j-\nu}(\log T_1)^{CK}e^{CKU^{2\delta_1}\log_2 U}(CK)^\nu\frac{1}{|t_1|^{u+v+\nu+2+a+b}},
\tag{6.47}
$$

and therefore we may move the contour $\mathcal{L}$ back to the imaginary axis with a semicircle of radius $1/\log U$ centered and to the left of $s_1 = 0$. The contribution to $\mathcal{C}_j$, $1 \le j \le k$, from the integral along the imaginary axis is

$$\ll (CK^2)^j (\log U)^{CK} e^{CK \log_2 U} \ll (CK^2)^j e^{CK \log_2 U} \tag{6.48}$$

since by (6.16) we have

$$L := \max(u, v) \le CK. \tag{6.49}$$

We bound the contribution to $\mathcal{C}_j$, $1 \le j \le d-1$, from the semicircle contour similarly to the above as

$$\ll (CK^2)^j e^{CK \log_2 U}. \tag{6.50}$$

We conclude that for arbitrarily large fixed M we have

$$\mathcal{C}_j(\mathcal{H}_1, \mathcal{H}_2) \ll_M (\log\log 10h)^{C'_j} \tag{6.51}$$

with constants C'_j depending on M. This proves Theorem 1′.

We have to be somewhat more careful, however, in order to show Theorem 1″.

By (6.45)–(6.50) we have for the sum in I_2 by (6.17)

$$\begin{aligned} I'_2 &\ll \frac{(\log R)^{d-1}}{(d-1)!} \sum_{j=0}^{d-1} \left(\frac{CK^2}{\log R}\right)^j e^{CK \log_2 U} \\ &\ll \frac{(\log R)^{d-1} e^{CK \log_2 U}}{(d-1)!}. \end{aligned} \tag{6.52}$$

If we compare the estimate (6.52) for I'_2 with the asymptotic value of the main term in (6.31), we can see that in order to prove our Theorem 1″ we have to show that

$$\frac{(d+v+u)!}{(d-1)!} e^{CK \log_2 U} = O\big((\log R)^{u+v+\frac{1}{3}}\big) \tag{6.53}$$

which by (6.49) will follow from

$$L \log K + K(\log_2 K + \log_3 10h) \le cL \log_2 R. \tag{6.54}$$

Now this condition will be satisfied if

$$K \le (\log R)^c \tag{6.55}$$

and

$$\frac{K}{L} \log_2 K \le c \log_2 R, \tag{6.56}$$

further

$$\frac{K}{L} \le \begin{cases} c \log_2 R / \log_4 R & \text{if } h \ll (\log R)^C \\ c \log_2 R / \log_3 R & \text{if } h \ll R^C. \end{cases} \tag{6.57}$$

The conditions (6.55)–(6.57) clearly hold if K satisfies (6.37) or (6.38), respectively.

Now (6.55)–(6.57) prove our Theorem 1″, too, the final restriction on K being (6.37)–(6.38).

7. Proof of Theorem 2

We use the same notation as earlier. In addition, let

$$\Theta(x;q,a) := \sum_{\substack{p\le x\\ p\equiv a (\bmod\ q)}} \log p = [(a,q)=1]\frac{x}{\phi(q)} + E'(x;q,a),$$
$$E'(x,q) := \max_{a;(a,q)=1} |E'(x,q,a)|, \qquad E^*(X,q) = \max_{x\le X} E'(x,q). \tag{7.1}$$

where $[S]$ is 1 if the statement S is true and is 0 if S is false. To prove Theorems 2 and 2′ we need the Bombieri–Vinogradov Theorem which we formulate in the form that for any $A>0$ there is a number $B=B(A)$ so that

$$\sum_{q\le Q} E^*(X,q) \ll \frac{X}{(\log X)^A} \tag{7.2}$$

provided $1\le Q\le \frac{X^{\frac12}}{(\log X)^B}$. We prove the following stronger version of Theorem 2. Let

$$C_R(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0) = \begin{cases} 1, & \text{if } h_0\notin\mathcal{H};\\ \frac{(\ell_1+\ell_2+1)\log R}{(\ell_1+1)(r+\ell_1+\ell_2+1)}, & \text{if } h_0\in\mathcal{H}_1 \text{ and } h_0\notin\mathcal{H}_2;\\ \frac{(\ell_1+\ell_2+2)(\ell_1+\ell_2+1)\log R}{(\ell_1+1)(\ell_2+1)(r+\ell_1+\ell_2+1)}, & \text{if } h_0\in\mathcal{H}_1\cap\mathcal{H}_2, \end{cases} \tag{7.3}$$

and by relabeling the variables we obtain the corresponding form if $h_0\in\mathcal{H}_2$ and $h_0\notin\mathcal{H}_1$. Then we have with the notation $\vartheta(n)$ in (2.15)

Theorem 2′. *Suppose $h\ll R$. Given any positive A, there is a $B=B(A,M)$ such that for $R\ll_{M,A} \frac{N^{\frac14}}{(\log N)^B}$ and $R,N\to\infty$,*

$$\begin{aligned}\sum_{n=1}^{N} &\Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\vartheta(n+h_0)\\ &= \frac{C_R(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)}{(r+\ell_1+\ell_2)!}\binom{\ell_1+\ell_2}{\ell_1}\mathfrak{S}(\mathcal{H}^0)N(\log R)^{r+\ell_1+\ell_2}\\ &\quad + N\sum_{j=1}^{r}\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)(\log R)^{r+\ell_1+\ell_2-j}\\ &\quad + O_{M,A}\left(\frac{N}{(\log N)^A}\right)\end{aligned} \tag{7.4}$$

where the $\mathcal{D}_j(\ell_1,\ell_2,\mathcal{H}_1,\mathcal{H}_2,h_0)$'s are functions independent of R and N which satisfy the bound

$$\mathcal{D}_j(\mathcal{H}_1,\mathcal{H}_2,h_0) \ll_M (\log U)^{C_j} \ll_M (\log\log 10h)^{C'_j} \tag{7.5}$$

for some positive constants C_j, C'_j depending on M. Assuming the Elliott–Halberstam conjecture (7.4) holds for $R\ll_M N^{\frac12-\varepsilon}$ and $h\le R^{\varepsilon}$, for any given $\varepsilon>0$.

Proof. We will assume that both $\mathcal{H}_1$ and $\mathcal{H}_2$ are non-empty and thus $k_1\ge 1$, $k_2\ge 1$. The proof in the case when one of these sets is empty is much easier and may be obtained by an argument analogous to that of Section 5. We have

(7.6)
$$\begin{aligned}&\tilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0)\\ &:=\sum_{n=1}^{N}\Lambda_R(n;\mathcal{H}_1,\ell_1)\Lambda_R(n;\mathcal{H}_2,\ell_2)\vartheta(n+h_0)\\ &=\frac{1}{(k_1+\ell_1)!(k_2+\ell_2)!}\sum_{d,e\le R}\mu(d)\mu(e)\left(\log\frac{R}{d}\right)^{k_1+\ell_1}\left(\log\frac{R}{e}\right)^{k_2+\ell_2}\sum_{\substack{1\le n\le N\\ d|P_{\mathcal{H}_1}(n)\\ e|P_{\mathcal{H}_2}(n)}}\vartheta(n+h_0).\end{aligned}$$

For the inner sum, we let $d=a_1a_{12}$, $e=a_2a_{12}$ where $(d,e)=a_{12}$, and thus a_1, a_2, and a_{12} are pairwise relatively prime. As in the last section, the n for which $d|P_{\mathcal{H}_1}(n)$ and $e|P_{\mathcal{H}_2}(n)$ cover certain residue classes modulo $[d,e]$. If $n\equiv b\,(\mathrm{mod}\,a_1a_2a_{12})$ is such a residue class, then letting $m=n+h_0\equiv b+h_0(\mathrm{mod}\,a_1a_2a_{12})$ we see this residue class contributes to the inner sum

(7.7)
$$\begin{aligned}\sum_{\substack{1+h_0\le m\le N+h_0\\ m\equiv b+h_0\,(\mathrm{mod}\,a_1a_2a_{12})}}\vartheta(m)&=\Theta(N+h_0;a_1a_2a_{12},b+h_0)-\Theta(h_0;a_1a_2a_{12},b+h_0)\\ &=[(b+h_0,a_1a_2a_{12})=1]\frac{N}{\phi(a_1a_2a_{12})}+E(N;a_1a_2a_{12},b+h_0)+O(h\log N).\end{aligned}$$

We need to determine the number of these residue classes where $(b+h_0,a_1a_2a_{12})=1$ so that the main term is non-zero. If $p|a_1$ then $b\equiv -h_j\ (\mathrm{mod}\,p)$ for some $h_j\in\mathcal{H}_1$, and therefore $b+h_0\equiv h_0-h_j\ (\mathrm{mod}\,p)$. Thus, if h_0 is distinct modulo p from all the $h_j\in\mathcal{H}_1$ then all $\nu_p(\mathcal{H}_1)$ residue classes satisfy the relatively prime condition, while otherwise $h_0\equiv h_j(\mathrm{mod}\,p)$ for some $h_j\in\mathcal{H}_1$ leaving $\nu_p(\mathcal{H}_1)-1$ residue classes with a non-zero main term. We introduce the notation ${\nu_p}^*({\mathcal{H}_1}^0)$ for this number in either case, where we define for a set $\mathcal{G}$

$$ {\nu_p}^*(\mathcal{G})=\nu_p(\mathcal{G}^0)-1. \tag{7.8}$$

and

$$ \mathcal{G}^0=\mathcal{G}\cup\{h_0\}. \tag{7.9}$$

We extend this definition to ${\nu_d}^*({\mathcal{H}_1}^0)$ for squarefree numbers d by multiplicativity. (The function ${\nu_d}^*$ is familiar in sieve theory, see [6].) The same applies for $\nu_d^*(\mathcal{H}_2)$ and $\overline{\nu}_d^*\big((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\big)$, as in (6.5).

Next, the divisibility conditions $a_2|P_{\mathcal{H}_2}(n)$ $a_{12}|P_{\mathcal{H}_1}(n)$, and $a_{12}|P_{\mathcal{H}_2}(n)$ are handled as in the last section together with the above considerations, and we conclude, since $E(n;q,a)\ll(\log N)$ if $(a,q)>1$ and $q\le N$:

(7.10)
$$\begin{aligned}\sum_{\substack{1\le n\le N\\ d|P_{\mathcal{H}_1}(n)\\ e|P_{\mathcal{H}_2}(n)}}\vartheta(n+h_0)&={\nu_{a_1}}^*({\mathcal{H}_1}^0){\nu_{a_2}}^*({\mathcal{H}_2}^0){\overline{\nu}_{a_{12}}}^*\left((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\right)\frac{N}{\phi(a_1a_2a_{12})}\\ &\quad+O\left(d_k(a_1a_2a_{12})\left(\max_{\substack{b\\(b,a_1a_2a_{12})=1}}\left|E(N;a_1a_2a_{12},b)\right|+h(\log N)\right)\right).\end{aligned}$$

Substituting this into (7.6) we conclude

(7.11)

$$\begin{aligned}
&\tilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0)\\
&\quad = \frac{N}{(k_1+\ell_1)!(k_2+\ell_2)!}\sideset{}{'}\sum_{\substack{a_1a_{12}\le R\\ a_2a_{12}\le R}} \frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2{\nu_{a_1}}^*({\mathcal{H}_1}^0){\nu_{a_2}}^*({\mathcal{H}_2}^0){\overline{\nu}_{a_{12}}}^*\left((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\right)}{\phi(a_1a_2a_{12})}\\
&\qquad\qquad \times\left(\log\frac{R}{a_1a_{12}}\right)^{k_1+\ell_1}\left(\log\frac{R}{a_2a_{12}}\right)^{k_2+\ell_2}\\
&\quad + O\left((\log R)^M\sideset{}{'}\sum_{\substack{a_1a_{12}\le R\\ a_2a_{12}\le R}} d_k(a_1a_2a_{12})\max_{\substack{b\\(b,a_1a_2a_{12})=1}}\left|E(N;a_1a_2a_{12},b)\right|\right)\\
&\quad + O(hR^2(3\log N)^{M+3k+1})\\
&\quad = N\tilde{\mathcal{T}}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) + O\left((\log R)^M\mathcal{E}_k(N)\right) + O(hR^2(3\log N)^{M+3k+1}),
\end{aligned}$$

where the last error term was obtained by the estimate (6.7). To estimate the first error term we use Lemmas 2, 4 and the trivial estimate $|E(N;q,a)|\le\frac{2N\log N}{q}$ for $q\le N$, and (7.2) to find uniformly for $k\le\sqrt{(\log N)/18}$

$$\begin{aligned}
|\mathcal{E}_k(N)| &\le \sideset{}{^\flat}\sum_{q\le R^2} d_k(q)\max_{\substack{b\\(b,q)=1}}\left|E(N;q,b)\right|\sum_{q=a_1a_2a_{12}}1\\
&= \sideset{}{^\flat}\sum_{q\le R^2} d_k(q)d_3(q)E'(N,q)\\
&\le \sqrt{\sideset{}{^\flat}\sum_{q\le R^2}\frac{d_{3k}(q)^2}{q}}\sqrt{\sum_{q\le R^2}q(E^*(N,q))^2} \qquad (7.12)\\
&\le \sqrt{(\log N)^{9k^2}}\sqrt{2N\log N}\sqrt{\sum_{q\le R^2}E^*(N,q)}\\
&\ll N(\log N)^{(9k^2+1-A)/2},
\end{aligned}$$

provided $R^2\ll\frac{N^{\frac12}}{(\log N)^B}$. We conclude on relabeling that given any positive integers A and M there is a positive constant $B=B(A,M)$ so that for $R\ll\frac{N^{\frac14}}{(\log N)^B}$ and $h\le R$,

$$\tilde{\mathcal{S}}_R(N;\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) = N\tilde{\mathcal{T}}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) + O_M\left(\frac{N}{(\log N)^A}\right). \qquad (7.13)$$

Elliott and Halberstam conjectured that (7.2) holds whenever $Q\le x^{1-\varepsilon}$ for any given $\varepsilon>0$. This conjecture therefore gives (7.13) for the longer range $R\ll_M N^{\frac12-\varepsilon}$, $h\ll N^\varepsilon$.

Returning to the main term in (7.11), we have by (5.5) that

$$\mathcal{T}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0) = \frac{1}{(2\pi i)^2}\int_{(c_2)}\int_{(c_1)}F(s_1,s_2)\frac{R^{s_1}}{{s_1}^{k_1+\ell_1+1}}\frac{R^{s_2}}{{s_2}^{k_2+\ell_2+1}}ds_1ds_2, \qquad (7.14)$$

where, letting $s_j = \sigma_j + it_j$ and assuming $\sigma_1, \sigma_2 > 0$,

(7.15)

$$F(s_1,s_2) = \sideset{}{'}\sum_{1\le a_1,a_2,a_{12}<\infty} \frac{\mu(a_1)\mu(a_2)\mu(a_{12})^2 \nu_{a_1}{}^*(\mathcal{H}_1{}^0)\nu_{a_2}{}^*(\mathcal{H}_2{}^0)\overline{\nu}_{a_{12}}{}^*\left((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\right)}{\phi(a_1){a_1}^{s_1}\phi(a_2){a_2}^{s_2}\phi(a_{12}){a_{12}}^{s_1+s_2}}$$
$$= \prod_p \Big(1 - \frac{\nu_p{}^*(\mathcal{H}_1{}^0)}{(p-1)p^{s_1}} - \frac{\nu_p{}^*(\mathcal{H}_2{}^0)}{(p-1)p^{s_2}} + \frac{\overline{\nu}_p{}^*\left((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\right)}{(p-1)p^{s_1+s_2}}\Big).$$

We now consider three cases.

Case 1. Suppose $h_0 \notin \mathcal{H}$. Then for $p > h$ we have

$$\nu_p{}^*(\mathcal{H}_1{}^0) = k_1, \quad \nu_p{}^*(\mathcal{H}_2{}^0) = k_2, \quad \overline{\nu}_p{}^*\left((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\right) = r,$$

and therefore we define $G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)$ in this case by

$$F(s_1,s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)\frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1}\zeta(1+s_2)^{k_2}}. \tag{7.16}$$

We see G is analytic and uniformly bounded for $\sigma_1, \sigma_2 > -1/100$ by (6.13).

Case 2. Suppose $h_0 \in \mathcal{H}_1$ but $h_0 \notin \mathcal{H}_2$. (By relabeling this also covers the case where $h_0 \in \mathcal{H}_2$ and $h_0 \notin \mathcal{H}_1$.) Then for $p > h$ we have

$$\nu_p{}^*(\mathcal{H}_1{}^0) = k_1 - 1, \quad \nu_p{}^*(\mathcal{H}_2{}^0) = k_2, \quad \overline{\nu}_p{}^*\left((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\right) = r,$$

and therefore we define $G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)$ by

$$F(s_1,s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)\frac{\zeta(1+s_1+s_2)^r}{\zeta(1+s_1)^{k_1-1}\zeta(1+s_2)^{k_2}}, \tag{7.17}$$

and see G is analytic and uniformly bounded for $\sigma_1, \sigma_2 > -1/100$ by (6.13).

Case 3. Suppose $h_0 \in \mathcal{H}_1 \cap \mathcal{H}_2$. Then for $p > h$ we have

$$\nu_p{}^*(\mathcal{H}_1{}^0) = k_1 - 1, \quad \nu_p{}^*(\mathcal{H}_2{}^0) = k_2 - 1, \quad \overline{\nu}_p{}^*\left((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0\right) = r - 1,$$

and therefore we define $G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)$ in this case by

$$F(s_1,s_2) = G_{\mathcal{H}_1,\mathcal{H}_2}(s_1,s_2)\frac{\zeta(1+s_1+s_2)^{r-1}}{\zeta(1+s_1)^{k_1-1}\zeta(1+s_2)^{k_2-1}}, \tag{7.18}$$

where G is analytic and uniformly bounded for $\sigma_1, \sigma_2 > -1/100$ by (6.13).

We first show that in all three cases

$$G_{\mathcal{H}_1,\mathcal{H}_2}(0,0) = \mathfrak{S}(\mathcal{H}^0); \tag{7.19}$$

notice in the second two cases we have $\mathcal{H}^0 = \mathcal{H}$. To see this, by (4.1), (7.8) and (7.15) we have

(7.20)

$$G_{\mathcal{H}_1,\mathcal{H}_2}(0,0)$$
$$= \prod_p \left(1 - \frac{\nu_p(\mathcal{H}_1{}^0) + \nu_p(\mathcal{H}_2{}^0) - \overline{\nu}_p((\mathcal{H}_1\overline{\cap}\mathcal{H}_2)^0) - 1}{p-1}\right)\left(1 - \frac{1}{p}\right)^{-a(\mathcal{H}_1,\mathcal{H}_2,h_0)}$$
$$= \prod_p \left(1 - \frac{\nu_p(\mathcal{H}^0) - 1}{p-1}\right)\left(1 - \frac{1}{p}\right)^{-a(\mathcal{H}_1,\mathcal{H}_2,h_0)},$$

where in Case 1 $a(\mathcal{H}_1,\mathcal{H}_2,h_0)=k_1+k_2-r=k-r$, in Case 2 $a(\mathcal{H}_1,\mathcal{H}_2,h_0)=(k_1-1)+k_2-r=k-r-1$, and in Case 3 $a(\mathcal{H}_1,\mathcal{H}_2,h_0)=(k_1-1)+(k_2-1)-(r-1)=k-r-1$. Hence in Case 1 we have

$$\begin{aligned} G_{\mathcal{H}_1,\mathcal{H}_2}(0,0) &= \prod_p \left(\frac{p-\nu_p(\mathcal{H}^0)}{p-1}\right)\left(1-\frac{1}{p}\right)^{-(k-r)} \\ &= \prod_p \left(1-\frac{\nu_p(\mathcal{H}^0)}{p}\right)\left(1-\frac{1}{p}\right)^{-(k-r+1)} \\ &= \mathfrak{S}(\mathcal{H}^0), \end{aligned} \tag{7.21}$$

while in Cases 2 and 3 we have

$$\begin{aligned} G_{\mathcal{H}_1,\mathcal{H}_2}(0,0) &= \prod_p \left(\frac{p-\nu_p(\mathcal{H})}{p-1}\right)\left(1-\frac{1}{p}\right)^{-(k-r-1)} \\ &= \prod_p \left(1-\frac{\nu_p(\mathcal{H})}{p}\right)\left(1-\frac{1}{p}\right)^{-(k-r)} \\ &= \mathfrak{S}(\mathcal{H}) \quad (= \mathfrak{S}(\mathcal{H}^0).) \end{aligned} \tag{7.22}$$

We are now ready to evaluate $\mathcal{T}_R(\mathcal{H}_1,\mathcal{H}_2,\ell_1,\ell_2,h_0)$. There are two differences between the functions F and G that appear in (7.15), (7.16) and the earlier (6.9)–(6.11). The first is a factor of p in the denominator of the Euler product in (6.9) has been replaced by $p-1$, which only effects the value of constants in calculations. The second difference is the relationship between k_1, k_2, and r, which effects the residue calculations of the main terms. However the analysis of lower order terms and the error analysis is essentially unchanged and therefore we only need to examine the main terms. On the other hand, the evaluation of the main terms is covered by the very general treatment of the integral in (6.15), covering all Cases 1–3. Taking into account (7.16)–(7.18) we have

Case 1. $a=k_1,\quad b=k_2,\quad d=r \implies u=\ell_1,\quad v=\ell_2,\quad d=r$

Case 2. $a=k_1-1,\ b=k_2,\quad d=r \implies u=\ell_1+1,\ v=\ell_2,\quad d=r$

Case 3. $a=k_1-1,\ b=k_2-1,\ d=r-1 \implies u=\ell_1+1,\ v=\ell_2+1,\ d=r-1.$

Now, the general formula (6.31) and $G(0,0)=\mathfrak{S}(\mathcal{H}^0)$ immediately yields Theorems 2 and 2′.

Theorem 2 combined with Theorem 1 (with the choice of k_1,k_2,ℓ_1,ℓ_2 as arbitrarily large, fixed parameters) enables us to prove Theorems A, B and C, and thereby (2.18) but not (1.1). Another disadvantage of this treatment is that owing to the use of the Bombieri–Vinogradov theorem the whole result is ineffective, due to the eventual appearance of Siegel zeros. The next two sections will be devoted to solve the arising problems and thus prepare the way towards an effective proof of (1.1).

8. A modified Bombieri–Vinogradov theorem

The treatment of the error term in Section 7 contains a loss of a factor $O((\log N)^{Ck})$ due to the appearance of the factor $d_k(a_1a_2a_{12})$ in the second error term in (7.11) and allows a gain of an arbitrary power of $\log N$ due to the Bombieri–Vinogradov theorem. This clearly does not allow of any choice of an explicit function $K = K(N) \to \infty$ and one cannot hope therefore to show $p_{n+1}-p_n = o(\log p_n/y(n))$ in this way with any explicitly given function $y(n) \to \infty$.

In the course of proof in this section an essential role will be played by the following theorem of Heath-Brown [9]. We will show, however, in the next section an alternative way which avoids Heath-Brown's theorem. Let us formulate first

Hypothesis S. *There exists a constant c_0 such that for all $q \geq 2$ and all $\chi \pmod q$ we have $L(\sigma + it, \chi) \neq 0$ for $\sigma \geq 1 - c_0/\log(q(|t|+2)$.*

Theorem (Heath-Brown). *If S is false then there are infinitely many prime twins.*

This means that we are entitled to assume that Hypothesis S is true, at least when examining E_1, i.e., the difference between consecutive primes. This enables us to improve the Bombieri–Vinogradov theorem in the following way, thereby allowing $K = K(N) \to \infty$ explicitly (cf. (8.18)). We remark here that Heath-Brown's result is completely effective in the sense that more precisely (as it turns out from his work) we could choose a small fixed value of c_0 such that for a sufficiently large X

either (i) S holds with c_0 for all characters $\mathrm{mod}\, q \leq X$

or (ii) there are twin primes between $\log X$ and X^{500}.

Theorem 3. *On Hypothesis S we have for $Q^* = X^{1/2}\exp(-c^*\sqrt{\log X})$ (c^* arbitrary positive constant)*

$$\sum_{q \leq Q^*} E^*(X,q) \ll X\exp(-c_1\sqrt{\log X}), \tag{8.1}$$

where c_1 depends on c^* and c_0 (see Hypothesis S) in an explicitly calculable way.

Proof. Let $\mathcal{L} = \log X$. Using the explicit formula for primes in arithmetic progressions with $T^* = \sqrt{X}\log^2 X$ ($\varrho = \beta + i\gamma = 1 - \delta + i\gamma$ denotes a generic zero of an L-function) we obtain (cf. Davenport [2] § 19) for any a with $(a,q) = 1$, $q \leq Q^*$, $y \leq X$ the relation

$$E'(y,q,a) = -\frac{1}{\varphi(q)}\sum_{\chi(q)}\overline{\chi}(a)\sum_{\substack{\varrho=\varrho_\chi\\ \beta\geq 1/2, |\gamma|\leq T^*}}\frac{y^\varrho}{\varrho} + O\big(\sqrt{y}\big). \tag{8.2}$$

The effect of the last error term is clearly admissible, $O(Q^*\sqrt{X})$ in total. We can classify zeros of all primitive L-functions $\mathrm{mod}\, q \leq Q^*$ up to height T^* into $O(\mathcal{L}^3)$ classes $A(\kappa,\mu,\nu)$ by Hypothesis S, as

$$q \in [Q_\nu/2, Q_\nu) \quad \gamma \in [T_\mu/2, T_\mu) \quad \delta \in \left[\frac{\kappa c_0}{\mathcal{L}}, \frac{(\kappa+1)c_0}{\mathcal{L}}\right) \tag{8.3}$$

where

$$Q_\nu = 2^\nu \leq 2Q^*, \quad T_\mu = 2^\mu \leq 2T^*, \quad \frac{\kappa c_0}{\mathcal{L}} \leq \frac{1}{2} \tag{8.4}$$

with the additional class of index 0: $t \in [0,1) = [0, T_0)$. The triplets κ,μ,ν satisfying (8.4) with $\nu \geq 1$, $\mu \geq 0$, $\kappa \geq 1$ will be called admissible triplets and their set will be denoted by $\mathcal{A}$.

In this case we have clearly by (8.2)

$$\frac{1}{X}\sum_{q\leq Q^*} E^*(X,q) \ll \mathcal{L}^C \max_{\kappa,\mu,\nu\in\mathcal{A}} \frac{N^*(1-\frac{\kappa c_0}{\mathcal{L}}, Q_\nu, T_\mu)}{Q_\nu T_\mu} X^{-c_0\kappa/\mathcal{L}}, \tag{8.5}$$

where

$$N^*(\sigma,Q,T)=\sum_{\substack{\chi(q)\\ \chi\text{ primitive}}}\ \sum_{\substack{\varrho=\varrho_\chi\\ \beta\le\sigma,|\gamma|\le T}}1. \tag{8.6}$$

Thus, in order to show (8.1) it is enough to prove for any triplet δ, Q, T with the property

$$\frac{c_0}{\log(QT)}\le\delta\le\frac{1}{2},\quad 2\le Q\le Q^*,\quad 2\le T\le T^* \tag{8.7}$$

the crucial inequality

$$N^*(1-\delta,Q,T)\ll QTX^\delta e^{-c\sqrt{\log X}}, \tag{8.8}$$

with some positive absolute constant c.

We will use Theorem 12.2 of Montgomery [13]

$$N^*(1-\delta,Q,T)\ll (Q^2T)^{\frac{3\delta}{1+\delta}}(\log QT)^9. \tag{8.9}$$

(We do not need the stronger inequality of Theorem 12.2 of [13] with the exponent $3\delta/(1+\delta)$ replaced by $2\delta/(1-\delta)$ valid for $\delta\le 1/5$.) Since $3\delta\le 1+\delta$, (8.8) will follow from

$$(Q^*)^{\frac{6\delta}{1+\delta}-1}\ll X^\delta e^{-c_2\sqrt{\log X}}\text{ with } c_2=c^*/6. \tag{8.10}$$

Since in the range $0\le\delta\le 1/2$ we have $\frac{6\delta}{1+\delta}-1\le 2\delta$, this is really true by the definition $Q^*=X^{1/2}\exp(-c^*\sqrt{\log X})$, if $\delta\ge 1/12$.

In case of $\delta\le 1/12$ we have by (8.9)

$$N^*(1-\delta,Q,T)\ll (QT)^{1/2}. \tag{8.11}$$

If we have here $QT\ge\exp(\sqrt{\log X})$, then (8.11) directly implies (8.8), since

$$\frac{N^*(1-\delta,Q,T)}{QT}\ll (QT)^{-1/2}\ll\exp\big(-\sqrt{\log X}/2\big). \tag{8.12}$$

On the other hand, if $QT\le\exp\big(\sqrt{\log X}\big)$, then by Hypothesis S (cf. (8.7)) we have $\delta\ge c_0/\sqrt{\log X}$ and therefore we have

$$X^\delta\ge\exp(c_0\sqrt{\log X}), \tag{8.13}$$

which together with (8.11) shows (8.8) and so the Theorem is proved. Q.E.D.

In the applications we will choose our crucial parameter R in the course of proof of (1.1) in Section 10 as

$$R=\sqrt{Q'},\quad Q'=N^{1/2}\exp(-c^*\sqrt{\log N}). \tag{8.14}$$

By this choice of R we can clearly fulfill the later important condition (cf. (11.15))

$$R=N^{1/4-\xi}\quad\text{where}\quad \xi=\frac{c^*}{2\sqrt{\log N}}<\frac{0.3}{\sqrt{K}} \tag{8.15}$$

if the very weak condition

$$K<(5c^*/3)^{-2}\log N \tag{8.16}$$

is satisfied.

However, first we have to show how to improve the estimate for the quantity $\mathcal{E}_K(N)$, which is defined as the sum after $(\log R)^M$ in the first error term in (7.11). We will suppose $\max(k_1,k_2,\ell_1,\ell_2)\le K$. This implies $M\le 4K$.

Using the notation $R^2 = Q'$, we obtain from (8.1) by the trivial estimate $|E(N,q,a)| \le 2q^{-1}N\log N$ (for $q \le N$), Lemmas 4, 5 and by Hölder's inequality with parameters in $\alpha = \nu+1$, $\beta = (\nu+1)/\nu$ (so $\alpha/\beta = \nu$) where $\nu \in \mathbb{Z}^+$, $\nu \ge c'\log(K+1)$, uniformly for $K \le (\log N)/(2C)$

$$|\mathcal{E}_K(N)| \le \sum_{q\le Q'}{}^{\flat} d_K(q)E^*(N,q) \sum_{q=a_1a_2a_{12}} 1 \tag{8.17}$$

$$\begin{aligned}
&\le \sum_{q\le Q'}{}^{\flat} d_K(q)d_3(q)E^*(N,q)\\
&= \sum_{q\le Q'}{}^{\flat} \frac{d_{3K}(q)}{q^{1/\beta}}\cdot q^{1/\beta}E^*(N,q)\\
&\le \biggl(\sum_{q\le Q'}{}^{\flat} \frac{(d_{3K}(q))^{\beta}}{q}\biggr)^{1/\beta}\biggl(\sum_{q\le Q'} q^{\alpha/\beta}(E^*(N,q))^{\alpha}\biggr)^{1/\alpha}\\
&\le \Bigl(CK+\frac12\log N\Bigr)^{CK}(2N\log N)^{\nu/(\nu+1)}\biggl(\sum_{q\le Q'}E^*(N,q)\biggr)^{\frac{1}{\nu+1}}\\
&\ll (\log N)^{CK+1}N\exp\biggl(-\frac{c_1\sqrt{\log N}}{\nu+1}\biggr)\\
&\le N\exp\Bigl((CK+1)\log_2 N - c_1(\nu+1)^{-1}\sqrt{\log N}\Bigr)\\
&\le N\exp\Bigl(-c\sqrt{\log N}/\log(K+1)\Bigr),
\end{aligned}$$

if K satisfies the inequality

$$K\log_2 N < c\sqrt{\log N}/\log K \Leftrightarrow K < \widetilde{c}\sqrt{\log N}/(\log_2 N)^2. \tag{8.18}$$

Supposing (8.18), we have, finally

$$(\log R)^{4K}|\mathcal{E}_K(N)| \le N\exp\bigl(-c\sqrt{\log N}/\log(K+1)\bigr). \qquad \text{Q.E.D.} \tag{8.19}$$

The results proved in this section (for Theorem 2) combined with the analysis made in Section 6 will make possible the proof of the following generalizations of Theorems 1 and 2, where $k_i(N)$ and $\ell_i(N)$ can tend to infinity as $N \to \infty$.

Theorem 1″. *Suppose the conditions of Theorem 1. Suppose further that with a sufficiently small c we have*

$$\max(k_1,k_2,\ell_1,\ell_2) \le K, \quad \min(k_1,k_2) \ge cK, \quad \max(\ell_1,\ell_2) \le c\frac{K\log_2 K}{\log K}, \tag{8.20}$$

with a K satisfying

$$K \le \begin{cases} c\log_2 R/\log_4 R & \text{if } h \ll (\log R)^C\\ c\log_2 R/\log_3 R & \text{if } h \ll R^C. \end{cases} \tag{8.21}$$

Then the result (2.14) holds with the change that the error $o_M(1)$ can be substituted by

$$O\bigl((\log R)^{-2/3}\bigr) \tag{8.22}$$

where the constant implied by the O symbol is effective and absolute (it may depend on $c_0, c, C, \overline{c}, c^*$ however).

Remark. In (8.22) the exponent $2/3$ can be replaced by any $b < 1$, with the implied constant in (8.22) depending on b.

Theorem 2″. *Suppose the conditions of Theorem 2, further* (8.20)–(8.21) *and that Hypothesis S is true with some constant* c_0 *for* $q \le \sqrt{N}$. *Let us require beyond the original assumption of Theorem 2*

$$R \le N^{1/4} \exp\big(- c^* \sqrt{\log N}\big) \tag{8.23}$$

with an arbitrary fixed constant c^*. *Then* (2.16) *holds with* $o_M(1)$ *replaced by* (8.22).

Remark. Similarly to Theorem 1″ we may have

$$O_b\big((\log R)^{-b}\big) \tag{8.24}$$

with any fixed constant $b < 1$ too, in place of (8.22), i.e. $b = 2/3$.

Proof. We have only to check that the requirements in (6.37)–(6.38), (6.55)–(6.57) and (in case of Theorem 2″) (8.18) are satisfied. Now (6.37)–(6.38) are exactly the same as (8.21) while all other mentioned conditions are either distinctly weaker or in the worst case (for $L = 1$) equivalent to (8.21).

9. Another version of Bombieri–Vinogradov's theorem

This section will provide an alternative treatment of the modified Bombieri–Vinogradov theorem. This means that we can avoid Heath-Brown's deep theorem about twin primes and Siegel zeros, and still get effective results of Bombieri–Vinogradov type (with the stronger error term given by (8.1)). Another advantage of this method is that, in contrast to the original Bombieri-Vinogradov's theorem, it makes possible to obtain effective results in those cases (as the difference of $p_{n+\nu} - p_n$ for $\nu > 1$ and the distribution of almost primes), where we are not able to use Heath-Brown's result which refers for consecutive primes (although we do not assert that Heath-Brown's method might not be applicable in some of these problems).

A brief analysis of Hypothesis S and Theorem 3 in Section 8 reveals that

(i) Hypothesis S is fulfilled (with a suitable, effective c_0) apart from the case of real zeros $1 - \delta$ belonging to real L-functions, that is apart the possibly existing Siegel zeros (cf. e.g. Davenport [2], §21).

(ii) Instead of Hypothesis S we need actually (cf. (8.13)) the following weaker version of it with $Y = \exp(\sqrt{\log X})$ to show (8.1) for a given X (if we take into account (i) as well).

Hypothesis S′(Y). *If* $L(1 - \delta, \chi) = 0$ *for a* $\delta > 0$ *and a real primitive character* $\chi \pmod q$, $q \le Y$, *then*

$$\delta > \frac{1}{3 \log Y}, \tag{9.1}$$

for $Y > C_0$, *an explicitly calculable absolute constant.*

We note that we have the effective unconditional estimate ([4], [14]), valid for $q > q_0$:

$$\delta \ge \frac{1}{\sqrt{q}}. \tag{9.2}$$

A further observation (similar to that of Maier [12]) is that by Landau's theorem (cf. [2], §14 with a constant c in place of $1/3$, or Pintz [15] with $(1/2+o(1))/\log Y$) for any given Y there is at most one real primitive character χ_1 which does not fulfill (9.1). This makes possible to turn Hypothesis S′(Y) into a theorem, valid for a sequence $Y = Y_n \to \infty$ (for $n > n_0$ explicitly calculable absolute constant) with

$$Y_n \leq \exp\left(\sqrt{Y_{n-1}}\right). \tag{9.3}$$

In order to show this, suppose that (9.1) is false for a sufficiently large Y', i.e. by (9.2) there exists a $\chi_1 \bmod q_1 \leq Y'$ such that $L(1-\delta_1, \chi_1) = 0$ with

$$\frac{1}{\sqrt{Y'}} \leq \min\left(\frac{1}{\sqrt{q_1}}, c_0\right) \leq \delta_1 \leq \frac{1}{3\log Y'}. \tag{9.4}$$

Let us choose $\widetilde{Y} \geq Y'$ in such a way, that

$$\widetilde{Y} = \exp\left(\frac{1}{3\delta_1}\right) \Leftrightarrow \delta_1 = \frac{1}{3\log \widetilde{Y}}. \tag{9.5}$$

Then for any other zero $1-\delta_2$ belonging to a real primitive $\chi_2 \bmod q_2$, $q_2 \leq \widetilde{Y}$, we have by Landau's theorem in the version of Pintz [15]

$$\max(\delta_1, \delta_2) > \frac{1}{3\log \widetilde{Y}} \Leftrightarrow \delta_2 > \frac{1}{3\log \widetilde{Y}}. \tag{9.6}$$

Now, (9.4)–(9.6) show that (9.1) is true for a value $Y = \widetilde{Y}$ satisfying

$$Y' \leq \widetilde{Y} \leq \exp(\sqrt{Y'}). \tag{9.7}$$

We can formulate this as

Lemma 7. *Hypothesis S′(Y) holds for a sequence $Y_n \to \infty$ with* (9.3) *if $Y_0 > C_0$, an explicitly calculable absolute constant.*

Choosing $X_n = \exp(\log^2 Y_n) \Leftrightarrow Y_n = \exp(\sqrt{\log X_n})$ we obtain the required other alternative form of an unconditional, effective Bombieri–Vinogradov type theorem, valid for a sequence $X = X_n \to \infty$.

Theorem 4. *There exists a sequence $X_n \to \infty$ with*

$$X_n \leq \exp\left(\exp\sqrt{\log X_{n-1}}\right), \tag{9.8}$$

$X_0 > C_0'$, an explicitly calculable absolute constant, such that for $Q_n = X_n^{1/2}\exp\left(-c^*\sqrt{\log X_n}\right)$ we have

$$\sum_{q \leq Q_n} E^*(X_n, q) \ll X_n \exp\left(-c_1\sqrt{\log X_n}\right). \tag{9.9}$$

Proof. The proof is the same as that of Theorem 3 with the only change that (8.13) is true by Lemma 7 for $X = X_n = \exp(\log^2 Y_n)$.

10. The sum of the singular series $\mathfrak{S}(\mathcal{H})$

Let

$$B_h(k) = B(k) = \sum_{|\mathcal{H}|=k} \mathfrak{S}(\mathcal{H}), \tag{10.1}$$

where all sets $\mathcal{H} = \{h_1, h_2, \dots, h_k\}$, $h_i \in [1, h]$ are counted with $k!$ multiplicity according to all possible permutations of h_i.

By Gallagher's theorem [3] we have for fixed k as $h \to \infty$

$$B_h(k) = h^k\big(1 + O_{k,\varepsilon}(h^{-\frac{1}{2}+\varepsilon})\big). \tag{10.2}$$

This is not uniform in k but up to some level $k \le f(h)$ one could show still $B_h(k) \sim h^k$. However, we will use here a completely different approach. We do not prove (10.2), just (see Lemma 8) the weaker relation that $B_h(k)/h^k$ is (apart from a factor $(1 + o(1))$ non-decreasing as a function of k, at least as long as $k = o(h/\log h)$). This result is completely sufficient for our purposes, fortunately.

Let c be an arbitrary small constant, h, z and Z sufficiently large,

(10.3)

$$k \le \sqrt{z},\ \ h \le z \le (1-c)\log Z,\ \ P(z) = \prod_{p \le z} p,\ \ V = V_z = \prod_{p \le z}\left(1 - \frac{1}{p}\right)^{-1} \sim e^{\gamma}\log z.$$

$$Q := Q_z := \{n; (n, P(z)) = 1\}, \quad M := \sum_{1 \le n \le Z, n \in Q} 1 = \frac{Z}{V} + O(P(z)). \tag{10.4}$$

Then we have for a fixed set $\mathcal{H}$ consisting of k distinct elements $h_i \in [1, h]$, similarly to Section 5, the average number of z-quasi-prime tuples of pattern $\mathcal{H}$:

$$\begin{aligned} R(\mathcal{H}) &= \frac{1}{Z}\sum_{\substack{n=1\\ P_{\mathcal{H}}(n) \in Q}}^{Z} 1 = \frac{1}{Z}\sum_{n=1}^{Z}\sum_{d \mid (P(z), P_{\mathcal{H}}(n))} \mu(d) = \frac{1}{Z}\sum_{d \mid P(z)} \mu(d) \sum_{\substack{n=1\\ d \mid P_{\mathcal{H}}(n)}}^{Z} 1 \\ &= \frac{1}{Z}\sum_{d \mid P(z)} \mu(d)\nu_d(\mathcal{H})\left(\frac{Z}{d} + O(1)\right) = \sum_{d \mid P(z)} \frac{\mu(d)\nu_d(\mathcal{H})}{d} + O\Big(\sum_{d \mid P(z)} \nu_d(\mathcal{H})\Big) Z^{-1} \\ &= \prod_{p \le z}\left(1 - \frac{\nu_p(\mathcal{H})}{p}\right) + O\Big(\prod_{p \le z}(1 + k)\Big) Z^{-1} \\ &= V^{-k}\prod_{p \le z}\left(1 - \frac{\nu_p(\mathcal{H})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k} + O\Big(\exp(1 + o(1))\frac{z\log(k+1)}{\log z}\Big) Z^{-1} \\ &= V^{-k}\mathfrak{S}(\mathcal{H}) e^{O\left(\sum_{p > z} \frac{k^2}{p^2}\right)} + O(Z^{-1/3}) \\ &= V^{-k}\mathfrak{S}(\mathcal{H})\Big(1 + O\Big(\frac{1}{\log z}\Big)\Big) + O(Z^{-1/3}) \end{aligned} \tag{10.5}$$

uniformly in k, h, z, Z satisfying (10.3), if c is fixed. Let further

$$S^*(k) = \frac{1}{h^k}\sum_{|\mathcal{H}|=k} \mathfrak{S}(\mathcal{H}) = \frac{B_h(k)}{h^k}. \tag{10.6}$$

Lemma 8. *If* $k < h(\log z)^{-2}$, $z \le \max(h, k^2)$, *then*

$$S^*(k+1) \ge S^*(k)\Big(1 + O\Big(\frac{1}{\log z}\Big)\Big). \tag{10.7}$$

Proof. For $i \in [1, Z]$ let

$$f_i = \sum_{i+1 \le m \le i+h, m \in Q} 1,\ a_i = a_i(k) = f_i(f_i - 1)\dots(f_i - k + 1). \tag{10.8}$$

Then $a_i(k)$ is clearly the number of all k-tuples of z-quasiprimes of type $i + h_\nu$ ($\nu = 1, \dots, k$, $1 \le h_\nu \le h$, h_ν distinct), calculated with all $k!$ permutations, while f_i is clearly the number of z-quasiprimes in the interval $[i+1, i+h]$. We have obviously for every pair $i, j \in [1, h]$

$$f_i \ge f_j \Leftrightarrow a_i \ge a_j, \tag{10.9}$$

therefore

$$\frac{1}{Z}\sum_{i=1}^{Z} a_i f_i \ge \frac{\sum_{i=1}^{Z} f_i}{Z}\frac{\sum_{i=1}^{Z} a_i}{Z}. \tag{10.10}$$

The above formula follows from

$$2\left(Z\sum_{i=1}^{Z} a_i f_i - \sum_{i=1}^{Z} f_i \sum_{i=1}^{Z} a_i\right) = \sum_{i=1}^{Z}\sum_{j=1}^{Z}(f_i - f_j)(a_i - a_j) \ge O. \tag{10.11}$$

We have further $a_i(k+1) = a_i(k)(f_i - k) = a_i f_i - k a_i$ and by calculating in two different ways how many times all pairs $n, \mathcal{H}$ satisfy the relation $P_{\mathcal{H}}(n) \in Q$ we obtain

$$Z^{-1}\sum_{n=1}^{Z} a_n(k) = Z^{-1}\sum_{n=1}^{Z}\sum_{\substack{|\mathcal{H}|=k \\ P_{\mathcal{H}}(n)\in Q}} 1 = Z^{-1}\sum_{|\mathcal{H}|=k}\sum_{\substack{n=1 \\ P_{\mathcal{H}}(n)\in Q}}^{Z} 1 = \sum_{|\mathcal{H}|=k} R(\mathcal{H}), \tag{10.12}$$

$$\frac{1}{Z}\sum_{i=1}^{Z} f_i = \frac{1}{Z}(hM + O(h^2)) = \frac{h}{V} + O(Z^{-c/2}). \tag{10.13}$$

Thus (10.10) and (10.12) imply by $a_i f_i = a_i(k+1) + k a_i$ that

$$\frac{1}{Z}\sum_{i=1}^{Z} a_i(k+1) + k \cdot \frac{1}{Z}\sum_{i=1}^{Z} a_i(k) \ge \left(\frac{h}{V} + O(Z^{-c/2})\right) \cdot \frac{1}{Z}\sum_{i=1}^{Z} a_i(k). \tag{10.14}$$

Hence, using (10.12), we obtain

$$\sum_{|\mathcal{H}|=k+1} R(\mathcal{H}) \ge \left(\frac{h}{V} - 2k\right)\sum_{|\mathcal{H}|=k} R(\mathcal{H}). \tag{10.15}$$

Multiplying by V^{k+1} on both sides, we obtain by (10.5)

$$\sum_{|\mathcal{H}|=k+1} \mathfrak{S}(\mathcal{H}) \ge h\left(1 + O\left(\frac{kV}{h}\right) + O\left(\frac{1}{\log z}\right)\right)\sum_{|\mathcal{H}|=k} \mathfrak{S}(\mathcal{H}). \tag{10.16}$$

Now dividing by h^{k+1} on both sides we have by $k < h(\log z)^{-2}$

$$S^*(k+1) \ge \left(1 + O\left(\frac{1}{\log z}\right)\right) S^*(k). \tag{10.17}$$

By choosing $z = h^2$ we obtain from Lemma 8 the following

Corollary 5. *If $k < h/(4\log^2 h)$, then*

$$S^*(k+1) \geq S^*(k)\left(1 + O\left(\frac{1}{\log h}\right)\right).$$

We note the following corollary too, which is, however, not needed in the sequel.

Corollary 6. *We have*

$$S^*(k) \geq e^{-ck/\log z}. \tag{10.18}$$

Consequently,

$$S^*(k) \geq 1 + o(1) \tag{10.19}$$

if

$$k = o(h^{1/3}), \tag{10.20}$$

by choosing parameter z as

$$z = \exp(h^{1/3}). \tag{10.21}$$

Proof. Since for sets $\mathcal{H}$ of size 1 we have $\nu_p(\mathcal{H}) = 1$ for all primes, we have clearly for all $\mathcal{H}$

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{p}\right)^{-1} = 1 \tag{10.22}$$

and so

$$S^*(1) = h^{-1} \sum_{|\mathcal{H}|=1} \mathfrak{S}(\mathcal{H}) = 1. \tag{10.23}$$

Now the result follows from (10.17). Q.E.D.

Remark. With a somewhat more careful treatment in Lemma 8, the condition (10.20) may be weakened to $k = o(\sqrt{h/\log h})$.

11. Proof of (1.1)

In order to show our results about small gaps between primes we will examine the average (2.17) for the proofs of Theorems A and B.

(2.17) can clearly be composed from

$$M_R(N, K, \ell) = \frac{1}{Nh^{2K}} \sum_{n=N+1}^{2N} (\psi_R(K, \ell, n, h))^2 \tag{11.1}$$

and

$$\widetilde{M}_R(N, K, \ell) = \frac{1}{Nh^{2K+1}} \sum_{n=N+1}^{2N} \left(\sum_{p=n+1}^{n+h} \log p\right)(\psi_R(K, \ell, n, h))^2. \tag{11.2}$$

For simplicity we will consider first the case when $\nu = 1$ and K and ℓ are arbitrarily large, but bounded, $h, N, R \to \infty$.

For convenience we agree that in the definition (2.17) we consider every set of size k with a multiplicity $k!$ according to all permutations of the elements $h_i \in \mathcal{H}$, unless mentioned otherwise. Let us consider now (11.1) and let us group the pairs of sets $\mathcal{H}_1, \mathcal{H}_2$ according to the size of their intersection $r = |\mathcal{H}_1 \cap \mathcal{H}_2|$. In this case clearly $|\mathcal{H}| = |\mathcal{H}_1 \cup \mathcal{H}_2| = 2K - r$. Let us choose now a set $\mathcal{H}$ and here exceptionally we disregard the permutation of the elements in $\mathcal{H}$. (However at $\mathcal{H}_1, \mathcal{H}_2$ we take into account all permutations.) Let further

$$\lambda = \frac{h}{\log 3N}, \quad \Theta = \frac{\log R}{\log 3N}, \quad x = \frac{\log R}{h} = \frac{\Theta}{\lambda}. \tag{11.3}$$

Given the set $\mathcal{H}$ of size $2K - r$ we can choose $\mathcal{H}_1$ in $\binom{2K-r}{K}$ ways. Afterwards we can choose the intersection set in $\binom{K}{r}$ ways, finally we can arrange the elements both in $\mathcal{H}_1$ and $\mathcal{H}_2$ in $K!$ ways. This gives

$$\binom{2K-r}{K} \cdot \binom{K}{r} (K!)^2 = \binom{K}{r} \frac{(2K-r)!(K!)^2}{K!(K-r)!} = (2K-r)! \binom{K}{r}^2 \cdot r! \tag{11.4}$$

choices for $\mathcal{H}_1$ and $\mathcal{H}_2$, taking into account the permutation of the elements in $\mathcal{H}_1, \mathcal{H}_2$.

If we consider in the summation every union set $\mathcal{H}$ of size k just once, independently of the arrangement of the elements then Gallagher's theorem (10.2) may be formulated as ($\sum^*$ indicates we consider every set just once)

$$\sideset{}{^*}\sum_{|\mathcal{H}|=k} \mathfrak{S}(\mathcal{H}) = \frac{h^k}{k!} \big(1 + O(h^{-\frac{1}{2}+\varepsilon})\big). \tag{11.5}$$

Applying this we obtain by (11.4) and Theorem 2 (or 2″), taking into account the Remark after Theorem 2′,

$$\widetilde{M}_R(N, K, \ell) \sim \frac{1}{Nh^{2K+1}} \sum_{r=0}^{K} (2K-r)! \binom{K}{r}^2 \cdot r! \cdot Z_r, \tag{11.6}$$

where abbreviating $a = \frac{2\ell+1}{\ell+1} = \binom{2\ell+1}{\ell+1}\binom{2\ell}{\ell}^{-1} = \binom{2\ell+2}{\ell+1}\binom{2\ell}{\ell}^{-1} \cdot \frac{1}{2}$ we have

$$Z_r := \binom{2\ell}{\ell} \frac{(\log R)^{2\ell+r}}{(r+2\ell)!} \Bigg\{ r \sideset{}{^*}\sum_{|\mathcal{H}|=2K-r} 2a \frac{\log R}{r+2\ell+1} \mathfrak{S}(\mathcal{H}) N \tag{11.7}$$

$$+ (2K-2r) \sideset{}{^*}\sum_{|\mathcal{H}|=2K-r} a \cdot \frac{\log R}{r+2\ell+1} \mathfrak{S}(\mathcal{H}) N + \sideset{}{^*}\sum_{|\mathcal{H}|=2K-r} \sum_{\substack{h_0=1\\ h_0 \notin \mathcal{H}}}^{h} \mathfrak{S}(\mathcal{H}^0) N \Bigg\}$$

$$\sim N \binom{2\ell}{\ell} \frac{(\log R)^{2\ell+r}}{(r+2\ell)!} \left\{ \frac{h^{2K-r}}{(2K-r)!} \cdot \frac{2aK \log R}{r+2\ell+1} + \frac{2K-r+1}{(2K-r+1)!} h^{2K-r+1} \right\},$$

where in the last sum we took into account which element of $\mathcal{H}^0$ is h_0, which can be chosen in $2K - r + 1$ ways. So we obtain

(11.8)

$$\widetilde{M}_R(N, K, \ell) \sim \binom{2\ell}{\ell} (\log R)^{2\ell} \sum_{r=0}^{K} \binom{K}{r}^2 \cdot \frac{x^r}{(r+1)\dots(r+2\ell)} \left(\frac{2aK}{r+2\ell+1} x + 1 \right).$$

Similarly we get from Theorem 1 or 1″

(11.9)
$$M_R(N,K,\ell) \sim \frac{1}{Nh^{2K}} \sum_{r=0}^{K} (2K-r)! \binom{K}{r}^2 r! \cdot \binom{2\ell}{\ell} \frac{(\log R)^{2\ell+r}}{(r+2\ell)!} \cdot N \sum_{|\mathcal{H}|=2K-r}^{*} \mathfrak{S}(\mathcal{H})$$
$$\sim \binom{2\ell}{\ell} (\log R)^{2\ell} \sum_{r=0}^{K} \binom{K}{r}^2 \frac{x^r}{(r+1)\dots(r+2\ell)}.$$

Now (11.8) and (11.9) together imply for the crucial quantity in (2.17) in case of $\nu = 1$

(11.10)
$$\begin{aligned} S_R(N,K,\ell,1) &= \frac{1}{Nh^{2K+1}} \sum_{n=N+1}^{2N} \Big(\sum_{p=n+1}^{n+h} \log p - \log 3N \Big) (\psi_R(K,\ell,n,h))^2 \\ &= \widetilde{M}_R(N,K,\ell) - \frac{\log 3N}{h} M_R(N,K,\ell) \\ &\sim \binom{2\ell}{\ell} (\log R)^{2\ell} P_{K,\ell}(x) \end{aligned}$$

where writing further $\varphi = 1/(\ell+1)$ (that is, $a = 2 - \varphi$)

$$P_{K,\ell}(x) = \sum_{r=0}^{K} \binom{K}{r}^2 \frac{x^r}{(r+1)\dots(r+2\ell)} \left(1 + x \left(\frac{4(1-\frac{\varphi}{2})K}{r+2\ell+1} - \frac{1}{\Theta} \right) \right). \tag{11.11}$$

Let us choose now a sufficiently large ℓ and K so that

$$K = 16(\ell+1)^2 = 16\varphi^{-2} \ (\ell > \ell_0 \Leftrightarrow \varphi < \varphi_0) \tag{11.12}$$

In our present case as mentioned earlier K and ℓ will be arbitrarily large.

Let $x = K/100$, $r_0 = (1-2\varphi)K$, $r_1 = (1-\varphi)K$, $\Theta = \frac{1}{4} - \xi$, so $\Theta^{-1} = \frac{4}{1-4\xi} \le 4(1+5\xi)$ if $\xi \le 0.05$, which we will suppose.

By (11.12) we have $K\varphi = 16\varphi^{-1}$, so

$$r_1 + 2\ell + 1 = \left(1 - \varphi + \frac{2}{K\varphi}\right) K - 1 = \left(1 - \frac{7}{8}\varphi\right) K - 1, \tag{11.13}$$

and therefore we get for $r \le r_1$

$$\frac{4(1-\frac{\varphi}{2})K}{r+2\ell+1} - \frac{1}{\Theta} \ge 4\left(1 + \frac{3}{8}\varphi\right) - 4(1+5\xi) > 0 \tag{11.14}$$

if $3\varphi > 40\xi$. Consequently (11.14) holds if

$$\frac{12}{\sqrt{K}} > 40\xi \Longleftrightarrow \xi < \frac{0.3}{\sqrt{K}} \ \text{ and } \ r \le r_1. \tag{11.15}$$

Now let us compare the term

$$\binom{K}{r}^2 \frac{x^r}{(r+1)\dots(r+2\ell)} = f(r) \tag{11.16}$$

for

$$r = r_0 = (1-2\varphi)K \ \text{ and any } \ r_2 > r_1 = (1-\varphi)K. \tag{11.17}$$

This covers all terms which may be in the negative range with respect to the factor in (11.14). We have by $\varphi^2 = 16/K$, $x = K/100$, φ sufficiently small,

(11.18)
$$\frac{f(r_2)}{f(r_0)} < \prod_{r_0 \le r < r_2} \left(\left(\frac{K-r}{r+1} \right)^2 x \right) < \left(\frac{(2\varphi)^2}{(1-2\varphi)^2} x \right)^{\varphi K} < \left(\frac{65}{K} x \right)^{4\sqrt{K}} < e^{-\sqrt{K}},$$

so the total contribution of the negative terms is (since the last factor is at most $1 + x$) in absolute value

$$< K \left(\frac{K}{100} + 1 \right) e^{-\sqrt{K}} f(r_0) < e^{-\sqrt{K}/2} f(r_0), \tag{11.19}$$

which means

$$P_{K,\ell}(x) > 0 \Longleftrightarrow S_R(N, K, \ell, 1) > 0, \tag{11.20}$$

if (11.15) holds. This means that we have to choose

$$\xi < \frac{0.3}{\sqrt{K}} \Longleftrightarrow \Theta > \frac{1}{4} - \frac{0.3}{\sqrt{K}}. \tag{11.21}$$

This does not make any problem if K is an arbitrary large but fixed constant, in which case we obtain that the choice

$$x = \frac{\log R}{h} = \frac{\Theta}{\lambda} = \frac{\frac{1}{4} - \xi}{\lambda} = \frac{K}{100} \tag{11.22}$$

is admissible, that is we have by $\xi < 1/\sqrt{K}$

$$\lambda = \frac{25 + O(1/\sqrt{K})}{K} \Longleftrightarrow h = \frac{25 \log N (1 + O(1/\sqrt{K}))}{K} \tag{11.23}$$

for arbitrary K, which is equivalent with

$$\varliminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} = 0. \tag{11.24}$$

In order to show the stronger result (1.1) the main point is that according to Theorems $1''$ and $2''$ we may choose our K as large as

$$K = c \frac{\log_2 R}{\log_4 R} \sim c \frac{\log_2 N}{\log_4 N} \tag{11.25}$$

(with the restriction $\sqrt{K}/4 \in \mathbb{Z}$) which will lead finally to the admissible choice of

$$h = \frac{C \log N}{\log_2 N} \log_4 N \tag{11.26}$$

as stated in (1.1). (We remark here that as $\ell_1 = \ell_2 = \ell = 4\sqrt{K} - 1$ the weak bound for ℓ_i in (8.20) will be satisfied.) Formulas (11.25)–(11.26) mean that the condition of Corollary 5 in Section 10 is satisfied so we may use it in place of Gallagher's theorem (11.5) and so we obtain with the notation $\eta_1 = (\log R)^{-\frac{2}{3}} < \eta_2 = (\log h)^{-1}$

and (10.6) for the resulting analogue $P^*_{K,\ell}(x)$ of $P_{K,\ell}(x)$ the inequality:

$$\begin{aligned}P^*_{K,\ell}(x) &= \sum_{r=0}^{K}\binom{K}{r}^2 \frac{1}{(r+1)\dots(r+2\ell)}\bigg\{S^*(2K-r+1)x^r(1+O(\eta_1))\\ &\quad + S^*(2K-r)x^{r+1}\left(\frac{4\left(1-\frac{\varphi}{2}\right)K}{r+2\ell+1}-\frac{1}{\Theta}+O(\eta_1)\right)\bigg\}\\ &\geq \sum_{r=0}^{K}\binom{K}{r}^2 S^*(2K-r)\frac{x^r}{(r+1)\dots(r+2\ell)}\times\\ &\quad\times\left(1+O(\eta_2)+x\left(\frac{4\left(1-\frac{\varphi}{2}\right)K}{r+2\ell+1}-\frac{1}{\Theta}+O(\eta_1)\right)\right).\end{aligned} \tag{11.27}$$

First we can note that by the choice $R = N^{1/4-\xi}$, $\xi = c/\sqrt{\log N}$ we have the terms with $r \leq r_1 = (1-\varphi)K$ again positive, since similarly to (11.14) we have now again

$$\frac{4(1-\frac{\varphi}{2})K}{r+2\ell+1} - \frac{1}{\Theta} + O(\eta_1) \geq 4\left(1+\frac{3}{8}\varphi\right) - \frac{4}{1-4\xi} + O(\xi^{\frac{4}{3}}) > 0 \tag{11.28}$$

if $3\varphi/8 > 5\xi$, $N > N_0$, that is, if (11.15) is satisfied for K and ξ, which is clearly the case by (11.25). So we can again compare all eventually negative terms of indices $r_2 \geq r_1$ with that of index r_0.

Since the only change compared to (11.11) is the appearance of the additional factor $S^*(2k-r)$ in the r^{th} term which is monotonically decreasing for increasing values of r up to a factor $1+O(\eta_2)$, the critical inequalities (11.18) and (11.19) remain valid with an eventual loss of $1+O(\eta_2)$ in every term in the product in (11.18), which means that the formula

$$\frac{(2\varphi)^2 x}{(1-2\varphi)^2}(1+o(\eta_2)) < \frac{65}{K}x \tag{11.29}$$

remains valid for any choice $\eta_2 < c$ that is $h > h_0$ already and thereby all formulas (11.19)–(11.23) remain valid. The choice of K in (11.25) thus really leads to the required result for h in (11.26) which proves (1.1).

12. Chains of small gaps between consecutive primes

This section will be devoted to the proof of Theorem C. Let $\nu \geq 2$ and $\vartheta_0 \in [1/2, 1]$, be fixed, with $\vartheta_0 < 1$ in case of $\nu = 2$. (The case $\nu = 2$, $\vartheta_0 = 1$, $E_2 = 0$ is already covered by Theorem B.) We will work similarly to Section 11. We will choose first ε_0 as a sufficiently small fixed positive number. Then we will use a choice of K and ℓ, similar to (11.12), but we will choose now ℓ sufficiently large, depending on ν, ϑ_0, ε_0, so we will have now

$$K = \ell^2 = \varphi^{-2}, \qquad \ell > \ell_0(\nu, \vartheta_0, \varepsilon_0) \Leftrightarrow \varphi < \varphi_0(\nu, \vartheta_0, \varepsilon_0). \tag{12.1}$$

Further we choose, similarly to (11.3)

$$\Theta = \frac{\log R}{\log 3N} = \frac{\vartheta_0(1-\varphi)}{2}, \quad \lambda = \frac{h}{\log 3N}, \quad x = \frac{\log R}{h} = \frac{\Theta}{\lambda} \tag{12.2}$$

and according to our present situation, analogously to (11.11), let

$$(12.3)\quad P_{K,\ell,\nu}(x)=\sum_{r=0}^{K}\binom{K}{r}^2\frac{x^r}{(r+1)\dots(r+2\ell)}\left(1+x\left(\frac{4\left(1-\frac{\varphi}{2}\right)K}{r+2\ell+1}-\frac{\nu}{\Theta}\right)\right).$$

Let further, in addition to (11.16) and (12.2)

$$(12.4)\qquad y=\frac{1}{x}=\frac{h}{\log R}=\frac{\lambda}{\Theta},\qquad z=\sqrt{y},$$

$$(12.5)\qquad g(r)=\binom{K}{r}^2x^r=\left(\binom{K}{r}z^{-r}\right)^2,\quad j(r)=\frac{g(r)}{g(r-1)}.$$

If we examine the function $j(r)$ we can see that

$$(12.6)\qquad j(r)=\left(\frac{K-r+1}{r}z^{-1}\right)^2\gtreqless 1\Leftrightarrow r\lesseqgtr\frac{K+1}{z+1}.$$

Thus the crucial part of $f(r)$ in (11.16), our function $g(r)$ is monotonically increasing for $r\le(K+1)/(z+1)$, monotonically decreasing for $r\ge(K+1)/(z+1)$ and the maximum appears at $r=\left[\frac{K+1}{z+1}\right]$ or $r=\left\lceil\frac{K+1}{z+1}\right\rceil$.

Since our function on the right-hand side of (1.12) is positive, that is $a(\nu,\vartheta_0):=\left(\sqrt{\nu}-\sqrt{2\vartheta_0}\right)^2>0$, choosing ℓ and therefore K sufficiently large we will ensure that we have to consider intervals of length $h=\lambda\log 3N$, $\lambda=\lambda_0(1+o(1))$, $\lambda_0:=a(\nu,\vartheta_0)$. Let

$$(12.7)\quad z:=z_0(1+\varepsilon_0):=b(\nu,\vartheta_0)(1+\varepsilon_0),\quad z_0=b=\sqrt{2a/\vartheta_0}=\sqrt{2\nu/\vartheta_0}-2.$$

Consequently the maximum of $g(r)$ appears strictly inside the range $r\in[1,K]$:

$$(12.8)\qquad \frac{K+1}{z+1}\sim d(\nu,\vartheta_0)K,\quad d:=\left(\sqrt{\frac{2a}{\vartheta_0}}+1\right)^{-1}\in(0,1).$$

Now, an easy calculation shows that with the above choice of the variables we have (cf. (12.7))

$$(12.9)\qquad 1+\frac{1}{z_0^2}\left(\frac{4K}{\frac{K}{z_0+1}}-\frac{2\nu}{\vartheta_0}\right)=0\Leftrightarrow(z_0+2)^2=\frac{2\nu}{\vartheta_0}.$$

Let us choose now

$$(12.10)\qquad r_0=\left[\frac{K+1}{z_0+1}\right],\quad r_1=r_0+\varphi K=r_0+\ell.$$

Then, by the choice $z=z_0(1+\varepsilon_0)$ in (12.7), the crucial expression in $P_{K,\ell,\nu}(x)$ will be by (12.9) for $r\le r_1$

$$(12.11)\qquad \begin{aligned}&1+x\left(\frac{4(1-\varphi/2)K}{r+2\ell+1}-\frac{\nu}{\Theta}\right)\\ &=1+\frac{1}{z_0^2(1+\varepsilon_0)^2}\left(\frac{4K(1+O(\varphi))}{\frac{K}{z_0+1}+O(K\varphi)}-\frac{2\nu}{\vartheta_0(1-\varphi)}\right)\\ &=1+\frac{-z_0^2+O(\sqrt{\nu}\varphi)+(\nu\varphi)}{z_0^2(1+\varepsilon_0)^2}>c(\nu,\vartheta_0)\varepsilon_0\ \text{ if }\ \varphi<\varphi_0(\nu,\vartheta,\varepsilon_0),\end{aligned}$$

with a positive constant $c(\nu, \vartheta_0)$. On the other hand, for $r_2 > r_1$ we have, similarly to (11.18), by (12.6) and (12.10) using the notation (11.16),

$$\frac{f(r_2)}{f(r_0)} < \prod_{r_0 < r \le r_2} \left(\frac{K+1-r}{r} \cdot \frac{1}{z} \right)^2 < \prod_{r_0 < r \le r_1} \left(\frac{K+1-r}{r} \cdot \frac{1}{z} \right)^2 < \left(\left(\frac{K+1}{r_0} - 1 \right) \frac{1}{z} \right)^{2\ell} \le \left(\frac{z_0+1-1}{z_0(1+\varepsilon_0)} \right)^{2\ell} < e^{-\varepsilon_0 \ell}. \tag{12.12}$$

Thus the total contribution of the negative terms will be for sufficiently large ℓ at most

$$K \left(1 + \frac{4(K+\nu)}{z^2} \right) e^{-\varepsilon_0 \ell} f(r_0) < e^{-\varepsilon_0 \ell/2} f(r_0), \tag{12.13}$$

while that of the single term r_0 will be by (12.11) at least

$$c(\nu, \vartheta_0) \varepsilon_0 f(r_0) > e^{-\varepsilon_0 \ell/2} f(r_0) \quad \text{if } \ell > \ell_0(\nu, \vartheta_0, \varepsilon_0). \tag{12.14}$$

This shows, similarly to (11.20)

$$P_{K,\ell,\nu}(x) > 0 \Leftrightarrow S_R(N, K, \ell, \nu) > 0, \tag{12.15}$$

and so, by the notation (2.17) we must have at least ν primes in some interval of type

$$[n+1, n+h] = [n+1, n+\lambda \log N], \quad n \in [N+1, 2N] \tag{12.16}$$

where by (12.4), (12.7)

$$\lambda = \Theta y = \Theta z^2 < \frac{\vartheta_0}{2} z_0^2 (1+\varepsilon_0)^2 = (1+\varepsilon_0)^2 \left(\sqrt{\nu} - \sqrt{2\vartheta_0} \right)^2. \tag{12.17}$$

Since ε_0 can be chosen arbitrarily small, this proves Theorem C.

References

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

[2] H. Davenport, *Multiplicative Number Theory*, Second Edition, Revised by Hugh L. Montgomery, Springer, Berlin, Heidelberg, New York, 1980.

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

[4] D. M. Goldfeld and A. Schinzel, On Siegel's zero, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **2** (1975), 571–583.

[5] D. A. Goldston, C. Y. Yildirim, Small gaps between primes I, preprint.

[6] H. Halberstam and H. -E. Richert, *Sieve methods*, Academic Press, London, New York, 1975.

[7] 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.

[8] D. R. Heath-Brown, Almost-prime k-tuples, *Mathematika* **44** (1997), 245–266.

[9] D. R. Heath-Brown, Prime twins and Siegel zeros, *Proc. London Math. Soc.* (3) **47** (1983), no. 2, 193–224.

[10] M. N. Huxley, On the differences of primes in arithmetical progressions, *Acta Arith.* **15** (1968/69), 367–392.

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

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

[13] H. L. Montgomery, *Topics in Multiplicative Number Theory*, Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, New York, 1971.

[14] J. Pintz, Elementary methods in the theory of $\mathcal{L}$-functions II, On the greatest real zero of a real $\mathcal{L}$-function, *Acta Arith.* **31** (1976), 273–289.

[15] J. Pintz, Elementary methods in the theory of $\mathcal{L}$-functions VIII, Real zeros of real $\mathcal{L}$-functions, *Acta Arith.* **33** (1977), 89–98.
[16] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, Second edition, Edited and with a preface by D. R. Heath-Brown, The Clarendon Press, Oxford University Press, New York, 1986.

Department of Mathematics, San Jose State University, San Jose, CA 95192, USA
E-mail address: goldston@math.sjsu.edu

Rényi Mathematical Institute of the Hungarian Academy of Sciences, H-1364 Budapest, P.O.B. 127, Hungary
E-mail address: pintz@renyi.hu

Department of Mathematics, Boğaziçi University, Bebek, Istanbul 34342 Turkey & Feza Gürsey Enstitüsü, Çengelköy, Istanbul, P.K. 6, 81220 Turkey
E-mail address: yalciny@boun.edu.tr