The Goldston–Pintz–Yıldırım sieve and maximal gaps

by

Hakan Ali-John Seyalioglu (Los Angeles, CA)

1. Introduction. One field of particular interest in Number Theory concerns the gaps between consecutive primes. Within the last few years, very important results have been achieved on how small these gaps can be. The strongest of these results were obtained by Dan Goldston, János Pintz and Cem Yalçın Yıldırım. The present work begins by generalizing their results so that they can be applied to related problems in a more direct manner. Additionally, we improve the bound for F_2 (concerning the maximal gap in a block of three primes) obtained by the results of [4] with our generalization.

1.1. *Previous work.* The first result of Goldston, Pintz and Yıldırım [4] states (^{1})

$$\Delta_r = \liminf_{n \to \infty} \frac{p_{n+r} - p_n}{\log p_n} \le (\sqrt{r} - 1)^2$$

and in particular $\Delta_1 = 0$. Using related methods and incorporating ideas from Maier's matrix method [8], the authors were able to later improve this result by a factor of $e^{-\gamma}$ [5]. In proving this result, they developed a new sieve method, based closely on that of Selberg, to estimate the number of primes within an interval. Let $\mathcal{H} = \{h_1, \ldots, h_k\}$ and $P(n, \mathcal{H}) = (n+h_1) \ldots (n+h_k)$. Using the notation of [3], for $l \ge 0$, we set

$$\Lambda_R(n, \mathcal{H}, k+l) = \frac{1}{(k+l)!} \sum_{\substack{d \mid P(n,\mathcal{H}) \\ d \le R}} \mu(d) \log\left(\frac{R}{d}\right)^{k+l}.$$

The main results of [4] follow from two related estimates, given below. Throughout this paper C and c are absolute constants which may differ at

(^{1}) In earlier work [6], the constant Δ_r is referred to as E_r. However, this has been dropped in more recent works [4, 5] to avoid a notational conflict.

every occurrence. If an implied constant depends on a value this dependence will be indicated by a subscript (for example, $\ll_M$, $o_M(1)$, C_M denote dependence on M). Define $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2 \subset [1, h]$, $|\mathcal{H}_i| = k_i$, $|\mathcal{H}_1 \cap \mathcal{H}_2| = r$, $M = k_1 + k_2 + l_1 + l_2$.

PROPOSITION 1. *If $R \ll N^{1/2} (\log N)^{-4M}$ and $h \leq R^C$ for any $C > 0$, then, as $R, N \to \infty$,*

$$\sum_{N < n \leq 2N} \Lambda_R(n, \mathcal{H}_1, k_1 + l_1) \Lambda_R(n, \mathcal{H}_2, k_2 + l_2)$$
$$= \binom{l_1 + l_2}{l_1} \frac{(\log R)^{r + l_1 + l_2}}{(r + l_1 + l_2)!} (\mathfrak{S}(\mathcal{H}) + o_M(1)) N.$$

With this proposition, the authors are able to understand how the weights behave in an unmodified fashion over an interval. Their second task is to see how these weights are modified by incorporating the θ function ($\theta(n) = \log n$ if n is a prime and 0 otherwise).

PROPOSITION 2. *Take $h_0 \notin \mathcal{H}$. If $R \ll N^{1/4} (\log N)^{-B_M}$ for a sufficiently large constant B_M, and $h \leq R$, then, as $R, N \to \infty$,*

$$\sum_{N < n \leq 2N} \theta(n + h_0) \Lambda_R(n, \mathcal{H}_1, k_1 + l_1) \Lambda_R(n, \mathcal{H}_2, k_2 + l_2)$$
$$= \binom{l_1 + l_2}{l_1} \frac{(\log R)^{r + l_1 + l_2}}{(r + l_1 + l_2)!} (\mathfrak{S}(\mathcal{H} \cup h_0) + o_M(1)) N.$$

With the observation that if $n + h_0$ is a prime and $h_0 \in \mathcal{H}$, $\Lambda_R(n, \mathcal{H}, k + l) = \Lambda_R(n, \mathcal{H} \setminus h_0, (k - 1) + (l + 1))$ it is then possible to work around the restriction that $h_0 \notin \mathcal{H}$, providing a result for general h_0.

1.2. *Results.* Summing over all values for h_0 in an interval, the authors of [4] are able to get a sum of the logarithms of all primes that lay in the designated interval when multiplied by the Λ weight functions. However, a natural problem is how to sum over more complex distributions of primes. Instead of a sum over all primes in an interval, consider the problem of determining a sum over all primes p such that $p + j_1$ and $p + j_2$ are also primes. This problem is too complex for a strict asymptotics in this fashion; however, by modifying the sieve to deal with four Λ functions instead of two, we give a method by which an upper bound can be reached for the logarithms of such pairs of primes multiplied by the weight function. Say we need an upper bound for

$$(1) \qquad \sum_{n = N+1}^{2N} \theta(n + h_0) \theta(n + h_1) \Big(\sum_{\mathcal{H}} \Lambda_R(n, \mathcal{H}, k + l) \Big)^2.$$

The Λ function behaves very predictably when $n + h_0$ and $n + h_1$ are prime. Consulting the definition, one quickly derives that (assuming

$h, R < N)$

$$\theta(n + h_0)\theta(n + h_1) \leq (\log 3N)^2 \left(\Lambda(n, \{h_0, h_1\}, 2) \frac{2}{(\log R)^2} \right)^2.$$

Additionally, if $n + h_0$ and $n + h_1$ are both prime, h_0 and h_1 do not affect the second Λ function. Therefore, letting $\Omega := 4(\log 3N)^2/(\log R)^4$, we obtain

$$(1) \leq \Omega \sum_{n=N+1}^{2N} \Lambda(n, \{h_0, h_1\}, 2)^2 \Big(\sum_{\mathcal{H}} \Lambda(n, \mathcal{H} \setminus \{h_0, h_1\}, k+l) \Big)^2.$$

From this we can see that one way to attack the above mentioned and similar problems is to understand how four Λ functions act when the first two and second two take disjoint sets as their second arguments. Theorem 1 below addresses this problem. Define $|\mathcal{H}_i| = k_i$, $\mathcal{H}_i \subset [1, h]$, $|\mathcal{H}_1 \cap \mathcal{H}_2| = r_1$, $|\mathcal{H}_3 \cap \mathcal{H}_4| = r_2$ and $M = \sum_{i=1}^4 k_i + l_i$.

THEOREM 1. *Suppose that* $(\mathcal{H}_1 \cup \mathcal{H}_2) \cap (\mathcal{H}_3 \cup \mathcal{H}_4) = \emptyset$ *and* $R \ll N^{1/4}(\log N)^{-C_M}$ *for a sufficiently large* C_M, *and* $h \ll R^C$ *for any* $C > 0$. *Then, letting* $u = l_1 + l_2 + r_1$ *and* $v = l_3 + l_4 + r_2$ *we have, as* $R, N \to \infty$,

$$\sum_{N<n\leq 2N} \Lambda_R(n, \mathcal{H}_1, k_1+l_1) \Lambda_R(n, \mathcal{H}_2, k_2+l_2) \Lambda_R(n, \mathcal{H}_3, k_3+l_3) \Lambda_R(n, \mathcal{H}_4, k_4+l_4)$$

$$= \binom{l_1 + l_2}{l_1} \binom{l_3 + l_4}{l_3} N \frac{(\log R)^{u+v}}{u!v!} \left(\mathfrak{S}(\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3 \cup \mathcal{H}_4) + o_M(1) \right).$$

When applying this sieve result, however, a second problem presents itself if one does not want the $\mathcal{H}_i$ sets all taken uniformly from the same interval. In applying the propositions of [4] a result of Patrick Gallagher [2] is used on the average of singular series when the sets under consideration are taken uniformly from an interval. But, if both pairs of sets are taken from different intervals, their union (which is considered in the singular series) may not be uniformly varying over all k-element subsets of a given interval, but instead over a more complex distribution (e.g., over all sets $\mathcal{H} \subset [1, h]$ with two elements from $[1, h']$ and one element from $[h', h'']$). It is with this in mind that we present Theorem 2, a more general version of Gallagher's result. We show that instead of varying the set uniformly over one interval, if we instead take several subsets which vary uniformly over subintervals, the singular series of their union will still average to 1 assuming the subintervals obey a certain growth condition. Kevin Ford's [1] recent simplification of Gallagher's proof is the foundation for the extension presented. Let $\sum_{i=1}^l k_i = r$ and $\Omega_{\mathcal{H}}(p)$ be the residue classes occupied by the elements of $\mathcal{H}$ modulo p.

THEOREM 2. *Take an interval* $[0, h]$ *and take* l *subintervals* $[B_i(h), C_i(h)]$ $\subset [0, h]$ *where* $C_i(h) - B_i(h) = d_i(h)$. *Assume that for some* $1 > \delta > 0$, *and*

for all $i \leq l$, $h^\delta = o(d_i(h))$. Then, as $h \to \infty$,

$$\sum_{\substack{A_1,\ldots,A_l \\ A_i \subset [B_i(h),C_i(h)], |A_i|=k_i}} \mathfrak{S}\Big(\bigcup_{i=1}^{l} A_i\Big) = \prod_{i=1}^{l} \frac{d_i^{k_i}(h)}{k_i!}\,(1 + o_{r,\delta}(1)).$$

Finally, we apply these two results to a concrete situation. Restating the usual definition (notice F_n is trivially bounded above by Δ_n), we set

$$F_r = \liminf_{n\to\infty}\ \max_{1\leq i\leq r}\ \frac{p_{n+i} - p_{n+i-1}}{\log p_n}.$$

THEOREM 3. *There exists a $c > 0$ such that $F_2 \leq c < (\sqrt{2} - 1)^2$.*

Here the c is explicitly computable and we give such a c. While this falls short of improving the best known result for F_2, which is the current best bound for Δ_2 given in [5] as $e^{-\gamma}(\sqrt{2} - 1)^2$, since the previous bound is the result for Δ_2 of [4] with Maier's matrix method applied to it, it seems likely that a similar application of the matrix method would provide a corresponding improvement. The best c our method gives is approximately .1707, a modest improvement over [4]'s .1716. Our proof relies on breaking the interval considered in [4] into three and modifying the weighted difference to ensure that positivity implies either a prime occurring in the middle interval or three primes occurring in an end interval. Theorems 1 and 2 are provided with precise error terms in their corresponding sections.

The only widely available work which has managed to distance the best known bounds for Δ_n and F_n for any n is that of Huxley [6] (for the case $n = 2$ it was shown approximately that $\Delta_2 \leq 1.4105$ and $F_2 \leq 1.3624$). The lack of other results should not imply a disinterest in the F_n constants; the question whether $F_2 < 1$ is attributed to Erdős in [8]. As improving the F_n and Δ_n constants have proven very difficult, Erdős' problem was not resolved until it was shown that $\Delta_2 < 1$ by Maier [8] applying his matrix method to Huxley's results. Currently the best bound we have for F_2 is the trivial one afforded by Δ_2 in [5]. This paper is a first step in improving the two constants, with the only conjectured additional result needed to give an improvement by an application of Maier's method [7]. It should be stressed that applying Maier's method is not at all trivial ([5] and [8] are entirely devoted to applying the method to [4] and [6] respectively).

2. Proof of Theorem 1

THEOREM 1. *Suppose that $(\mathcal{H}_1 \cup \mathcal{H}_2) \cap (\mathcal{H}_3 \cup \mathcal{H}_4) = \emptyset$, and $h \ll R^C$ for any $C > 0$. Then, for any $\gamma > 0$, letting $u = l_1 + l_2 + r_1$ and $v = l_3 + l_4 + r_2$*

we have, as $R, N \to \infty$,

$$\sum_{N < n \leq 2N} \Lambda_R(n, \mathcal{H}_1, k_1+l_1) \Lambda_R(n, \mathcal{H}_2, k_2+l_2) \Lambda_R(n, \mathcal{H}_3, k_3+l_3) \Lambda_R(n, \mathcal{H}_4, k_4+l_4)$$

$$= \binom{l_1 + l_2}{l_1} \binom{l_3 + l_4}{l_3} N \frac{(\log R)^{u+v}}{u! v!} \mathfrak{S}\Big(\bigcup_{0 < i \leq 4} \mathcal{H}_i \Big)$$

$$+ O_{M,\gamma}(N(\log N)^{u+v-1+\gamma} + R^4 (\log R)^{C_M}).$$

2.1. *Outline.* The proof of Theorem 1 follows the same main outline as the proof of Proposition 1 from [4], with a few alterations to allow four weights instead of two. It will be necessary to use Lemma 3 from [4] in the analysis, which is stated in Section 2.6. We outline the proof below.

(1) We translate the product of Λ functions into a complex integral and translate part of the integrand into an Euler product (Section 2.2).

(2) We estimate the error term from the translation (Section 2.3).

(3) We are now left with a complex integral over four variables to estimate. We introduce a series of zeta functions which estimate the function and prove a lemma on how well these products of zeta functions estimate our integrand (Lemma 2.1). It is at this point that our assumption about the disjointness of the unions of the two pairs of sets is vital (Section 2.4).

(4) By our choice of the zeta weights we are almost able to separate the integral over four variables into two double integrals; however, there is some interplay in the G function (which represents the error with which the product of the zeta functions estimates our product). We show that the G function is small enough when two of the variables are fixed and non-negative to use Lemma 2.2 twice (Section 2.7).

2.2. *Rewriting the product.* First, let

$$\lambda_R(d; a) = \begin{cases} 0 & \text{if } d > R, \\ \dfrac{1}{a!} \mu(d) \Big(\log \dfrac{R}{d} \Big)^a & \text{if } d \leq R. \end{cases}$$

If we let $\Omega_i(p)$ be defined as the set of different residue classes among $-h$ mod p where $h \in \mathcal{H}_i$ and extend it multiplicatively (as in [3] and [4]), then

$$\sum_{N < n \leq 2N} \Lambda_R(n, \mathcal{H}_1, k_1+l_1) \Lambda_R(n, \mathcal{H}_2, k_2+l_2) \Lambda_R(n, \mathcal{H}_3, k_3+l_3) \Lambda_R(n, \mathcal{H}_4, k_4+l_4)$$

$$= \sum_{d_1, d_2, d_3, d_4} \lambda_R(d_1, k_1 + l_1) \lambda_R(d_2, k_2 + l_2) \lambda_R(d_3, k_3 + l_3) \lambda_R(d_4, k_4 + l_4)$$

$$\times \sum_{\substack{N < n \leq 2N \\ n \in \Omega_1(d_1), n \in \Omega_2(d_2) \\ n \in \Omega_3(d_3), n \in \Omega_4(d_4)}} 1.$$

From this one derives that

$$\sum_{N<n\leq 2N} \Lambda_R(n,\mathcal{H}_1,k_1+l_1)\Lambda_R(n,\mathcal{H}_2,k_2+l_2)\Lambda_R(n,\mathcal{H}_3,k_3+l_3)\Lambda_R(n,\mathcal{H}_4,k_4+l_4)$$
$$= NT + T'$$

where

$$T = \sum_{d_1,d_2,d_3,d_4} \frac{|\Phi(d_1,d_2,d_3,d_4)|}{[d_1,d_2,d_3,d_4]}$$
$$\times \lambda_R(d_1,k_1+l_1)\lambda_R(d_2,k_2+l_2)\lambda_R(d_3,k_3+l_3)\lambda_R(d_4,k_4+l_4),$$

$$T' = O\Big(\sum_{d_1,d_2,d_3,d_4} |\Phi(d_1,d_2,d_3,d_4)|$$
$$\times |\lambda_R(d_1,k_1+l_1)\lambda_R(d_2,k_2+l_2)\lambda_R(d_3,k_3+l_3)\lambda_R(d_4,k_4+l_4)|\Big).$$

Here $\Phi(\cdot,\cdot,\cdot,\cdot)$ is defined by $\Phi(p^{\beta_1},p^{\beta_2},p^{\beta_3},p^{\beta_4}) = |\bigcap_{i:\beta_i=1}\Omega_i(p)|$ for p prime and $\beta_i \in \{0,1\}$, and extended multiplicatively: $\Phi(n_1,n_2,n_3,n_4) = \prod_p \Phi(p^{\beta_1},p^{\beta_2},p^{\beta_3},p^{\beta_4})$ where $\beta_i = 1$ if $p \mid n_i$ and 0 otherwise.

2.3. *The first error term.* In this section we will show that $T' = O_M(R^4(\log R)^{C_M})$, giving the second error term in Theorem 1. Notice that the $\lambda_R(d,k+l)$ factors can be bounded by a constant power of $\log R$ depending only on k and l. It remains to bound

$$\sum_{d_1,d_2,d_3,d_4<R} \Phi(d_1,d_2,d_3,d_4) \leq \prod_{0<i\leq 4}\Big[\sum_{d<R}|\Omega_i(d)|\Big]^4 \leq R^4(\log R)^{C_M}$$

because $|\Omega_i(d)|$ is bounded by the k_ith generalized divisor function. It is worth noting the R^4 in this error term. In all applications which do not assume Elliot–Halberstam type results on the distribution of primes, Goldston, Pintz and Yıldırım take $R = N^{1/4-\varepsilon}$ due to bounds given by the Bombieri–Vinogradov theorem. In our application, we have this restriction in an unrelated point in the analysis.

2.4. *Introducing the zeta weights.* The next step is to write the formula as an Euler product. Using the complex analytic equality (the integral is taken over s with $\Re(s)=1$)

$$\lambda_R(d;a) = \frac{\mu(d)}{2\pi i}\int_{(1)}\left(\frac{R}{d}\right)^s\frac{ds}{s^{a+1}},$$

we obtain

$$T = \frac{1}{(2\pi i)^4}\int_{(1)}\int_{(1)}\int_{(1)}\int_{(1)}\frac{F(s_1,s_2,s_3,s_4)R^{s_1+s_2+s_3+s_4}}{s_1^{k_1+l_1+1}s_2^{k_2+l_2+1}s_3^{k_3+l_3+1}s_4^{k_4+l_4+1}}\,ds_1\,ds_2\,ds_3\,ds_4,$$

where

$$F(s_1, s_2, s_3, s_4) = \sum_{d_1,d_2,d_3,d_4} \mu(d_1)\mu(d_2)\mu(d_3)\mu(d_4) \frac{|\Phi(d_1,d_2,d_3,d_4)|}{[d_1,d_2,d_3,d_4]d_1^{s_1}d_2^{s_2}d_3^{s_3}d_4^{s_4}}$$

$$= \prod_p \left(1 - \frac{|\Omega_1(p)|}{p^{s_1+1}} - \frac{|\Omega_2(p)|}{p^{s_2+1}} + \frac{|\Omega_1(p) \cap \Omega_2(p)|}{p^{s_1+s_2+1}} - \frac{|\Omega_3(p)|}{p^{s_3+1}} - \frac{|\Omega_4(p)|}{p^{s_4+1}}\right.$$

$$+ \frac{|\Omega_3(p) \cap \Omega_4(p)|}{p^{s_3+s_4+1}} + \frac{|\Omega_1(p) \cap \Omega_3(p)|}{p^{s_1+s_3+1}} + \frac{|\Omega_2(p) \cap \Omega_3(p)|}{p^{s_2+s_3+1}}$$

$$+ \frac{|\Omega_1(p) \cap \Omega_4(p)|}{p^{s_1+s_4+1}} + \frac{|\Omega_2(p) \cap \Omega_4(p)|}{p^{s_2+s_4+1}} - \frac{|\Omega_1(p) \cap \Omega_2(p) \cap \Omega_3(p)|}{p^{s_1+s_2+s_3+1}}$$

$$- \frac{|\Omega_1(p) \cap \Omega_2(p) \cap \Omega_4(p)|}{p^{s_1+s_2+s_4+1}} - \frac{|\Omega_1(p) \cap \Omega_3(p) \cap \Omega_4(p)|}{p^{s_1+s_3+s_4+1}}$$

$$\left. - \frac{|\Omega_2(p) \cap \Omega_3(p) \cap \Omega_4(p)|}{p^{s_2+s_3+s_4+1}} + \frac{|\Omega_1(p) \cap \Omega_2(p) \cap \Omega_3(p) \cap \Omega_4(p)|}{p^{s_1+s_2+s_3+s_4+1}}\right).$$

In order to express this in a simpler fashion, we define the following function. Take $T \subset \{1,2,3,4\}$ and let $s(T) = \sum_{t \in T} s_t + 1$, $\Omega_T(p) = \bigcap_{t \in T} \Omega_t(p)$ and $E_T(p) = |\Omega_T(p)|/p^{s(T)}$. Rewrite the F function as

$$F(s_1, s_2, s_3, s_4) = \prod_p \left[1 + \sum_{\substack{T \subset \{1,2,3,4\} \\ T \neq \emptyset}} (-1)^{|T|} E_T(p)\right].$$

We now define our version of the G function and prove a lemma on its growth that we will need later. We set

$$G(s_1, s_2, s_3, s_4)$$
$$= F(s_1, s_2, s_3, s_4) \frac{\zeta(s_1+1)^{k_1}\zeta(s_2+1)^{k_2}}{\zeta(s_1+s_2+1)^{r_1}} \frac{\zeta(s_3+1)^{k_3}\zeta(s_4+1)^{k_4}}{\zeta(s_3+s_4+1)^{r_2}}.$$

First, let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3 \cup \mathcal{H}_4$ and $\Delta := \prod_{h_i,h_j \in \mathcal{H}, h_i \neq h_j} |h_j - h_i|$.

Choose $U := Ck^2 \log h$ so that $\log \Delta \leq U$.

LEMMA 2.1. *Let* $\beta_i = \max(-\Re(s_i), 0)$ *and assume* $\beta_i < 1/4 - \xi$ *for all* $0 < i \leq 4$ *for some* $\xi > 0$. *Then there exists a constant* C *such that*

$$G(s_1, s_2, s_3, s_4) \ll_{M,\xi} \exp(CMU^{\sum \beta_i} \log \log U).$$

Proof. We will divide the total product into three separate products and use the Euler product expansion of the zeta function to bound each part of the product.

2.4.1. *The product of primes under* U. We can divide the product into three parts, one which corresponds to terms of the F function, one which corresponds to the zeta functions in the numerator of G, and a third which corresponds to the zeta functions in its denominator.

We bound the product in the F function for a fixed T by standard means:

$$\prod_{p \leq U}(1 + |E_T(p)|) \leq \prod_{p \leq U}\left(1 + \frac{\max(k_i)}{p^{1-\sum \beta_i}}\right) \leq \exp\left(\sum_{p \leq U}\frac{\max(k_i)}{p^{1-\sum \beta_i}}\right)$$

$$\leq \exp\left(\max(k_i)U^{\sum \beta_i}\sum_{p \leq U}\frac{1}{p}\right) \ll \exp(\max(k_i)U^{\sum \beta_i}\log\log U).$$

Since we can bound the total F function by a fixed number of these products, this portion of the product is $\ll \exp(C\max(k_i)U^{\sum \beta_i}\log\log U)$.

Similarly to above, we invoke the results of [4] to bound the ζ functions in the numerator:

$$\prod_{p \leq U}\left|1 - \frac{1}{p^{s_i+1}}\right|^{-k_i} \ll \exp(3k_i U^{\beta_i}\log\log U).$$

Since r_i is bounded above by $\max(k_i)$ we can bound the ζ functions in the denominator as in [4],

$$\prod_{p \leq U}\left|1 - \frac{1}{p^{s_i+s_j+1}}\right|^{r_i} \leq \left(\prod_{p \leq U}\left(1 + \frac{1}{p^{1-\sum \beta_i}}\right)^{-1}\right)^{r_i}$$

$$\ll \exp(\max(k_i)U^{\sum \beta_i}\log\log U).$$

Therefore, the final product for primes less than U is

$$\ll \exp(C\max(k_i)U^{\sum \beta_i}\log\log U).$$

2.4.2. *Primes above U which divide Δ.* For this, notice that

$$(2) \qquad \prod_{\substack{p|\Delta \\ p>U}}\left(1 + \frac{\max(k_i)}{p^{1-\sum \beta_i}}\right) \leq \exp\left(\sum_{\substack{p|\Delta \\ p>U}}\frac{\max(k_i)}{p^{1-\sum \beta_i}}\right).$$

Similarly to the analysis in [4], there are fewer than $(1+o(1))\log\Delta < U$ primes such that $p|\Delta$. We can therefore replace the sum above with the first U numbers greater than U. Hence,

$$(2) \leq \exp\left(\max(k_i)\sum_{U<n\leq 2U}\frac{1}{n^{1-\sum \beta_i}}\right) \leq \exp(C\max(k_i)U^{\sum \beta_i}).$$

And each of the factors can be bounded identically.

2.4.3. *Primes above U which do not divide Δ.* At this point the choice of the zeta functions and our restriction on the intersection of the sets becomes important. By their selection, the size of the intersections of the $\Omega_i(p)$ functions are exactly known in this category. If two of the values occupied the same residue class modulo a given prime, then their difference would be

a factor in computing Δ, and therefore the prime would divide Δ, and so it is not in this product. So for all p such that $p \nmid \Delta$,

$$|\Omega_i| = k_i, \qquad |\Omega_1 \cap \Omega_2| = r_1, \qquad |\Omega_3 \cap \Omega_4| = r_2$$

and all other combinations are empty by our initial assumption that

$$(\mathcal{H}_1 \cup \mathcal{H}_2) \cap (\mathcal{H}_3 \cup \mathcal{H}_4) = \emptyset.$$

In this case, we note the following about the product of terms belonging to the F function. Assuming U is greater than a certain threshold depending only on $\max(k_i)$ and ξ, we have

$$\prod_{\substack{p \nmid \Delta \\ p > U}} \left[1 + \sum_{\substack{T \subset \{1,2,3,4\} \\ T \neq \emptyset}} (-1)^{|T|} E_T(p) \right]$$

$$= \prod_{\substack{p \nmid \Delta \\ p > U}} \left(1 - \frac{k_1}{p^{1+s_1}} - \frac{k_2}{p^{1+s_2}} + \frac{r_1}{p^{1+s_1+s_2}} - \frac{k_3}{p^{1+s_3}} - \frac{k_4}{p^{1+s_4}} + \frac{r_2}{p^{1+s_3+s_4}} \right)$$

$$\leq \prod_{\substack{p \nmid \Delta \\ p > U}} \left| \left(1 - \frac{k_1}{p^{1+s_1}} \right) \left(1 - \frac{k_2}{p^{1+s_2}} \right) \left(1 + \frac{r_1}{p^{1+s_1+s_2}} \right) \left(1 - \frac{k_3}{p^{1+s_3}} \right) \right.$$

$$\left. \times \left(1 - \frac{k_4}{p^{1+s_4}} \right) \left(1 + \frac{r_2}{p^{1+s_3+s_4}} \right) \left(1 + \frac{C \max(k_i)}{p^{2-4\max(\beta_i)}} \right) \right|$$

$$\ll_{M,\xi} \prod_{\substack{p \nmid \Delta \\ p > U}} \left| \left(1 - \frac{k_1}{p^{1+s_1}} \right) \left(1 - \frac{k_2}{p^{1+s_2}} \right) \left(1 + \frac{r_1}{p^{1+s_1+s_2}} \right) \left(1 - \frac{k_3}{p^{1+s_3}} \right) \right.$$

$$\left. \times \left(1 - \frac{k_4}{p^{1+s_4}} \right) \left(1 + \frac{r_2}{p^{1+s_3+s_4}} \right) \right|.$$

The final inequality follows from the fact that $\beta_i < 1/4 - \xi$. Now each of these factors can be paired with its corresponding zeta function. There are two distinct such pairs, and only two lines of analysis are necessary. We use the following results from [4]:

$$\prod_{\substack{p \nmid \Delta \\ p > U}} \left| \left(1 - \frac{k_i}{p^{1+s_i}} \right) \left(1 - \frac{1}{p^{1+s_i}} \right)^{-k_i} \right| \leq \exp(2k_i U^{\beta_i}),$$

$$\prod_{\substack{p \nmid \Delta \\ p > U}} \left| \left(1 + \frac{r_i}{p^{1+s_i+s_j}} \right) \left(1 - \frac{1}{p^{1+s_i+s_j}} \right)^{r_i} \right| = \prod_{\substack{p \nmid \Delta \\ p > U}} \left(1 + O_M \left(\frac{1}{p^{2-2s_i-2s_j}} \right) \right)$$

where the product on the right hand side of the second inequality is convergent depending only on ξ due to our assumption that $\beta_i < 1/4 - \xi$. Therefore, the entire product for primes in this category is $\ll \exp(C \max(k_i) U^{\sum \beta_i})$.

Combining the results of all three sections, we see that the total product is

$$\ll_{M,\xi} \exp(CMU^{\sum \beta_i} \log\log U)$$

for some $C > 0$ which does not depend on any k_i value. This growth condition is necessary to invoke the result of [4]. It should also be noted that as in previous works, G is analytic in the region described in this lemma; this property will be needed later (it follows from the definition of G and the bound we have just exhibited). We now give a few results on the zeta function cited in [4] before stating the necessary lemma. ∎

2.5. *Some facts on the zero free region of* ζ. First, there is a small constant $\bar{c} \leq 10^{-2}$ such that $\zeta(\sigma + it) \neq 0$ in the region $\sigma \geq 1 - 4\bar{c}/\log(|t| + 3)$. Furthermore, in this region,

$$\zeta(\sigma + it) - \frac{1}{\sigma - 1 + it} \ll \log(|t| + 3), \qquad \frac{1}{\zeta(\sigma + it)} \ll \log(|t| - 3).$$

2.6. *A necessary lemma.* In the analysis presented in this paper, only a weaker version of the lemma of [4] is needed. The version needed is stated below. Let

$$T_R^*(a, b, d, u, v, h) := \frac{1}{(2\pi i)^2} \int\limits_{(1)} \int\limits_{(1)} \frac{D(s_1, s_2) R^{s_1 + s_2}}{s_1^{u+1} s_2^{v+1} (s_1 + s_2)^d} \, ds_1 \, ds_2$$

where

$$D(s_1, s_2) := \frac{G(s_1, s_2) W^d(s_1 + s_2)}{W^a(s_1) W^b(s_2)} \quad \text{and} \quad W(s) := s\zeta(1 + s).$$

Assume $G(s_1, s_2)$ is regular on and to the right of the line

$$(3) \qquad\qquad s = -\frac{\bar{c}}{\log(|t| + 3)} + it$$

and satisfies the bound

$$G(s_1, s_2) \ll_M \exp(CMU^{\beta_1 + \beta_2} \log\log U), \quad \text{where} \quad U = CM^2 \log 2h.$$

LEMMA 2.2. *Suppose that*

$$0 \leq a, b, d, u, v \leq M, \qquad a + u \geq 1, \qquad b + v \geq 1, \qquad d \leq \min(a, b)$$

where M is any large constant. Let $h \ll R^C$ for any $C > 0$. Then, as $R \to \infty$,

$$T_R^*(a, b, d, u, v, h) = \binom{u + v}{u} \frac{(\log R)^{u+v+d}}{(u + v + d)!} G(0, 0)$$
$$+ O_M((\log R)^{u+v+d-1}(\log\log R)^{C_M}).$$

2.7. *Splitting the four integrals.* The goal of this section is to split the four integrals in the expression for T into two pairs of two integrals and use Lemma 2.2 twice. We will apply Lemma 2.1 to show that each of the pairs

of integrals is acceptable to use with Lemma 2.2. Let $k_i + l_i + 1 = u_i$ and introduce the following notation (let $s_j = \sigma_j + it_j$):

$$\frac{\zeta(s_1 + s_2 + 1)^{r_1}}{\zeta(s_1 + 1)^{k_1}\zeta(s_2 + 1)^{k_2}} = \zeta_1(s_1, s_2), \qquad \frac{\zeta(s_3 + s_4 + 1)^{r_2}}{\zeta(s_3 + 1)^{k_3}\zeta(s_4 + 1)^{k_4}} = \zeta_2(s_3, s_4).$$

This allows a simplification of the expression for T as (let $\bar{d} = ds_1 ds_2 ds_3 ds_4$)

$$T = \int\limits_{(1)} \int\limits_{(1)} \int\limits_{(1)} \int\limits_{(1)} G(s_1, s_2, s_3, s_4)\zeta_1(s_1, s_2)\zeta_2(s_3, s_4) \frac{R^{s_1+s_2+s_3+s_4}}{s_1^{u_1} s_2^{u_2} s_3^{u_3} s_4^{u_4}} \, \bar{d}.$$

The integrand above is analytic in s_3 to the right of the line $\Re(z) = 0$ as long as $\Re(s_i) > 0$ for all other s_i, and the same holds true for s_4. Checking Lemma 2.1, one sees that the integrand in T vanishes as either $|t_3| \to \infty$ or $|t_4| \to \infty$. One can therefore shift the integral over both variables to the vertical line L which passes through $1/\log N$. Therefore, with a quick substitution,

$$T = \frac{1}{(2\pi i)^2} \int\limits_{(L)} \int\limits_{(L)} Q(s_3, s_4)\zeta_2(s_3, s_4) \frac{R^{s_3+s_4}}{s_3^{u_3} s_4^{u_4}} \, ds_3 \, ds_4$$

where

$$Q(s_3, s_4) := \frac{1}{(2\pi i)^2} \int\limits_{(1)} \int\limits_{(1)} G(s_1, s_2, s_3, s_4)\zeta_1(s_1, s_2) \frac{R^{s_1+s_2}}{s_1^{u_1} s_2^{u_2}} \, ds_1 \, ds_2.$$

2.8. *Applying the lemma.* By Lemma 2.1, when s_3, s_4 are fixed on the line $\Re(s_3) = \Re(s_4) = 1/\log N$ (which implies $\beta_3 = \beta_4 = 0$) and taking ξ as a small enough absolute constant depending only on the absolute constant $\bar{c}$ to ensure the region $\beta_i \leq 1/4 - \xi$ includes the zero free region described previously, we have

$$G(s_1, s_2, s_3, s_4) \ll_M \exp(CMU^{\beta_1+\beta_2} \log\log U)$$

to the right of the line described before Lemma 2.2. Therefore the G function is acceptable to use Lemma 2.2 where s_3 and s_4 are fixed and positive and the integral is being evaluated over the s_1 and s_2 variables. Using the substitution with the W function defined as in Lemma 2.2, we obtain

$$D_{s_3,s_4}(s_1, s_2) = \frac{G(s_1, s_2, s_3, s_4)W^{r_1}(s_1, s_2)}{W^{k_1}(s_1)W^{k_2}(s_2)},$$

$$Q(s_3, s_4) = \frac{1}{(2\pi i)^2} \int\limits_{(1)} \int\limits_{(1)} D_{s_3,s_4}(s_1, s_2) \frac{R^{s_1+s_2}}{s_1^{l_1+1} s_2^{l_2+1}(s_1+s_2)^{r_1}} \, ds_1 \, ds_2.$$

Because the analyticity of G in the s_1 and s_2 variables is maintained when s_3 and s_4 are fixed and positive, the lemma can be applied by letting $a = k_1$, $b = k_2$, $u = l_1$, $v = l_2$ and $d = r_1$. As long as $k_1 + l_1 \geq 1$ and $k_2 + l_2 \geq 1$ (for the rest of the section, assume that each k_i is positive,

implying the previous inequality; the case when at least one of the k_i is zero will be addressed separately at the end of the section). Lemma 2.2 implies, as long as $\Re(s_3), \Re(s_4) \geq 0$,

$$Q(s_3, s_4) = \binom{l_1 + l_2}{l_1} \frac{(\log R)^{l_1 + l_2 + r_1}}{(l_1 + l_2 + r_1)!} G(0, 0, s_3, s_4)$$
$$+ O_M((\log N)^{l_1 + l_2 + r_1 - 1} (\log \log N)^{C_M}).$$

This implies

$$T = \binom{l_1 + l_2}{l_1} \frac{(\log R)^{l_1 + l_2 + r_1}}{(l_1 + l_2 + r_1)!} \int\limits_{(L)} \int\limits_{(L)} G(0, 0, s_3, s_4) \zeta_2(s_3, s_4) \frac{R^{s_3 + s_4}}{s_3^{u_3} s_4^{u_4}} \, ds_3 \, ds_4$$

$$+ O((\log N)^{l_1 + l_2 + r_1 - 1} (\log \log N)) \int\limits_{(L)} \int\limits_{(L)} \left| \zeta_2(s_3, s_4) \frac{R^{s_3 + s_4}}{s_3^{u_3} s_4^{u_4}} \right| \, ds_3 \, ds_4.$$

Since the first integrand vanishes as $t_1 \to \infty$ or $t_2 \to \infty$, one can shift the lines of integration of the first line above back to the line $\Re(z) = 1$. There are therefore two values left to evaluate:

$$T_1 = \int\limits_{(1)} \int\limits_{(1)} G(0, 0, s_3, s_4) \zeta_2(s_3, s_4) \frac{R^{s_3 + s_4}}{s_3^{u_3} s_4^{u_4}} \, ds_3 \, ds_4,$$

$$T_2 = \int\limits_{(L)} \int\limits_{(L)} \left| \zeta_2(s_3, s_4) \frac{R^{s_3 + s_4}}{s_3^{u_3} s_4^{u_4}} \right| \, ds_3 \, ds_4.$$

2.9. *Evaluating T_1, the main term and some error.* Since the roles of (s_1, s_2) and (s_3, s_4) are symmetric in Lemma 2.1 (when $s_1 = s_2 = 0$, $\beta_1 = \beta_2 = 0$), one can switch the roles of s_1, s_2 with s_3, s_4 and reapply Lemma 2.2 to get

$$T_1 = \binom{l_3 + l_4}{l_3} \frac{(\log R)^{l_3 + l_4 + r_2}}{(l_3 + l_4 + r_2)!} G(0, 0, 0, 0)$$
$$+ O_M((\log N)^{l_3 + l_4 + r_2 - 1} (\log \log N)^{C_M}).$$

2.10. *Evaluating T_2, the second error term.* Fix any absolute constant $\gamma \in (0, 1)$. Since $R^{s_3 + s_4}$ is absolutely bounded on the line $1/\log N$,

$$T_2 \ll \int\limits_{(L)} \int\limits_{(L)} \left| \frac{\zeta(s_3 + s_4 + 1)^{r_2}}{\zeta(s_3 + 1)^{k_3} \zeta(s_4 + 1)^{k_4}} \frac{1}{s_3^{k_3 + l_3 + 1} s_4^{k_4 + l_4 + 1}} \right| \, ds_3 \, ds_4.$$

Now observe that with $k_3 \geq 1$, when $\Re(s_3) \geq 0$ and $|s_3| \leq 1$,

$$\left| \frac{1}{\zeta(s_3 + 1)^{k_3} s_3^{k_3}} \right| \ll_M 1 \ll_M \frac{1}{|s_3^\gamma|},$$

and when $|s_3| > 1$, there is the general inequality that follows from the growth conditions enumerated in Section 2.4 (since if $\Re(s_3) \geq 0$, $s_3 + 1$

trivially falls in the region described),

$$\left| \frac{1}{\zeta(s_3+1)^{k_3} s_3^{k_3}} \right| \ll_M \left| \frac{1}{\zeta(s_3+1)^{k_3} s_3} \right| \ll_M \frac{(\log(|t_3|+3))^{k_3}}{|s_3^{\gamma}|},$$

and finally, the inequality

$$|\zeta(s_3+s_4+1)^{r_2}| \ll_M (\log(|t_3+t_4|+3))^{r_2} \max\left(1, \frac{1}{|s_3+s_4|}\right)^{r_2}.$$

Substituting $\omega_3 = x_3 + iy_3 = s_3 \log N$ and letting $\bar{d} = ds_3 \, ds_4, \hat{d} = d\omega_3 \, d\omega_4$ gives

$$T_2 \ll_M \int\limits_{(L)} \int\limits_{(L)} \frac{(\log(|t_3|+3))^{k_3}(\log(|t_4|+3))^{k_4}(\log(|t_3+t_4|+3))^{r_2}}{|s_3^{l_3+1+\gamma} s_4^{l_4+1+\gamma}|}$$
$$\times \max\left(1, \frac{1}{|s_3+s_4|}\right)^{r_2} \bar{d}$$

$$\ll_M (\log N)^{r_2} \int\limits_{(L)} \int\limits_{(L)} \frac{(\log(|t_3|+3))^{k_3}(\log(|t_4|+3))^{k_4}(\log(|t_3+t_4|+3))^{r_2}}{|s_3^{l_3+1+\gamma} s_4^{l_4+1+\gamma}|} \bar{d}$$

$$\ll_M (\log N)^{l_3+l_4+r_2+2\gamma}$$
$$\times \int\limits_{(1)} \int\limits_{(1)} \frac{\left(\log\left(\left|\frac{y_3}{\log N}\right|+3\right)\right)^{k_3} \left(\log\left(\left|\frac{y_4}{\log N}\right|+3\right)\right)^{k_4} \left(\log\left(\left|\frac{y_3+y_4}{\log N}\right|+3\right)\right)^{r_2}}{|\omega_3^{l_3+1+\gamma} \omega_4^{l_4+1+\gamma}|} \hat{d}$$

$$\ll_{M,\gamma} (\log N)^{l_3+l_4+r_2+2\gamma},$$

because the final integral is absolutely convergent since $l_i \geq 0$ and $\gamma > 0$.

2.11. *Combining the results.* Finally, using the fact that $G(0,0,0,0) = \mathfrak{S}(\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3 \cup \mathcal{H}_4)$ and labeling $l_1 + l_2 + r_1 = u$ and $l_3 + l_4 + r_2 = v$, we obtain

$$T = \binom{l_1+l_2}{l_1} \binom{l_3+l_4}{l_3} \frac{(\log R)^{u+v}}{u! \, v!} \mathfrak{S}(\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3 \cup \mathcal{H}_4)$$

$$+ O_{M,\gamma}((\log N)^{u+v-1+3\gamma}).$$

Since we can pick γ as any positive value, this implies the theorem when combined with the additional error term from Section 2.3. Now, let us address the case where some k_i are zero. If two k_i values are zero, the theorem is implied by the result of Proposition 1 from [4] because there are only two remaining weight functions. The remaining case is when only one k_i value is zero. In this case, instead of four integrals, there are only three remaining. The analysis up to Section 2.9 is identical, with the only change being that there are three integrals instead of four. At Section 2.9 instead of invoking Lemma 2.1, one would use the analysis of Proposition 1 (Special Case) of [4], which is the equivalent statement of Proposition 1 with only

one weight function instead of two (it is only explicitly shown for $l = 0$ but as mentioned in Section 6 of [4], the analysis generalizes to all $l \geq 0$). The corresponding analysis of Section 2.10 follows identically with one integral instead of two.

3. Proof of Theorem 2, Gallagher extension. In this section, we will show that Kevin Ford's [1] simplification of P. X. Gallagher's proof can be extended to give an estimate for a more involved sum over singular series which is useful in applying Theorem 1. The importance of the theorem is that it allows the two pairs of weights we consider to vary over different intervals.

THEOREM 2. *Take an interval $[0, h]$ and take l subintervals $[B_i(h), C_i(h)] \subset [0, h]$ where $C_i(h) - B_i(h) = d_i(h)$. Assume that for some $1 > \delta > 0$, and for all $i \leq l$, $h^\delta = o(d_i(h))$. Then*

$$\sum_{\substack{A_1,\ldots,A_l \\ A_i \subset [B_i(h),C_i(h)], |A_i|=k_i}} \mathfrak{S}\Big(\bigcup_{i=1}^l A_i\Big) = \prod_{i=1}^l \frac{d_i^{k_i}(h)}{k_i!}\left(1 + O_{r,\delta}\left(\frac{1}{\log\log h}\right)\right).$$

First, recall the definition of the singular series:

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{|\Omega_\mathcal{H}(p)|}{p}\right)\left(1 - \frac{1}{p}\right)^{-|\mathcal{H}|}.$$

Similarly to the proof of Lemma 2.1 define, with $\mathcal{H} = \bigcup A_i$,

$$\Delta := \prod_{h \neq h' \in \mathcal{H}} |h - h'|, \quad y := (\delta/2)\log h, \quad r := \sum_{0 < i \leq l} k_i.$$

The first statement to show will be that it suffices to consider all A_i such that $i \neq j \Rightarrow A_i \cap A_j = \emptyset$ because the number of such sets vastly exceeds all others. The singular series itself can be bounded above without much trouble; notice that the product for all $p > h$ is $\ll_r 1$ since the sizes of $|\mathcal{H}|$ and $|\Omega_\mathcal{H}(p)|$ will be equal. The product over primes h is $\ll \log^r h$ by Mertens' theorem. Since h^δ grows more slowly than any $d_i(h)$,

$$(4) \qquad \sum_{\substack{A_1,\ldots,A_l \\ A_i \subset [B_i(h),C_i(h)], |A_i|=k_i}} \mathfrak{S}\Big(\bigcup_{i=1}^l A_i\Big)$$

$$= \sum_{\substack{A_1,\ldots,A_l \\ A_i \subset [B_i(h),C_i(h)] \\ |A_i|=k_i, |\bigcup A_i|=r}} \mathfrak{S}\Big(\bigcup_{i=1}^l A_i\Big) + O_r\Big(h^{-\delta}\log^r h \prod_{i=1}^l d_i^{k_i}(h)\Big).$$

The sum on the right hand side is easier to evaluate because the exponent in the definition of the singular series will now be a constant $-r$. Now, fix

any $A_1, \ldots, A_l$ which fall within the subintervals such that $|\bigcup A_i| = r$. It is shown in [1] that

$$\prod_{p > y} \left(1 - \frac{|\Omega_{\mathcal{H}}(p)|}{p}\right)\left(1 - \frac{1}{p}\right)^{-r} = 1 + O_{r,\delta}\left(\frac{1}{\log \log h}\right).$$

The sum on the right hand side of (4) is equal to

$$(5) \quad \left(1 + O_{r,\delta}\left(\frac{1}{\log \log h}\right)\right)\prod_{p \leq y}\left(1 - \frac{1}{p}\right)^{-r}$$

$$\times \sum_{\substack{A_1, \ldots, A_l \\ A_i \subset [B_i(h), C_i(h)] \\ |A_i| = k_i, |\bigcup A_i| = r}} \prod_{p \leq y}\left(1 - \frac{|\Omega_{\mathcal{H}}(p)|}{p}\right).$$

Let $P = \prod_{p \leq y} p$ and note that $P = e^{y + o(y)} = h^{\delta/2 + o(1)}$. The right-most product in (5) is $1/P$ times the number of n, $0 \leq n < P$, such that $(\prod_{\alpha \in \mathcal{H}}(n + \alpha), P) = 1$. We can now also eliminate the $|\bigcup A_i| = r$ condition with an error term $O_r(h^{-\delta}\prod_i d_i^{k_i}(h))$; this leaves the second factor of (5) as

$$\sum_{\substack{A_1, \ldots, A_l \\ A_i \subset [B_i(h), C_i(h)] \\ |A_i| = k_i}} \frac{1}{P}\sum_{n=0}^{P-1}\prod_{\alpha \in \mathcal{H}}\sum_{e|(n+\alpha, P)}\mu(e) + O_r\left(h^{-\delta}\prod_{i=1}^{l} d_i^{k_i}(h)\right)$$

$$= \frac{1}{P}\sum_{n=0}^{P-1}\sum_{\substack{A_1, \ldots, A_l \\ A_i \subset [B_i(h), C_i(h)] \\ |A_i| = k_i}}\prod_{\alpha \in \mathcal{H}}\sum_{e|(n+\alpha, P)}\mu(e) + O_r\left(h^{-\delta}\prod_{i=1}^{l} d_i^{k_i}(h)\right).$$

Let $Q(\alpha, e_i) = 1$ if $e_i \mid n + \alpha$ and 0 otherwise. Then, letting $\mathcal{H}' = \{\alpha_1, \ldots, \alpha_r\}$ where the first k_1 elements are in $[B_1(h), C_1(h)]$, the next k_2 are from $[B_2(h), C_2(h)]$ and so on, with the last k_l being from the interval $[B_l(h), C_l(h)]$, we have

$$\sum_{\substack{A_1, \ldots, A_l \\ A_i \subset [B_i(h), C_i(h)] \\ |A_i| = k_i}}\prod_{\alpha \in \mathcal{H}}\sum_{e|(n+\alpha, P)}\mu(e) = \left[\prod_{i=1}^{l}\frac{1}{k_i!}\right]\sum_{\mathcal{H}'}\prod_{\alpha \in \mathcal{H}'}\sum_{e|(n+\alpha, P)}\mu(e)$$

$$= \left[\prod_{i=1}^{l}\frac{1}{k_i!}\right]\sum_{\mathcal{H}'}\sum_{e_1, \ldots, e_r | P}\mu(e_1)\ldots\mu(e_r)\left[\prod_{j=1}^{r}Q(\alpha_j, e_j)\right],$$

and therefore the second factor of (5) is, apart from the error term, equal to

$$\frac{1}{P}\sum_{n=0}^{P-1}\left[\prod_{i=1}^{l}\frac{1}{k_i!}\right]\sum_{e_1, \ldots, e_r | P}\mu(e_1)\ldots\mu(e_r)\left[\sum_{\mathcal{H}'}\prod_{j=1}^{r}Q(\alpha_j, e_j)\right].$$

For fixed $e_1, \ldots, e_r$, $Q(\alpha_i, e_i)$ will be 1 a total of $d_j(h)/e_i + O(1)$ times over all choices of a_i from $[B_j(h), C_j(h)]$ independent of all other a_j. By the definition, exactly k_j of the a_i were chosen from $[B_j(h), C_j(h)]$. Since P is an upper bound for each e_i and h^δ grows much more slowly than any $d_i(h)$, it follows that

$$\sum_{\mathcal{H}'} \prod_{j=1}^r Q(\alpha_j, e_j) = \frac{d_1^{k_1}(h) \ldots d_l^{k_l}(h)}{e_1 \ldots e_r} (1 + O_r(h^{-\delta}P)).$$

We have now eliminated the dependence on n and therefore the P and $\sum_{n=0}^{P-1}$ cancel out, leaving the right side of (5) as

$$\left[\prod_{i=1}^l \frac{d_i^{k_i}(h)}{k_i!} \right] \sum_{e_1, \ldots, e_r | P} \frac{\mu(e_1) \ldots \mu(e_r)}{e_1 \ldots e_r} (1 + O_r(h^{-\delta}P)) + O_r\left(h^{-\delta} \prod_{i=1}^l d_i^{k_i}(h) \right).$$

After substituting this in the previous bound on the growth of P in terms of h, we obtain

$$\left[\prod_{i=1}^l \frac{d_i^{k_i}(h)}{k_i!} \right] \sum_{e_1, \ldots, e_r | P} \frac{\mu(e_1) \ldots \mu(e_r)}{e_1 \ldots e_r} (1 + O_r(h^{-\delta/2+o(1)}))$$
$$+ O_r\left(h^{-\delta} \prod_{i=1}^l d_i^{k_i}(h) \right)$$
$$= \left[\prod_{i=1}^l \frac{d_i^{k_i}(h)}{k_i!} \right] \prod_{p \leq y} \left(1 - \frac{1}{p} \right)^r (1 + O_r(h^{-\delta/2+o(1)})) + O_r\left(h^{-\delta} \prod_{i=1}^l d_i^{k_i}(h) \right).$$

Multiplying the above with the left side of (5) and incorporating the error term from (4) shows that the original sum equals

$$\prod_{i=1}^l \frac{d_i^{k_i}(h)}{k_i!} \left(1 + O_{r,\delta}\left(\frac{1}{\log \log h} \right) \right) + \prod_{p \leq y} \left(1 - \frac{1}{p} \right)^{-r} O_{r,\delta}\left(h^{-\delta/2+o(1)} \prod_{i=1}^l d_i^{k_i}(h) \right).$$

Using Mertens' theorem again to bound the product over $p \leq y$ as $\ll_r (\log y)^r$ allows us to incorporate the second error term into the first, which implies the theorem.

4. Application to F_2. In this section, we will show an example application of the previous two theorems. We will use them to show an improvement over the result for F_2 obtained in [4].

THEOREM 3. *There exists a $c > 0$ such that $F_2 \leq c < (\sqrt{2} - 1)^2$.*

Following Goldston, Pintz and Yıldırım's lead, some sets of size k will be counted with multiplicity of $k!$ according to their permutations. If a subset is meant to be taken with multiplicity in this fashion, we will use the notation $\subset^*$. Goldston, Pintz and Yıldırım proved their result for Δ_2 by the following method. They showed that assuming $h > (\sqrt{2} - 1)^2 \log N$, the

difference

$$(6) \quad A(\mathbf{v}) := \sum_{n=N+1}^{2N} \Big[\sum_{1 \le h_0 \le h} \theta(n+h_0) - 2 \log 3N \Big] \Big[\sum_{\substack{\mathcal{H} \subset^* \{1,\dots,h\} \\ |\mathcal{H}|=k}} \Lambda_R(n; \mathcal{H}, k+l) \Big]^2,$$

where

$$R = N^{\Theta}, \quad h = \lambda \log 3N \quad \text{and} \quad \mathbf{v} = \lambda, \Theta, k, l, N,$$

is positive, which implies there are three primes in the interval $n + H :=$ $n + \{1, \dots, h\}$ ($H = [1, h]$) for some $n \in [N+1, \dots, 2N]$. For our derivation, we will modify their analysis in the following way. Instead of considering one, we will consider three intervals, where $h' = \delta h$, $\delta < 1/2$:

$$H_1 := \{i \in \mathbb{Z} : 0 < i \le h'\}, \quad H_2 := \{i \in \mathbb{Z} : h' < i < h - h'\},$$
$$H_3 := \{i \in \mathbb{Z} : h - h' \le i \le h\}.$$

If we could modify the difference $A(\mathbf{v})$ in such a way that positivity not only implied that there are three primes in the interval $n + H$ but also that one of these primes came from $n + H_2$ we could then guarantee that $F_2 \le (h - h')/\log 3N = \lambda(1 - \delta)$ whenever the modified difference is positive. If no primes come from the central interval and there are in total fewer than three primes, the difference (6) is already negative. Moreover, if three primes lie in $n + H_1$ or three lie in $n + H_3$, it would imply $F_2 \le h'$, which would imply $F_2 < \lambda(1 - \delta)$ since $\delta < 1/2$. If there are five or more primes, either three primes come from an end interval or there is one in the central interval. With a little consideration, we can see that in order for positivity to imply our bound on F_2 all we need to do is add another negative term to the difference (6) which would ensure that the following "bad cases" also lead to the difference not being positive:

(1) Two primes in $n + H_1$, one in $n + H_3$ and no other in $n + H$.
(2) One prime in $n + H_1$, two in $n + H_3$ and no other in $n + H$.
(3) Two primes both in $n + H_1$ and in $n + H_3$ and no other in $n + H$.

If the added term made these three cases negative as well, we could guarantee that $F_2 \le \lambda(1 - \delta)$. The proof strategy now relies on the fact that the number of triples of primes coming from these very short intervals (H_1 and H_3) should be approximately proportionate to δ^3 times the number of triples of primes coming from the whole interval (if we assign each number n an independent probability of $1/\log n$ of being prime). As we expand the interval by a factor of $1 + \delta$, the positive contribution from the first term in (6) grows linearly with respect to this factor. The negative contribution adds substantially less to the overall sum than the positive term increases when δ is small. This provides the leverage we need to lower the bound for F_2 even though our total interval is bigger than the one used in [4]'s proof for Δ_2.

Consider the following term:

$$B_1(\mathbf{v}, \delta) := \sum_{n=N}^{2N}{}^* \Big(\sum_{|\mathcal{H}|=k} \Lambda_R(n, \mathcal{H}, k+l) \Big)^2 \log 3N,$$

where the starred summation indicates that the term n is only counted if the interval $n+H_1$ has exactly two primes and $n+H_3$ has at least one prime. Let B_2 be the same summation where $n+H_3$ has exactly two primes and $n+H_1$ has at least one. One can see that subtracting $B_1(\mathbf{v}, \delta) + B_2(\mathbf{v}, \delta)$ would satisfy the requirements that all three cases above would not contribute positivity to the overall sum. So, if we can show for a given choice of λ, δ, that there exist $\Theta < 1/4$ and $k, l \in \mathbb{N}$ such that $A(\mathbf{v}) - B_1(\mathbf{v}, \delta) - B_2(\mathbf{v}, \delta) > 0$ it will imply $F_2 \leq \lambda(1 - \delta)$.

Now, we use a second set of weights to find an upper bound for $B_1(\mathbf{v}, \delta)$ (the analysis can be repeated identically for $B_2(\mathbf{v}, \delta)$):

$$B_1(\mathbf{v}, \delta)$$

$$\leq \Omega_1 \sum_{n=N}^{2N} \sum_{\substack{A_1 \subset H_1, A_2 \subset H_3 \\ |A_1|=2, |A_2|=1}} \Lambda_R(n, A_1 \cup A_2, 3)^2 \Big[\sum_{\substack{\mathcal{H} \subset^* H \\ |\mathcal{H}|=k}} \Lambda_R(n, \mathcal{H} \setminus (A_1 \cup A_2), k+l) \Big]^2$$

where $\Omega_1 := (36 \log(3N)/(\log R)^6)$. This bound holds because if for all $a \in A_1 \cup A_2$, $n+a$ is prime, then

$$\Lambda_R(n, A_1 \cup A_2, 3)^2 = \frac{(\log R)^6}{36},$$

$$\Lambda_R(n, \mathcal{H}, k+l) = \sum_{|\mathcal{H}|=k} \Lambda_R(n, \mathcal{H} \setminus (A_1 \cup A_2), k+l).$$

And in all other cases, the square of the terms ensures positivity. Let $\sum_{A_1, A_2}$ be the sum over all sets $A_1 \subset H_1$, $A_2 \subset H_3$ such that $|A_1| = 2$, $|A_2| = 1$, and

$$S_1(\mathbf{v}, \delta) := \sum_{n=N}^{2N} \sum_{A_1, A_2} \Lambda_R(n, A_1 \cup A_2, 3)^2 \Big[\sum_{\substack{\mathcal{H} \subset H \\ |\mathcal{H}|=k}} \Lambda_R(n, \mathcal{H} \setminus (A_1 \cup A_2), k+l) \Big]^2.$$

This yields the bound (notice that we do not consider the sets with multiplicity $k!$ in S_1)

$$B_1(\mathbf{v}, \delta) \leq S_1(\mathbf{v}, \delta)(k!)^2 \frac{36}{(\log R)^6} \log 3N.$$

4.1. *Simplifying the equation.* The goal of this section is to provide an estimate for $S_1(\mathbf{v}, \delta)$. Fix the sets A_1, A_2. Then

$$\sum_{\substack{\mathcal{H} \subset H \\ |\mathcal{H}|=k}} \Lambda_R(n, \mathcal{H} \setminus (A_1 \cup A_2), k+l) = \sum_{j=0}^{3} \binom{3}{j} \sum_{\substack{\mathcal{H} \subset H, |\mathcal{H}|=k-j \\ (A_1 \cup A_2) \cap \mathcal{H} = \emptyset}} \Lambda_R(n, \mathcal{H}, k+l).$$

Here the $\binom{3}{j}$ terms result from the choice of which elements of $A_1 \cup A_2$ were removed from the set under consideration. Letting $\sum_{\mathcal{H}_1, \mathcal{H}_2}^{j_1, j_2}$ be the sum over all $\mathcal{H}_1, \mathcal{H}_2 \subset \mathcal{H}$ such that $|\mathcal{H}_1| = k - j_1$ and $|\mathcal{H}_2| = k - j_2$ and $(\mathcal{H}_1 \cup \mathcal{H}_2) \cap (A_1 \cup A_2) = \emptyset$ (where the dependence on A_1 and A_2 of the summation is understood but not indicated), we obtain

$$S_1(\mathbf{v}, \delta) = \sum_{n=N}^{2N} \sum_{A_1, A_2} \Lambda_R(n, A_1 \cup A_2, 3)^2 \left[\sum_{j=0}^{3} \binom{3}{j} \sum_{\substack{\mathcal{H} \subset H, |\mathcal{H}| = k-j \\ (A_1 \cup A_2) \cap \mathcal{H} = \emptyset}} \Lambda_R(n, \mathcal{H}, k+l) \right]^2$$

$$= \sum_{n=N}^{2N} \sum_{A_1, A_2} \Lambda_R(n, A_1 \cup A_2, 3)^2$$

$$\times \sum_{j_1, j_2 = 0}^{3} \binom{3}{j_1} \binom{3}{j_2} \sum_{\mathcal{H}_1, \mathcal{H}_2}^{j_1, j_2} \Lambda_R(n, \mathcal{H}_1, k+l) \Lambda_R(n, \mathcal{H}_2, k+l)$$

$$= \sum_{j_1, j_2 = 0}^{3} \sum_{r=0}^{\min(k-j_1, k-j_2)} \binom{3}{j_1} \binom{3}{j_2} S_1'(\mathbf{v}, \delta, j_1, j_2, r),$$

where

$$S_1'(\mathbf{v}, \delta, j_1, j_2, r)$$
$$:= \sum_{n=N}^{2N} \sum_{A_1, A_2} \sum_{\mathcal{H}_1, \mathcal{H}_2}^{j_1, j_2, r} \Lambda_R(n, A_1 \cup A_2, 3)^2 \Lambda_R(n, \mathcal{H}_1, k+l) \Lambda_R(n, \mathcal{H}_2, k+l),$$

and $\sum_{\mathcal{H}_1, \mathcal{H}_2}^{j_1, j_2, r}$ is the sum over all $\mathcal{H}_1, \mathcal{H}_2 \subset \mathcal{H}$ such that $|\mathcal{H}_1| = k - j_1$ and $|\mathcal{H}_2| = k - j_2$, $(\mathcal{H}_1 \cup \mathcal{H}_2) \cap (A_1 \cup A_2) = \emptyset$ and $|\mathcal{H}_1 \cap \mathcal{H}_2| = r$ (once again there is a dependence on A_1 and A_2). By Theorem 1 (recall that if $|\mathcal{H}_1| = k - j_1$ and the third argument of the Λ function is $k + l$, the l value increases to $l + j_1$),

$$S_1'(\mathbf{v}, \delta, j_1, j_2, r) = N \binom{2l + j_1 + j_2}{l + j_1} \frac{(\log R)^3 (\log R)^{2l+r+j_1+j_2}}{3! (2l + r + j_1 + j_2)!}$$

$$\times \sum_{A_1, A_2} \sum_{\mathcal{H}_1, \mathcal{H}_2}^{j_1, j_2, r} (\mathfrak{S}(A_1 \cup A_2 \cup \mathcal{H}_1 \cup \mathcal{H}_2) + o_M(1)).$$

Since $|H_1| = |H_3| = \delta |H|$ where δ is an absolute constant, the growth condition for Theorem 2 is satisfied where the subintervals are taken as $[B_1, C_1] = H_1$, $[B_2, C_2] = H_2$, $[B_3, C_3], [B_4, C_4], [B_5, C_5] = H$ and two elements are from $[B_1, C_1]$, one from $[B_2, C_2]$, $k - j_1 - r$ from $[B_3, C_3]$, $k - j_2 - r$ from $[B_4, C_4]$, and r from $[B_5, C_5]$. The elements from the first interval correspond to the elements of A_1, from the second to A_2, from the third to

$\mathcal{H}_1 \setminus \mathcal{H}_2$, from the fourth to $\mathcal{H}_2 \setminus \mathcal{H}_1$, and from the fifth to $\mathcal{H}_1 \cap \mathcal{H}_2$. We can almost use the theorem, except that we have the additional restriction that the five sets are pairwise disjoint. (By definition the A_i sets are disjoint from each other and any $\mathcal{H}_i$. Obviously the sets $\mathcal{H}_1 \setminus \mathcal{H}_2$, $\mathcal{H}_1 \cap \mathcal{H}_2$ and $\mathcal{H}_2 \setminus \mathcal{H}_1$ are pairwise disjoint.) Since we showed in the proof of Theorem 2 that each of the singular series is bounded above by a constant power of $\log h$, the sum of the singular series where the sets are not disjoint is $\ll_M (\log h)^{C_M} h^{2k-r-j_1-j_2+2}$. From the theorem and the previous fact that the sum over disjoint sets we consider dominates the entire sum, it follows that

$$\sum_{A_1, A_2} \sum_{\mathcal{H}_1, \mathcal{H}_2}^{r} \mathfrak{S}(A_1 \cup A_2 \cup \mathcal{H}_1 \cup \mathcal{H}_2) = \frac{\delta^3 h^{2k-r-j_1-j_2+3}}{2!(k-j_1-r)!(k-j_2-r)!r!}$$
$$\times (1 + o_M(1)).$$

Therefore,

$$S_1'(\mathbf{v}, \delta, j_1, j_2, r) = N \binom{2l+j_1+j_2}{l+j_1} \frac{(\log R)^3}{6} \frac{(\log R)^{2l+r+j_1+j_2}}{(2l+r+j_1+j_2)!}$$
$$\times \frac{\delta^3 h^{2k-r-j_1-j_2+3}}{2(k-j_1-r)!(k-j_2-r)!r!} (1 + o_M(1)).$$

It is now possible to step back to evaluate $S_1(\mathbf{v}, \delta)$. Define the following notation (with an empty product being defined as 1):

$$\gamma(j_i, k, r) = (k-r)(k-r-1)\ldots(k-r-j_i+1),$$
$$\beta(j_1, j_2, l, r) = (r+2l+1)(r+2l+2)\ldots(r+2l+j_1+j_2),$$
$$a(j_1, j_2, l) = \binom{2l+j_1+j_2}{l+j_1}\binom{2l}{l}^{-1},$$
$$\mu(j_1, j_2, k, l, r) = \frac{\gamma(j_1, k, r)\gamma(j_2, k, r)a(j_1, j_2, l)}{\beta(j_1, j_2, r, l)}\binom{3}{j_1}\binom{3}{j_2}.$$

We define this in order to put our previous derivation into a form more comparable with the work of [4]. Notice that

$$\frac{k!k!}{(k-j_1-r)!(k-j_2-r)!r!(2l+r+j_1+j_2)!}$$
$$= \binom{k}{r}^2 \frac{r!}{(2l+r)!} \frac{\gamma(j_1, k, r)\gamma(j_2, k, r)}{\beta(j_1, j_2, r, l)}$$
$$= \binom{k}{r}^2 \frac{1}{(r+1)(r+2)\ldots(r+2l)} \frac{\gamma(j_1, k, r)\gamma(j_2, k, r)}{\beta(j_1, j_2, r, l)}.$$

With this, it is possible to rewrite the S_1 function. Defining $x = (\log R)/h$, and $\chi := N(\log R)^{2l+6}h^{2k}/(k!)^2$, we obtain

$$
S_1(\mathbf{v},\delta) \sim \chi\binom{2l}{l} \sum_{j_1,j_2=0}^{3} \sum_{r=0}^{\min(k-j_1,k-j_2)} \binom{k}{r}^2 \frac{\delta^3 x^{r+j_1+j_2-3}\mu(j_1,j_2,k,l,r)}{12(r+1)\dots(r+2l)}
$$

$$
\leq \chi\binom{2l}{l} \sum_{j_1,j_2=0}^{3} \sum_{r=0}^{k} \binom{k}{r}^2 \frac{\delta^3 x^{r+j_1+j_2-3}\mu(j_1,j_2,k,l,r)}{12(r+1)\dots(r+2l)}
$$

$$
= \chi\binom{2l}{l} \sum_{r=0}^{k} \binom{k}{r}^2 \frac{x^r}{(r+1)\dots(r+2l)} \sum_{j_1,j_2=0}^{3} \frac{\delta^3 x^{j_1+j_2-3}\mu(j_1,j_2,k,l,r)}{12}.
$$

And therefore, an upper bound for our original sum is

$$
B_1(\mathbf{v},\delta) = \sum_{n=N+1}^{2N}{}^{*} \Big(\sum_{|\mathcal{H}|=k} \Lambda_R(n,\mathcal{H},k+l) \Big)^2 \log 3N
$$

$$
\leq N(\log R)^{2l}h^{2k}(\log 3N)\binom{2l}{l}\sum_{r=0}^{k}\binom{k}{r}^2\frac{x^r}{(r+1)\dots(r+2l)}
$$

$$
\times \sum_{j_1,j_2=0}^{3} 3\delta^3 x^{j_1+j_2-3}\mu(j_1,j_2,k,l,r)(1+o_M(1))
$$

$$
\sim Nh^{2k+1}\binom{2l}{l}(\log R)^{2l}\sum_{r=0}^{k}\binom{k}{r}^2\frac{x^r}{(r+1)\dots(r+2l)}
$$

$$
\times \Big[\frac{3\delta^3}{\Theta x^2}\Big]\sum_{j_1,j_2=0}^{3} x^{j_1+j_2}\mu(j_1,j_2,k,l,r),
$$

with identical reasoning giving the same bound for $B_2(\mathbf{v},\delta)$. In [4] the authors derive two facts:

$$
\sum_{n=N+1}^{2N} 2(\log N)\Big(\sum_{|\mathcal{H}|=k}\Lambda_R(n,\mathcal{H},k+l)\Big)^2
$$

$$
\sim 2Nh^{2k}(\log 3N)\binom{2l}{l}(\log R)^{2l}\sum_{r=0}^{k}\binom{k}{r}^2\frac{x^r}{(r+1)\dots(r+2l)},
$$

$$
\sum_{n=N+1}^{2N}\Big(\sum_{1\leq h_0\leq h}\theta(n+h_0)\sum_{|\mathcal{H}|=k}\Lambda_R(n,\mathcal{H},k+l)\Big)^2
$$

$$
\sim Nh^{2k+1}\binom{2l}{l}(\log R)^{2l}\sum_{r=0}^{k}\binom{k}{r}^2\frac{x^r}{(r+1)\dots(r+2l)}\Big(\frac{2a(1,0,l)k}{r+2l+1}x+1\Big).
$$

It follows that if after factoring out

$$\binom{2l}{l} Nh^{2k+1}(\log R)^{2l}$$

(which controls all dependence on N) the remaining factor is positive, then $F_2 \le \lambda(1 - \delta)$. Therefore, it suffices to show (where the 3 in the final term becomes a 6 because we are considering both B_1 and B_2) that

$$\sum_{r=0}^{k} \binom{k}{r}^2 \frac{x^r}{(r+1)\dots(r+2l)}$$

$$\times \left(\frac{2a(1,0,l)k}{r+2l+1} x + 1 - \frac{2x}{\Theta} - \frac{6\delta^3}{\Theta x^2} \sum_{j_1,j_2=0}^{3} x^{j_1+j_2} \mu(j_1, j_2, k, l, r) \right)$$

is positive. This will imply $F_2 \le \lambda(1 - \delta)$.

4.2. *The derivation.* First notice that by bounding each component individually, we obtain

$$\mu(j_1, j_2, k, l, r) \le 2^{j_1+j_2} \left(\frac{k-r}{r} \right)^{j_1+j_2} \binom{3}{j_1}\binom{3}{j_2},$$

which implies

$$\sum_{j_1,j_2=0}^{3} x^{j_1+j_2} \mu(j_1, j_2, k, l, r) \le \left(\sum_{j=0}^{3} \binom{3}{j} 2^j x^j \left(\frac{k-r}{r} \right)^j \right)^2 = \left(2x \frac{k-r}{r} + 1 \right)^6.$$

There are now two parts of the proof remaining. Let (similarly to [4])

$$f(r) = \binom{k}{r}^2 \frac{x^r}{(r+1)(r+2)\dots(r+2l)},$$

$$P(r, \delta) = \frac{2a(1,0,l)k}{r+2l+1} x + 1 - \frac{2x}{\Theta} - \frac{6\delta^3}{\Theta x^2} \left(2x \frac{k-r}{r} + 1 \right)^6.$$

We will show that with good choices for $k, l, \lambda, \Theta, \delta$, $\sum_{r=0}^{k} f(r)P(r, \delta)$ will be positive, giving the bound $F_2 \le \lambda(1 - \delta)$. In a sense, $f(r)$ will contribute to the magnitude of the rth term while $P(r, \delta)$ will control the sign. First, notice that the term $f(r)$ is maximized when (one can justify this heuristic by a little computation, but it is not necessary, the decay when the terms are much bigger or smaller will be shown shortly)

$$r \sim \frac{k}{z+1} \quad \text{where} \quad z = \frac{1}{\sqrt{x}}.$$

Let $r_0 = [k/(z+1)]$.

We would now like to find, with a given choice of x, the maximal δ such that the maximal term is positive. (It is possible that a given x will have no valid corresponding δ if the interval we are considering is simply too small. In these cases there will be a contradiction with the definition of δ, giving a value that is not in $[0, 1/2)$ as δ was defined to be.) If we took $\Theta = 1/4$ while in reality it can only be taken to be $1/4 - \varepsilon$, and $k/l = k$ while in reality it can only be taken as $k(1 - \varepsilon)$ (in applications in [4], $l = o(k)$) we would obtain, in a sense, the function $P(r, \delta)$ "approaches". Call this function $P'(r, \delta)$ and notice

$$P'\left(\frac{k}{z+1}, \delta\right) = \frac{4k}{\frac{k}{z+1}} x + 1 - 8x - \frac{24\delta^3}{x^2}\left(2x \frac{k - \frac{k}{z+1}}{\frac{k}{z+1}} + 1\right)^6 > 0$$

(since $x \neq 0$ in this application, we can divide by it; also, recall $x = z^{-2}$)

$$\Leftrightarrow 4(z+1) + z^2 - 8 - 24\delta^3(z+2)^6 > 0$$
$$\Leftrightarrow (z+2)^2 - 8 - 24\delta^3(z+2)^6 > 0.$$

This is equivalent to

$$\delta^3 \leq \frac{(z+2)^2 - 8}{24(z+2)^6}.$$

So, if δ is exactly the cube root of the value on the right above, then $P'(k/(z+1)) = 0$. If it is below that root, one can check easily that $P'(k/(z+1))$ is positive.

THEOREM 4.1. *Let* $\lambda > 0$, $z = 2\sqrt{\lambda}$ *and*

$$\delta' = \sqrt[3]{\frac{(z+2)^2 - 8}{24(z+2)^6}}.$$

If $\delta' \in [0, 1/2)$, *then* $F_2 \leq \lambda(1 - \delta')$.

Proof. Fix λ and δ' for the rest of the proof that satisfy the above conditions (any constants from here on may depend on λ and δ). We claim that for any $\delta < \delta'$ the total sum is positive assuming the k, l values are sufficiently large with $l = o(k)$ and $\Theta = (1/4)(1 - 1/l)$. This will imply that $F_2 \leq \lambda(1 - \delta)$. Since δ can be taken arbitrarily close to δ' this will imply $F_2 \leq \lambda(1 - \delta')$.

LEMMA 4.2. *Let* $\delta = \delta'\sqrt[3]{1 - \varepsilon}$. *Then there is a small constant* $\nu > 0$, *which depends on* ε, *such that for* $k, l > C(\varepsilon)$ *and* $l/k < c(\varepsilon)$, *we have* $P(r) > c'(\varepsilon) > 0$ *for* $r \in [r_0 - \nu k, r_0 + \nu k]$.

Proof. Take $r \in [r_0 - \nu k, r_0 + \nu k]$. We will analyze each term of $P(r, \delta)$ separately (recalling that we have set x as a constant $+ O(l^{-1})$ ($x = \Theta/\lambda$)

and k/r_0 as a constant $+ O(k^{-1})$). As $k, l \to \infty$ with $l = o(k)$ and $\nu \to 0$,

$$\frac{2a(1,0,l)k}{r+2l+1} x \geq \frac{4}{r_0/k} x + O(l/k) + O(l^{-1}) + O(\nu),$$

$$1 - \frac{2x}{\Theta} = 1 - 8x + O(l^{-1}),$$

$$\frac{6\delta^3}{x^2\Theta}\left(2x\,\frac{k-r}{r} + 1\right)^6 \leq \frac{24\delta'^3}{x^2}\left(2x\left(\frac{k}{r_0}-1\right)+1\right)^6 + O(\nu) - c(\varepsilon) + O(l^{-1}),$$

with $c(\varepsilon) > 0$. Combining these, one sees a correlation with $P'(k/z+1, \delta')$,

$$P(r, \delta) \geq P'(k/z+1, \delta') + c(\varepsilon) + O(l/k) + O(l^{-1}) + O(k^{-1}) + O(\nu)$$
$$\geq c(\varepsilon) + O(l/k) + O(l^{-1}) + O(k^{-1}) + O(\nu),$$

which proves the statement. We first fix ε and then select l, k large enough and ν small enough such that $P(r, \delta) \geq c(\varepsilon)/2$ for all $r \in [r_0 - \nu k, r_0 + \nu k]$. ∎

The rest of the proof will proceed in the following manner: First we will show that for $r < r_0 - \nu k/2 = r_1$ or $r > r_0 + \nu k/2 = r_2$, the values $f(r)$ rapidly decrease by at least a constant factor in magnitude. This will imply the negative terms (which are smaller than $r_0 - \nu k$ and greater than $r_0 + \nu k$) will all be exponentially small in k and their total sum can be bounded by the r_0 term. For notational simplicity, let $\nu' = \nu/2$.

First, we begin analyzing the terms below r_1. Take any $r < r_1$. We have

$$\frac{f(r+1)}{f(r)} = \left(\frac{k-r}{r+1}\right)^2 \frac{x(r+1)}{r+2l+1} > \left(\frac{k-r}{r+1}\right)^2 \frac{xr}{r+2l+1}.$$

Notice first that, since $r < ck$ where $c < 1$,

$$\frac{(k-r)(r)}{(r+2l+1)(r+1)}$$

$$= \left(\frac{r^2}{kr-r^2} + \frac{(2l+2)r}{kr-r^2} + \frac{2l+1}{kr-r^2}\right)^{-1} = \left(\frac{r}{k-r} + O(l/k)\right)^{-1}$$

$$\geq \left(\frac{r_0 - \nu' k}{k - r_0 + \nu' k} + O(l/k)\right)^{-1} = \left(\frac{k - r_0 + \nu' k}{r_0 - \nu' k}\right)(1 + O(l/k))$$

$$\geq \left(\frac{k}{r_0}(1 + c(\nu)) - 1\right)(1 + O(l/k)) \quad \text{where } c(\nu) > 0$$

$$\geq \left(\frac{k}{r_0} - 1\right)(1 + c(\nu)/2) \quad \text{for } k/l \text{ small enough.}$$

Secondly,

$$\frac{k-r}{r+1} \geq \frac{k-r_0}{r_0} \quad \text{for } k, l \text{ sufficiently large.}$$

Therefore,

$$\frac{f(r+1)}{f(r)} > \left(\frac{k-r}{r+1}\right)^2 \frac{x(r+1)}{r+2l+1}$$

$$\geq \left(\frac{k-r_0}{r_0}\right)^2 x(1+c(\nu)/2) \qquad \text{by the definition of } r_0$$

$$= 1 + c'(\nu)/2 + O(k^{-1}) + O(l^{-1}).$$

So, as the terms go below the r_1st, the ratio between any two terms decreases by a constant factor. Therefore, the sum of all terms below $r_0 - \nu k$ is a polynomial in k (there are at most k such terms and $P(r,\delta)$ is bounded easily by a polynomial in k) times an inverse exponential in k times the r_0 term. The r_0 term will therefore be greater in magnitude than (any constant multiple of) the sum of all negative terms r with $r < r_1$ for k sufficiently large. Now we will show the same holds with $r > r_2$ (let $c(\nu)$ and $c'(\nu)$ denote small positive constants depending on ν):

$$\frac{f(r+1)}{f(r)} \leq \left(\frac{k-r}{r+1}\right)^2 x \leq \left(\frac{k}{r} - 1\right)^2 x \leq \left(\frac{k}{r_0 + k\nu'} - 1\right)^2 x$$

$$\leq \left(\frac{k}{r_0}(1-c(\nu)) - 1\right)^2 x \leq (1-c'(\nu))\left(\frac{k}{r_0} - 1\right)^2 x = 1 - c'(\nu) + O(k^{-1} + l^{-1}).$$

Therefore, these terms decay exponentially. Since, as before, $P(r,\delta)$ is bounded above by a polynomial in k and $f(r)$ is bounded above by an inverse exponential in k when compared to the r_0 term, the sum of all such $r > r_2$ can easily be bounded by half the term at r_0. This completes the proof of positivity, which implies the theorem. ∎

Notice that if we take $\lambda = (\sqrt{2} - 1)^2$, this implies $\delta = 0$ and the necessary condition is satisfied. This is precisely the result implied by [4]. One can check that if we increase λ by a very small amount so that the δ' value increases and stays within the allotted interval (since it varies continuously with respect to λ) the value $\lambda(1 - \delta')$ will initially decrease. This proves the theorem. One can numerically check that taking $\lambda = .172$ implies $\delta' \sim .007794$, which implies $F_2 < .172(1 - .007794) < .17066$.

References

[1]   K. Ford, *Simple proof of Gallagher's singular series sum estimate*, 2007, http://www.math.uiuc.edu/~ford/gallagher_sum.pdf.

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

[3]   D. A. Goldston, Y. Motohashi, J. Pintz, and C. Y. Yıldırım, *Small gaps between primes exist*, Proc. Japan Acad. Ser. A Math. Sci. 82 (2006), 61–65.

[4]   D. A. Goldston, J. Pintz, and C. Y. Yıldırım, *Primes in tuples I*, Ann. of Math., to appear.

[5]   —, —, —, *Primes in tuples III*, Funct. Approx. Comment. Math. 35 (2006), 79–89.

[6]   M. N. Huxley, *On the differences of primes in arithmetical progressions*, Acta Arith. 15 (1968/1969), 367–392.

[7]   H. Maier, *Primes in short intervals*, Michigan Math. J. 32 (1985), 221–225.

[8]   —, *Small differences between prime numbers*, ibid. 35 (1988), 323–344.

Department of Mathematics
University of California – Los Angeles
Los Angeles, CA 90024, U.S.A.
E-mail: hseyalioglu@ucla.edu