

Algebra & Number Theory

Volume 19
2025
No. 1

**A modification of the linear sieve,
and the count of twin primes**

Jared Duker Lichtman



A modification of the linear sieve, and the count of twin primes

Jared Duker Lichtman

We introduce a modification of the linear sieve whose weights satisfy strong factorization properties, and consequently equidistribute primes up to size x in arithmetic progressions to moduli up to $x^{10/17}$. This surpasses the level of distribution $x^{4/7}$ with the linear sieve weights from well-known work of Bombieri, Friedlander, and Iwaniec, and which was recently extended to $x^{7/12}$ by Maynard. As an application, we obtain a new upper bound on the count of twin primes. Our method simplifies the 2004 argument of Wu, and gives the largest percentage improvement since the 1986 bound of Bombieri, Friedlander, and Iwaniec.

1. Introduction

Given a finite set \mathcal{A} of positive integers, sieve methods offer a broad framework for estimating the number of elements in \mathcal{A} all whose prime factors exceed z , denoted by $S(\mathcal{A}, z)$, in terms of the approximate density $g_{\mathcal{A}}(d) = g(d)$ of multiples of d in \mathcal{A} , denoted by \mathcal{A}_d . Note one often expects

$$S(\mathcal{A}, z) \approx |\mathcal{A}| \prod_{p < z} (1 - g(p)).$$

Combinatorial sieves may be viewed as refinements of the basic inclusion-exclusion principle, and are described by a sequence of weights $\lambda(d) \in \{-1, 0, 1\}$ supported on integers up to some level $D \geq 1$. We refer the reader to Opera de Cribro [Friedlander and Iwaniec 2010] for a more thorough introduction to the subject.

In particular, the upper bound weights $\lambda^+(d)$ for the linear sieve satisfy

$$S(\mathcal{A}, z) \leq |\mathcal{A}| \prod_{p < z} (1 - g(p)) \left(F \left(\frac{\log D}{\log z} \right) + o(1) \right) + \sum_{\substack{d \leq D \\ p | d \Rightarrow p < z}} \lambda^+(d) (|\mathcal{A}_d| - |\mathcal{A}| g(d)) \quad (1-1)$$

as $D \rightarrow \infty$, provided $g = g_{\mathcal{A}}$ satisfies some mild conditions. Here the function $F : \mathbb{R}_{\geq 1} \rightarrow \mathbb{R}_{\geq 1}$ is defined by a delay-differential equation, as in (2-5). For sets \mathcal{A} sufficiently equidistributed in arithmetic progressions the second sum over $d \leq D$ in (1-1) contributes negligibly, in which case the main term is

$$S(\mathcal{A}, z) \lesssim |\mathcal{A}| \prod_{p < z} (1 - g(p)) F(s),$$

MSC2020: primary 11N35, 11N36; secondary 11N05.

Keywords: linear sieve, well-factorable weights, level of distribution, switching principle, Buchstab identity.

© 2025 The Author, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via [Subscribe to Open](#).

where $z = D^{1/s}$. In fact, $F(s) \rightarrow 1$ as $s \rightarrow \infty$ so the main term confirms the naïve expectation in this case. Moreover, F is optimal in the sense that the bound (1-1) is attained sharply for a particular set \mathcal{A} .

We introduce this sieve theory setup more formally in Section 2.2 below, and define the sieve weights λ^+ explicitly in Section 3; see [Friedlander and Iwaniec 2010, Section 12] for further details on the linear sieve ($\beta = 2$), as well as [loc. cit., Section 11] for its generalization to the β -sieve.

The linear sieve is powerful when combined with equidistribution estimates which make the final sum in (1-1) small. For example, the Bombieri–Vinogradov theorem shows that for every $\varepsilon, A > 0$, letting $Q = x^{1/2-\varepsilon}$ we have

$$\sum_{q \leq Q} \sup_{(a,q)=1} \left| \pi(x; q, a) - \frac{\pi(x)}{\varphi(q)} \right| \ll_{\varepsilon, A} \frac{x}{(\log x)^A}. \quad (1-2)$$

So by taking $D = Q$, (1-1) can give a good upper bound when the set \mathcal{A} is related to the primes, such as when $\mathcal{A} = \{p+2 : p \leq x\}$, in which case (1-1) gives an upper bound for the count of twin primes.

The estimate (1-2) may be viewed as an assertion of the generalized Riemann hypothesis on average over moduli up to $Q = x^{1/2-\varepsilon}$. It remains an important open problem to extend the range to $Q = x^{1/2+\delta}$ for some fixed $\delta > 0$. Indeed, Elliott and Halberstam [1970] conjectured such an extension up to $Q = x^{1-\varepsilon}$ for any $\varepsilon > 0$.

In some contexts it suffices to relax the setup in (1-2) in order to raise the level of distribution. In particular, in the case of a fixed residue class $a \in \mathbb{Z}$, and the absolute values replaced by well-factorable weights $\lambda(q)$ (see Definition 2.1), the celebrated result of Bombieri, Friedlander and Iwaniec [Bombieri et al. 1986] raised the level up to $Q = x^{4/7-\varepsilon}$,

$$\sum_{\substack{q \leq Q \\ (q,a)=1}} \lambda(q) \left(\pi(x; q, a) - \frac{\pi(x)}{\varphi(q)} \right) \ll_{a, A, \varepsilon} \frac{x}{(\log x)^A}. \quad (1-3)$$

While the linear sieve weights are not themselves well-factorable, Iwaniec [1980] constructed a well-factorable variant $\tilde{\lambda}^+$ of the weights λ^+ (and so (1-3) holds with $\lambda = \tilde{\lambda}^+$), which are only slightly altered from λ^+ so that $\tilde{\lambda}^+$ enjoys an analogous linear sieve bound as in (1-1), notably with an identical form of the main term,

$$S(\mathcal{A}, z) \leq |\mathcal{A}| \prod_{p < z} (1 - g(p)) \left(F \left(\frac{\log D}{\log z} \right) + o(1) \right) + \sum_{\substack{d \leq D \\ p | d \Rightarrow p < z}} \tilde{\lambda}^+(d) (|\mathcal{A}_d| - |\mathcal{A}| g(d)). \quad (1-4)$$

The bound (1-3) stood for several decades, but quite recently Maynard [2020] managed to extend the level in (1-3) further to $Q = x^{7/12-\varepsilon}$ in the case of the weights $\lambda = \tilde{\lambda}^+$. Given the currently available equidistribution estimates for primes, we note the level $x^{7/12}$ is a natural barrier for these weights.

In this article, we modify the technical construction of the linear sieve weights to avoid this barrier, and thereby produce new sieve weights that induce stronger equidistribution estimates for primes.

Theorem 1.1. *Let $D = x^{10/17-\varepsilon}$. There exists a sequence $\tilde{\lambda}^*(d) \in \{-1, 0, 1\}$ satisfying:*

(1) *Equidistribution for primes: For any fixed $a \in \mathbb{Z}$, $A, \varepsilon > 0$, we have*

$$\sum_{\substack{d \leq D \\ (d,a)=1}} \tilde{\lambda}^*(d) \left(\pi(x; d, a) - \frac{\pi(x)}{\varphi(d)} \right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

(2) *Sieve upper bound: For $s \geq 1$, $z = D^{1/s}$, we have*

$$S(\mathcal{A}, z) \leq |\mathcal{A}| \prod_{p < z} (1 - g(p))(F^*(s) + o(1)) + \sum_{\substack{d \leq D \\ p|d \Rightarrow p < z}} \tilde{\lambda}^*(d)(|\mathcal{A}_d| - |\mathcal{A}|g(d)),$$

where $F^*(s) \leq 1.000081F(s)$ when $1 \leq s \leq 3$, for the linear sieve function F as in (2-5).

The key feature of Theorem 1.1 is to obtain equidistribution up to level $x^{10/17}$ at the cost of only a tiny loss in the main term. See Theorem 2.12 and Proposition 5.4 for full technical statements and additional variations that may be of independent interest.

1.1. Application to twin primes. We expect that Theorem 1.1 should give numerous improvements to sieve bounds related to the primes. As proof of concept in this direction, we give a new upper bound for the count of twin primes up to x , denoted by $\pi_2(x)$. Recall Hardy and Littlewood [1923] conjectured the asymptotic formula

$$\pi_2(x) \sim \frac{2x}{(\log x)^2} \prod_{p > 2} \frac{1 - 2/p}{(1 - 1/p)^2} =: \Pi(x). \quad (1-5)$$

Theorem 1.2. *As x tends to infinity, we have*

$$\pi_2(x) \lesssim 3.29956\Pi(x).$$

Theorem 1.2 gives a 2.94% refinement from the previous record bound of Wu [2004]. For reference, this gives the largest percentage improvement since the work of Bombieri, Friedlander, and Iwaniec [1986]. See Table 1 for a chronology of the known upper bounds on $\pi_2(x)/\Pi(x)$. Also see Siebert [1976], Riesel and Vaughan [1983, Lemma 5] for numerically explicit forms of Selberg's bound [1952].

The main ingredients for these results come from applying sieve bounds to the set $\mathcal{A} = \{p + 2 : p \leq x\}$, and using equidistribution of primes in arithmetic progressions to handle remainder terms. Bombieri and Davenport obtained $\pi_2(x)/\Pi(x) \lesssim 4$ as a consequence of the Bombieri–Vinogradov theorem (1-2) and a standard sieve upper bound of level $x^{1/2-\varepsilon}$. More generally, if one proves level of distribution $x^{\theta-\varepsilon}$ then one immediately obtains $\pi_2(x)/\Pi(x) \lesssim 2/\theta$. Bombieri, Friedlander and Iwaniec proved $\pi_2(x)/\Pi(x) \lesssim \frac{7}{2}$ by the well-factorable variant (1-3) level of distribution $x^{4/7-\varepsilon}$, together with the linear sieve with well-factorable remainder (1-4).

Year	Author(s)	$\pi_2(x)/\Pi(x) \lesssim$
1919	Brun [1919]	$O(1)$
1947	Selberg [1952]	8
1964	Pan [1964]	6
1966	Bombieri and Davenport [1966]	4
1978	Chen [1978]	3.9171
1983	Fouvry and Iwaniec [1983]	$3.7777 \dots = 34/9$
1984	Fouvry [1984]	$3.7647 \dots = 64/17$
1986	Bombieri, Friedlander and Iwaniec [1986]	3.5
1986	Fouvry and Grupp [1986]	3.454
1990	Wu [1990]	3.418
2003	Cai and Lu [2003]	3.406
2004	Wu [2004]	3.39951

Table 1. Upper bounds for $\pi_2(x)/\Pi(x)$.

The other key ingredient to subsequent improvements is the *switching principle*, introduced in Chen's celebrated result [1973] that there are infinitely many primes p such that $p + 2$ has at most two prime factors. The basic insight is to use a weighted sieve inequality to split the problem into multiple cases, apply sieve bounds to $\mathcal{A} = \{p + 2 : p \leq x\}$ in certain cases, and then reinterpret the remaining cases as new sieving problems for switched sets $\mathcal{B} = \{m - 2 \leq x\}$ where the numbers m are constructed from \mathcal{A} (as prescribed depending on the case).

1.2. Outline of main ideas in Theorem 1.1. Maynard's new equidistribution results show equidistribution of the primes with sieve weights $\tilde{\lambda}^+(d)$, provided $d = p_1 \cdots p_r$ is restricted to suitably well-factorable integers. Unfortunately, the original linear sieve weights only partially satisfy these well-factorable conditions. In particular for $\eta > 0$, when looking at the linear sieve of level $x^{7/12+\eta}$, some integers d in its support do not satisfy the conditions, which means that $x^{7/12}$ is the limit for the linear sieve given our current equidistribution technology. Nevertheless, the key observation here is that only a few exceptional d fail to satisfy these conditions. Moreover up to level $x^{10/17}$, i.e., $\eta < \frac{1}{204}$, the anatomy of exceptional d may be precisely characterized in terms of η (given specifically as $\mathcal{P}_4, \mathcal{P}_6$ in (3-6)). In particular, as $\eta > 0$ grows the family of exceptional integers contribute $O(\eta^5)$ to the sieve bound. However, we note this characterization breaks down when $\eta \geq \frac{1}{204}$, and the contribution becomes considerably larger and more complicated.

As such we carefully revise the construction of the linear sieve, altering a few particular inclusion-exclusion steps in order to avoid the exceptional integers d with bad factorizations. Once these terms no longer contribute to the sieve, this produces a worse and more complicated main term, but since there are only a very small number of such terms the resulting loss is small. And since these modified weights now satisfy stronger factorization properties in their support, we can now leverage the full strength of Maynard's equidistribution results.

Notation

We use the Vinogradov \ll and \gg asymptotic notation, and the big oh $O(\cdot)$ and $o(\cdot)$ asymptotic notation. We use $f \sim g$, $f \lesssim g$, and $f \gtrsim g$ to denote $f = (1 + o(1))g$, $f \leq (1 + o(1))g$, and $f \geq (1 + o(1))g$, respectively. Dependence on a parameter will be denoted by a subscript.

The letter p will always be reserved to denote a prime number, $\pi(x)$ is the prime counting function, and $\pi(x; d, a)$ is the count of primes up to x congruent to $a \pmod{d}$. We use φ to denote the Euler totient function, μ the Möbius function, and $e(x) := e^{2\pi i x}$ the complex exponential. We use $\mathbf{1}$ to denote the indicator function of a statement. For example, for a set A denote

$$\mathbf{1}_{a \in A} = \begin{cases} 1, & \text{if } a \in A, \\ 0, & \text{else,} \end{cases} \quad \mathbf{1}_{a_1, \dots, a_i \in A}^{a_0 \notin A} = \begin{cases} 1, & \text{if } a_1, \dots, a_i \in A \text{ and } a_0 \notin A, \\ 0, & \text{else.} \end{cases}$$

Finally, we refer to various sieve weights referred places throughout the article, so we take a moment to list them here:

We generically write λ to denote a sequence of weights in $\{0, \pm 1\}$. In particular, λ^+ and λ^- refer to the (upper and lower bound) weights of the linear sieve, given by restrictions of the Möbius function, $\lambda^\pm(d) = \mu(d)\mathbf{1}_{d \in \mathcal{D}^\pm}$. Analogously, the modified (upper bound) linear sieve weights λ^* are given by $\lambda^*(d) = \mu(d)\mathbf{1}_{d \in \mathcal{D}^*}$. Here the support sets \mathcal{D}^\pm and \mathcal{D}^* are defined in (3-1) and (3-5). We also write $\lambda^{(r)}$ to refer to the weights λ^+ or λ^- (and $\mathcal{D}^{(r)}$ to refer to \mathcal{D}^+ or \mathcal{D}^-), depending on whether r is odd or even.

The well-factorable weights $\tilde{\lambda}^\pm$ are defined in (5-17). Following Iwaniec, this construction involves certain auxiliary weights at intermediate steps, namely, $\lambda_{(D_1, \dots, D_r)}$ defined in (5-14), and $\lambda_{(D_1, \dots, D_r)}^{(r)} = \lambda_{(D_1, \dots, D_r)} * \lambda^{(r)}$ defined in (5-15). The analogous construction starting from λ^* gives modified weights $\tilde{\lambda}^*$, defined in (5-16).

2. Technical setup and results

2.1. Factorization of weights and their level of distribution.

Definition 2.1 (well-factorable). Let $Q \in \mathbb{R}_{\geq 1}$. A sequence $\lambda(q)$ is *well-factorable of level Q* , if for every factorization $Q = Q_1 Q_2$ into $Q_1, Q_2 \in \mathbb{R}_{\geq 1}$, there exist sequences γ_1, γ_2 such that:

- (1) $|\gamma_1(q_1)|, |\gamma_2(q_2)| \leq 1$ for all $q_1, q_2 \in \mathbb{N}$.
- (2) $\gamma_i(q) = 0$ if $q \notin [1, Q_i]$ for $i = 1, 2$.
- (3) We have $\lambda = \gamma_1 * \gamma_2$, i.e.,

$$\lambda(q) = \sum_{q=q_1 q_2} \gamma_1(q_1) \gamma_2(q_2).$$

Bombieri, Friedlander and Iwaniec [1986, Theorem 10] established level of distribution $x^{4/7-\varepsilon}$ with well-factorable weights.

Theorem 2.2 [Bombieri et al. 1986]. *Fix any $a \in \mathbb{Z}$ and let $A, \varepsilon > 0$. For any well-factorable sequence λ of level $Q \leq x^{A/7-\varepsilon}$, we have*

$$\sum_{\substack{q \leq Q \\ (q,a)=1}} \lambda(q) \left(\pi(x; q, a) - \frac{\pi(x)}{\varphi(q)} \right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

Maynard [2020] considered a natural strengthening of well-factorable sequences.

Definition 2.3 (triply well-factorable). Let $Q \in \mathbb{R}_{\geq 1}$. A sequence $\lambda(q)$ is *triply well-factorable of level Q* , if for every factorization $Q = Q_1 Q_2 Q_3$ into $Q_1, Q_2, Q_3 \in \mathbb{R}_{\geq 1}$, there exist sequences $\gamma_1, \gamma_2, \gamma_3$ such that:

- (1) $|\gamma_1(q_1)|, |\gamma_2(q_2)|, |\gamma_3(q_3)| \leq 1$ for all $q_1, q_2, q_3 \in \mathbb{N}$.
- (2) $\gamma_i(q) = 0$ if $q \notin [1, Q_i]$ for $i = 1, 2, 3$.
- (3) We have $\lambda = \gamma_1 * \gamma_2 * \gamma_3$, i.e.,

$$\lambda(q) = \sum_{q=q_1 q_2 q_3} \gamma_1(q_1) \gamma_2(q_2) \gamma_3(q_3).$$

The definitions of well-factorable and triply well-factorable sequences are quite natural and relatively simple from a conceptual standpoint. Maynard [2020, Theorem 1.1] obtains powerful equidistribution results for triply well-factorability that are beyond the scope of well-factorability. Unfortunately, triply well-factorability is too restrictive a condition for us to produce Theorem 1.1. As such we are forced to identify the precise mechanism that enables Maynard's equidistribution results, and extract the following technical definition that is implicit in [Maynard 2020].¹

Definition 2.4 (programmably factorable). Let $0 < \delta < 10^{-5}$. For $x \in \mathbb{R}_{> 1}$, a sequence $\lambda(q)$ is *programmably factorable of level Q (relative to x, δ)*, if for every $N \in [x^{2\delta}, x^{1/3+\delta/2}]$ there exists a factorization $Q = Q_1 Q_2 Q_3$ with $Q_1, Q_2, Q_3 \in \mathbb{R}_{\geq 1}$, satisfying the system

$$\begin{aligned} Q_1 &\leq Nx^{-\delta}, \\ N^2 Q_2 Q_3^2 &\leq x^{1-\delta}, \\ N^2 Q_1 Q_2^4 Q_3^3 &\leq x^{2-\delta}, \\ N Q_1 Q_2^5 Q_3^2 &\leq x^{2-\delta}. \end{aligned} \tag{2-1}$$

And for every such factorization $Q = Q_1 Q_2 Q_3$ there exist sequences $\gamma_1, \gamma_2, \gamma_3$ such that:

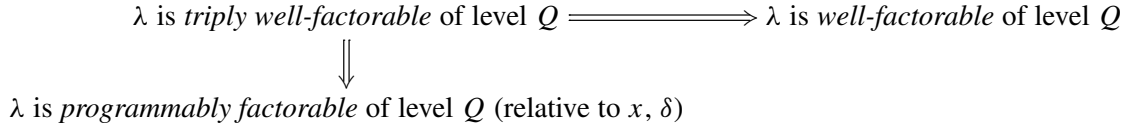
- (1) $|\gamma_1(q_1)|, |\gamma_2(q_2)|, |\gamma_3(q_3)| \leq 1$ for all $q_1, q_2, q_3 \in \mathbb{N}$.
- (2) $\gamma_i(q) = 0$ if $q \notin [1, Q_i]$ for $i = 1, 2, 3$.

¹Indeed, the definition of programmably factorable in the special case $Q_3 = 1$ gives the implicit condition (which is implied by well-factorable) that enables Bombieri, Friedlander and Iwaniec to get equidistribution (1-3); also see Lemma 5 in [Fouvry and Grupp 1986].

(3) We have $\lambda = \gamma_1 * \gamma_2 * \gamma_3$, i.e.,

$$\lambda(q) = \sum_{q=q_1q_2q_3} \gamma_1(q_1)\gamma_2(q_2)\gamma_3(q_3).$$

Programmable factorability is the key technical definition in this article. It is named in allusion to the linear programming-type system of inequalities (2-1) that the factors satisfy. The diagram below displays the various implications among the definitions:



In the key result [Maynard 2020, Theorem 1.1], Maynard extended the level of distribution up to $Q < x^{3/5}$ for programmably factorable weights. Note that level $x^{3/5}$ is the natural barrier for (2-1) to admit a solution.

Theorem 2.5 [Maynard 2020]. *Fix any $a \in \mathbb{Z}$ and let $A, \varepsilon > 0$. For any programmably factorable sequence λ of level $Q \leq x^{3/5-\varepsilon}$ (relative to $x, \varepsilon/50$), we have*

$$\sum_{\substack{q \leq Q \\ (q,a)=1}} \lambda(q) \left(\pi(x; q, a) - \frac{\pi(x)}{\varphi(q)} \right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

Remark 2.6. Theorem 1.1 in [Maynard 2020] was stated for triply factorable sequences, but its proof in fact gives the result for programmably factorable sequences.

Note the weights $\tilde{\lambda}^+$ are composed of well-factorable — but not necessarily programmably factorable — sequences of given level D . Nevertheless, Maynard showed the upper bound weights $\tilde{\lambda}^+$ of sieve level $D = x^{7/12-\varepsilon}$ are programmably factorable of level $D \leq Q = x^{3/5-\varepsilon}$ (relative to $x, \varepsilon/50$). By Theorem 2.5 this gives [Maynard 2020, Theorem 1.2] below.

Corollary 2.7 [Maynard 2020]. *For any fixed $a \in \mathbb{Z}$ and $A, \varepsilon > 0$, the weights $\tilde{\lambda}^+$ from (2-4) of sieve level $D = x^{7/12-\varepsilon}$ satisfy*

$$\sum_{\substack{d \leq D \\ (d,a)=1}} \tilde{\lambda}^+(d) \left(\pi(x; d, a) - \frac{\pi(x)}{\varphi(d)} \right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

Later, in Proposition 5.4, we shall obtain technical improvements of Corollary 2.7 for Iwaniec’s weights $\tilde{\lambda}^\pm$ (both upper and lower), in special cases where equidistribution is restricted to moduli which are smooth, or otherwise amenable to programmable factorization.

We may summarize the definitions and results of the section up to this point as follows:

λ is *triply well-factorable* of level Q :

- Equidistributed for $Q < x^{3/5}$.
- Can take $\lambda = \tilde{\lambda}^+$ for $D \leq Q^{2/3}$.

λ is *well-factorable* of level Q :

- Equidistributed for $Q < x^{4/7}$.
- Can take $\lambda = \tilde{\lambda}^+$ for $D \leq Q$.

λ is *programmably factorable* of level Q (relative to x, δ):

- Equidistributed for $Q < x^{3/5}$.
- Can take $\lambda = \tilde{\lambda}^+$ for $D \leq Q < x^{7/12}$.

For each type of sequence, we have outlined their corresponding levels of distribution, and the levels at which the type is satisfied by the upper bound weights for the linear sieve. Observe well-factorability is flexible enough to accommodate the linear sieve to any level, but has weaker equidistribution. On the other hand, triple well-factorability has stronger equidistribution, but is too rigid to accommodate the linear sieve (at nontrivial levels). Finally, programmable factorability also has strong equidistribution in addition to (nontrivially) accommodating the linear sieve, though at the cost of conceptual technicality.

Remark 2.8. In general, λ well-factorable of level Q directly implies λ triply well-factorable of level $Q^{2/3}$.² In particular, for $\lambda = \tilde{\lambda}^+$ the triply well-factorable level $Q^{2/3} < x^{2/5}$ is sharp.³

2.2. Sieve theory setup and bounds. We recall the standard sieve-theoretic notation. Given a finite set $\mathcal{A} \subset \mathbb{N}$, set of primes \mathcal{P} , and a threshold $z > 0$, we define $\mathcal{A}_d = \{n \in \mathcal{A} : d \mid n\}$ and remainder $r_{\mathcal{A}}$ via

$$|\mathcal{A}_d| = g(d)|\mathcal{A}| + r_{\mathcal{A}}(d),$$

where g is a multiplicative function, with $0 \leq g(p) < 1$ for $p \in \mathcal{P}$ (we assume $g(p) = 0$ if $p \notin \mathcal{P}$). Also define $P(z) = \prod_{p < z, p \in \mathcal{P}} p$ and $V(z) = \prod_{p \mid P(z)} (1 - g(p))$. The central object of interest is the *sifted sum*

$$S(\mathcal{A}, z) = S(\mathcal{A}, \mathcal{P}, z) = \sum_{n \in \mathcal{A}} \mathbf{1}_{(n, P(z))=1}. \quad (2-2)$$

Later for our application of interest, we will set $g(d) = 1/\varphi(d)$. For now, it suffices for us to assume for all $2 \leq w \leq z$,

$$\frac{V(w)}{V(z)} = \prod_{\substack{w \leq p < z \\ p \in \mathcal{P}}} (1 - g(p)) = \frac{\log z}{\log w} \left(1 + O\left(\frac{1}{\log w}\right) \right). \quad (2-3)$$

Remark 2.9. The proof of the upper bound for the standard linear sieve only requires a one-sided inequality for $V(w)/V(z)$, whereas our modification requires the above two-sided condition (2-3).

²Indeed, take any factorization $Q = Q_1 Q_2 Q_3$, with (say) $Q_1 \geq Q_2 \geq Q_3 \geq 1$. Note $Q_3 \leq Q^{1/3}$. If λ is well-factorability of level $Q/Q_3 = Q_1 Q_2$, there are sequences γ_1, γ_2 supported on $[1, Q_1], [1, Q_2]$ with $\lambda = \gamma_1 * \gamma_2 = \gamma_1 * \gamma_2 * \delta$. Here $\delta(q) = \mathbf{1}_{q=1}$. Hence λ is triply well-factorable of level $\inf_{Q=Q_1 Q_2 Q_3} Q/Q_3 \geq Q^{2/3}$.

³Indeed, consider the factorization $Q = Q_1 Q_2 Q_3$ with $(Q_1, Q_2, Q_3) = (Q^{1/3-\varepsilon}, Q^{1/3-\varepsilon}, Q^{1/3+2\varepsilon})$. Then for $q = p_1 p_2 p_3$ of size $p_1, p_2 \sim Q^{1/3}, p_3 \sim Q^{1/9}$, we see $p_1, p_2 > Q_1 = Q_2$. Thus all sequences γ_i supported on Q_i satisfy $\gamma_1 * \gamma_2 * \gamma_3(q) = 0$. In particular $\gamma_1 * \gamma_2 * \gamma_3 \neq \tilde{\lambda}^+$.

The basic result which we shall adapt is the linear sieve with well-factorable remainder, as in [Friedlander and Iwaniec 2010, Theorem 12.20].

Theorem 2.10 [Friedlander and Iwaniec 2010]. *Let $\varepsilon > 0$ and $D > 1$ be sufficiently small and large, respectively. Then for $s \geq 1$ and $z = D^{1/s}$, we have*

$$\begin{aligned} S(\mathcal{A}, z) &\leq |\mathcal{A}|V(z)(F(s) + O(\varepsilon)) + \sum_{d|P(z)} \tilde{\lambda}^+(d)r_{\mathcal{A}}(d), \\ S(\mathcal{A}, z) &\geq |\mathcal{A}|V(z)(f(s) + O(\varepsilon)) - \sum_{d|P(z)} \tilde{\lambda}^-(d)r_{\mathcal{A}}(d), \end{aligned}$$

where the implied constant only depends that of (2-3). Here the weights $\tilde{\lambda}^{\pm}$ are

$$\tilde{\lambda}^{\pm}(d) = \sum_{j \leq \exp(\varepsilon^{-3})} \lambda_j^{\pm}(d) \quad (2-4)$$

for some well-factorable sequences λ_j^{\pm} of level D . The functions $F, f : \mathbb{R}^+ \rightarrow \mathbb{R}$ satisfy the system of delay-differential equations

$$\begin{aligned} sF(s) &= 2e^{\gamma} \quad [s \leq 3] \quad (sF(s))' = f(s-1), \\ sf(s) &= 0 \quad [s \leq 2] \quad (sf(s))' = F(s-1). \end{aligned} \quad (2-5)$$

Remark 2.11. See [Iwaniec 1980, Theorem 1] for an alternate formulation and proof, which gives sharper quantitative bounds than $O(\varepsilon)$. However, it is more technical than necessary for our purposes.

The main result of this article is the following modification of the linear sieve with programmably factorable remainder.

Theorem 2.12. *Let \mathcal{A} be a finite set of positive integers with density function $g(d)$ satisfying (2-3), and $F(s)$ the function defined by the system (2-5). Let $\varepsilon > 0$ and $x > 1$ be sufficiently small and large, respectively. Then for $\eta \geq 0$, $D = x^{7/12+\eta}$, $s \geq 1$, and $z = D^{1/s}$, we have*

$$S(\mathcal{A}, z) \leq |\mathcal{A}|V(z)(F^*(s) + O(\varepsilon)) + \sum_{d|P(z)} \tilde{\lambda}^*(d)r_{\mathcal{A}}(d),$$

where the implied constant only depends that of (2-3). Here the weights $\tilde{\lambda}^*$ are

$$\tilde{\lambda}^*(d) = \sum_{j \leq \exp(\varepsilon^{-3})} \lambda_j^*(d) \quad (2-6)$$

for some programmably factorable sequences λ_j^* of level D (relative to $x, \varepsilon/50$). For $\eta < \frac{1}{204}$ we have $F^*(s) = F(s) + O(\eta^5)$, and $F^*(s) \leq 1.000081F(s)$ for $1 \leq s \leq 3$, $\eta = \frac{1}{204}$.

Note Theorem 2.5, applied to each $\lambda = \lambda_j^*$ above, immediately implies the following.

Corollary 2.13. *Given any fixed $a \in \mathbb{Z}$ and A . For $\eta < \frac{1}{204}$, the weights $\tilde{\lambda}^*$ as in (2-6) of level $D = x^{7/12+\eta}$ satisfy*

$$\sum_{\substack{d \leq D \\ (d,a)=1}} \tilde{\lambda}^*(d) \left(\pi(x; d, a) - \frac{\pi(x)}{\varphi(d)} \right) \ll_{a,A,\eta} \frac{x}{(\log x)^A}.$$

3. Programmably factorable support

The upper and lower bound weights λ^\pm for the linear sieve of level D are defined by

$$\lambda^\pm(d) = \mu(d) \mathbf{1}_{d \in \mathcal{D}^\pm},$$

where $\mathcal{D}^\pm = \mathcal{D}^\pm(D)$ are the standard support sets

$$\begin{aligned} \mathcal{D}^+ &= \{p_1 \cdots p_r : D^{1/2} \geq p_1 \geq \cdots \geq p_r, \text{ and } p_1 \cdots p_{l-1} p_l^3 \leq D \text{ for each odd } l \leq r\}, \\ \mathcal{D}^- &= \{p_1 \cdots p_r : D^{1/2} \geq p_1 \geq \cdots \geq p_r, \text{ and } p_1 \cdots p_{l-1} p_l^3 \leq D \text{ for each even } l \leq r\}. \end{aligned} \quad (3-1)$$

We may also write $\mathcal{D}^{(r)}$ to denote \mathcal{D}^+ or \mathcal{D}^- , when r is even or odd, respectively.

Observe that both sets satisfy the containment $\mathcal{D}^\pm(D) \subset \mathcal{D}^{\text{well}}(D)$, where

$$\mathcal{D}^{\text{well}} = \{p_1 \cdots p_r : D^{1/2} \geq p_1 \geq \cdots \geq p_r \text{ and } p_1 \cdots p_{l-1} p_l^2 \leq D \text{ for each } l \leq r\}. \quad (3-2)$$

We shall return to this observation later in the section.

Maynard [2020] deduces Corollary 2.7 for $\tilde{\lambda}^+$ from the general Theorem 2.5 by means of the following key result [loc. cit., Proposition 9.1] (along with a construction of Iwaniec we shall address in later sections), which programmably factorizes elements of the support \mathcal{D}^+ .

Proposition 3.1 [Maynard 2020]. *Let $0 < \delta < 10^{-3}$ and let $D = x^{7/12-50\delta}$, $N \in [x^{2\delta}, x^{1/3+\delta/2}]$. Then every $d \in \mathcal{D}^+(D)$ has a factorization $d = d_1 d_2 d_3$ such that $d_1 \leq N x^{-\delta}$ and*

$$N^2 d_2 d_3^2 \leq x^{1-\delta}, \quad N^2 d_1 d_2^4 d_3^3 \leq x^{2-\delta}, \quad N d_1 d_2^5 d_3^2 \leq x^{2-\delta}. \quad (3-3)$$

Remark 3.2. The level $x^{7/12}$ is sharp in this construction. Indeed, *heuristically speaking*, the linear sieve weights are not programmably factorable of level $D = x^{7/12+\eta}$ for any $\eta > 0$, because the support set contains obstructing (families of) elements $d \in \mathcal{D}^+(D)$ of the form $d = p_1 \cdots p_r$ where $p_1 \approx \cdots \approx p_6 \approx D^{1/7}$, or where $p_1 \approx p_2 \approx D^{2/7}$ and $p_3 \approx p_4 \approx D^{1/7}$. This heuristic description of the obstructions is made precise by the families $\mathcal{P}_4, \mathcal{P}_6$ (in (3-6) below), and thereby tells us how we should restrict the support set in order to increase the level (namely, to \mathcal{D}^* in (3-5) below).

For $\eta > 0$, level $D = x^{7/12+\eta}$, we define the modified weights $\lambda^* = \lambda_\eta^*$,

$$\lambda^*(d) = \mu(d) \mathbf{1}_{d \in \mathcal{D}^*}, \quad (3-4)$$

for the support set \mathcal{D}^* ,

$$\mathcal{D}^* = \mathcal{D}^+(x^{7/12}) \cup \{p_1 \cdots p_r \in \mathcal{D}^+(x^{7/12+\eta}) : p_1 \cdots p_i \notin \mathcal{P}_i, i \leq r, i \in \{4, 6\}\}. \quad (3-5)$$

Here \mathcal{P}_4 and $\mathcal{P}_6 = \mathcal{P}_{6,1} \cup \mathcal{P}_{6,2}$ are exceptional subsets of $\mathcal{D}^+(x^{7/12+\eta})$, given by

$$\begin{aligned} \mathcal{P}_4 &= \{p_1 \cdots p_4 : p_1 < x^{1/6+2\eta} \text{ and } p_2 p_4 > x^{1/4-3\eta}\}, \\ \mathcal{P}_{6,1} &= \{p_1 \cdots p_6 : p_1 p_2 < x^{1/6+2\eta} \text{ and } p_2 p_3 p_4 > x^{1/4-3\eta} \text{ and } p_6 > x^{1/12-5\eta}\}, \\ \mathcal{P}_{6,2} &= \{p_1 \cdots p_6 : p_1, p_2 p_3 < x^{1/6+2\eta} \text{ and } p_1 p_4, p_2 p_3 p_4 > x^{1/4-3\eta} \text{ and } p_6 > x^{1/12-5\eta}\}. \end{aligned} \quad (3-6)$$

The modified support set $\mathcal{D}^* = \mathcal{D}_\eta^*$ is understood to depend on $\eta > 0$ (as do $\mathcal{P}_4, \mathcal{P}_6$), but we will suppress this for notational convenience.

In this section, we establish a programmable factorization of the elements of the support \mathcal{D}^* provided $D < x^{10/17}$, i.e., $\eta < \frac{1}{204}$. This will serve as the key technical input for the proof of Theorem 2.12.

Proposition 3.3 (factorization of elements of \mathcal{D}^*). *Let $0 < \delta < 10^{-5}$, and take $0 < \eta < \frac{1}{204} - 3\delta$ and $N \in [x^\delta, x^{1/3-\delta/2}]$. If $d \in \mathcal{D}^*$ for $D = x^{7/12+\eta-50\delta}$, then we may factor $d = d_1 d_2 d_3$ such that $d_1 \leq Nx^{-\delta}$ and*

$$N^2 d_2 d_3^2 \leq x^{1-\delta}, \quad N^2 d_1 d_2^4 d_3^3 \leq x^{2-\delta}, \quad N d_1 d_2^5 d_3^2 \leq x^{2-\delta}. \quad (3-7)$$

On the first attempt working through technicalities, we encourage the reader to set $\delta = 0$ in order to better view the key features.

Before proving the proposition, we need some lemmas. The first gives a general-purpose criterion to factor an integer d .

Lemma 3.4. *Let $D = x^{7/12+\eta-50\delta}$ for $-\frac{1}{84} < \eta < \frac{1}{60}$. A factorization $d = d_1 d_2 d_3$ satisfies (3-7), provided $d_1, d_2, d_3 \geq 1$ satisfy*

$$d_2 \in [x^{1/6+2\eta}, x^{1/4-3\eta}], \quad d_1 \leq Nx^{-\delta} \text{ and } d_3 \leq D/Nd_2. \quad (3-8)$$

Proof. By (3-8), $Nd_3 \leq D/d_2$ and so

$$\begin{aligned} N^2 d_2 d_3^2 &\leq D^2/d_2 \leq x^{2(7/12+\eta-50\delta)-(1/6+2\eta)} = x^{1-\delta}, \\ N^2 d_1 d_2^4 d_3^3 &\leq D^3 d_2 \leq x^{3(7/12+\eta-50\delta)+(1/4-3\eta)} = x^{2-\delta}, \\ N d_1 d_2^5 d_3^2 &\leq D^2 d_3^2 \leq x^{2(7/12+\eta-50\delta)+3(1/4-3\eta)} = x^{23/12-7\eta-100\delta} < x^{2-\delta}, \end{aligned}$$

using $\eta \in (-\frac{1}{84}, \frac{1}{60})$. This gives (3-7). □

The above criterion implies factorizations in the following special cases.

Lemma 3.5. *Let $D = x^{7/12+\eta-50\delta}$ for $\eta < \frac{1}{60}$. For $r \geq 4$, let $x^{1/6+2\eta} > p_1 \geq \dots \geq p_r$ be primes for which $d = p_1 \cdots p_r \in \mathcal{D}^+(D)$. Suppose d_2 is one of the subproducts in $\{p_1 p_4, p_2 p_3, p_2 p_4, p_2 p_3 p_4\}$. Then d has a factorization $d = d_1 d_2 d_3$ satisfying (3-7), provided*

$$d_2 \in [x^{1/6+2\eta}, x^{1/4-3\eta}].$$

Proof. Let $C = D/Nd_2$ and note either $p_1 \leq N$ or $p_1 \leq C$, since

$$p_1^2 \leq D^{2/3} = x^{2/3(7/12+\eta-50\delta)} < x^{1/3+4\eta-50\delta} \leq D/d_2 = NC.$$

We proceed by induction on $r \geq 4$. As the base case $r = 4$, by Lemma 3.4 it suffices for each b to factor $p_1 \cdots p_4/d_2 = d_1 d_3$ for $d_1 \leq N$, $d_3 \leq C$. Indeed, this holds when $d_2 = p_2 p_3 p_4$ since $p_1^2 \leq AC$, and similarly:

- If $d_2 = p_2 p_4$ then $p_3^2 \leq D/p_1 p_2 p_3 \leq NC/p_1$ implies $p_1 p_3 = d_1 d_3$ for some $d_1 \leq N$, $d_3 \leq C$.
- If $d_2 = p_1 p_4$ then $p_3^2 \leq D/p_1 p_2 p_3 \leq NC/p_2$ implies $p_2 p_3 = d_1 d_3$ for some $d_1 \leq N$, $d_3 \leq C$.
- If $d_2 = p_2 p_3$ then $p_4^2 \leq D/p_1 p_2 p_3 \leq NC/p_1$ implies $p_1 p_4 = d_1 d_3$ for some $d_1 \leq N$, $d_3 \leq C$.

Now for $r \geq 5$, we inductively assume a factorization $p_1 \cdots p_{r-1} = d_1 d_2 d_3$ with $d_1 \leq N$, $d_3 \leq C$. Then $p_r^2 \leq D/p_1 \cdots p_{r-1} = NC/(ac)$ so either $d_1 p_r \leq N$ or $d_3 p_r \leq C$, extending the factorization. Hence Lemma 3.4 applies again, and completes the proof. \square

Finally, if the primes dividing d are small enough, we may use the greedy algorithm to factor d as follows.

Lemma 3.6. *Let $D = x^{7/12+\eta-50\delta}$ for $\eta < \frac{1}{60}$. For $r \geq 4$, let $x^{1/6+2\eta} > p_1 \geq \dots \geq p_r$ be primes for which $d = p_1 \cdots p_r \in \mathcal{D}^+(D)$, and $p_6 < x^{1/12-5\eta}$ if $r \geq 6$. Then d has a factorization $d = abc$ satisfying (3-7), provided there is a factorization $p_1 p_2 p_3 p_4 = d_1 d_2 d_3$ satisfying*

$$d_1 \leq Nx^{-\delta}, \quad d_3 \leq x^{1-2\delta}/DN, \quad d_2 \leq D^2/x^{1-3\delta} = x^{1/6+2\eta+3\delta}. \quad (3-9)$$

Proof. Let $D_1 = Nx^{-\delta}$, $D_2 = D^2/x^{1-3\delta}$, $D_3 = x^{1-2\delta}/(DN)$, so that $d_i \leq D_i$ by assumption.

We now greedily append primes to d_i while preserving $d_i \leq D_i$ for all i , i.e., where at the j -th step we replace $d_i \mapsto d_i p_j$ (for one of $i = 1, 2, 3$) provided $d_i p_j \leq D_i$. Starting from $j = 5$, we stop either when we have exhausted all primes (i.e., $j = r$), or $d_i p_j > D_i$ for each $i = 1, 2, 3$. In the former case, we have the desired $d_1 d_2 d_3 = d = p_1 \cdots p_r$ and $d_i \leq D_i$ so we easily get

$$\begin{aligned} D_1 &= Nx^{-\delta}, \\ N^2 D_2 D_3^2 &= x^{1-\delta}, \\ N^2 D_1 D_2^4 D_3^3 &= D^5 x^{-1+5\delta} \leq x^{5 \cdot (3/5) - 1 - 245\delta} < x^{2-\delta}, \\ N D_1 D_2^5 D_3^2 &= D^8 x^{-3+10\delta} \leq x^{8 \cdot (3/5) - 3 - 390\delta} < x^{2-\delta}. \end{aligned}$$

Thus $d_1 d_2 d_3 = d = p_1 \cdots p_r$ gives the desired factorization.

In the latter case, there exists a terminal index $j < r$ for which $d_i p_j > D_i$ for all $i = 1, 2, 3$. Note if j is odd, then $d_i p_j \leq D_i$ for some i , since

$$p_j^3 \leq \frac{D}{p_1 \cdots p_{j-1}} = \frac{D_1 D_2 D_3}{d_1 d_2 d_3}.$$

So the terminal j is even with $j \geq 6$. By assumption $p_j \leq p_6 \leq x^{1/12-5\eta}$ is smaller than the width of the interval $[x^{1/6+2\eta}, x^{1/4-3\eta}]$. And since $d_2 \leq D_2 = x^{1/6+2\eta} < d_2 p_j$, we deduce $e_2 := d_2 p_j$ lies in the interval $e_2 \in [x^{1/6+2\eta}, x^{1/4-3\eta}]$.

Thus letting $E_3 := D_2 D_3 / e_2$, for each $l > j$ in turn we shall greedily append the prime p_l onto either d_1 or d_3 while preserving $d_1 < D_1$ and $d_3 < E_3$. Indeed, for all $l > j$,

$$p_l^2 \leq \frac{D}{p_1 \cdots p_{l-1}} = \frac{D_1 D_2 D_3}{d_1 d_2 d_3 p_j \cdots p_{l-1}} \leq \frac{D_1 E_3}{d_1 d_3 p_{j+1} \cdots p_{l-1}},$$

so there is a factorization $e_1 e_3 = d_1 d_3 p_{j+1} \cdots p_l$ with $e_1 \leq D_1 = N$ and $e_3 \leq E_3 = D / (A e_2 x^{2\delta})$. Hence the result now follows by Lemma 3.4 for the factorization $e_1 e_2 e_3 = d_1 d_2 d_3 p_j \cdots p_l = p_1 \cdots p_l$. \square

Proof of Proposition 3.3. We shall consider 3 cases, depending on the sizes of p_1 and $p_2 p_3$ compared to the endpoints of the key interval $[x^{1/6+2\eta}, x^{1/4-3\eta}]$.

Case 1 ($p_1 \geq D^2/x = x^{1/6+2\eta}$). Let $d_2 := p_1$, $C := D/Nd_2$. Note $C = D/Nd_2 \geq D^{2/3}/N \geq 1$.

Next $D \geq p_1^3 \geq p_1 p_2^2$ implies $p_2^2 \leq D/p_1 = NC$, so either $p_2 \leq N$ or $p_2 \leq C$. Similarly, since $p_1 \cdots p_{j-1} p_j^2 \leq D$ for all $j \leq r$, we get $p_j^2 \leq AC/(p_2 \cdots p_{j-1})$ for $3 \leq j \leq r$. As such, we may factor $p_2 \cdots p_r = d_1 d_3$ for $d_1 \leq N$, $d_3 \leq C$. Hence by Lemma 3.4 $p_1 \cdots p_r = d_1 d_2 d_3$ satisfies (3-7).

In the remaining cases, we assume $p_1 < x^{1/6+2\eta}$. By Lemma 3.5, it remains to consider $p_2 p_3 > x^{1/4-3\eta}$ or $p_2 p_3 < x^{1/6+2\eta}$. Note

$$p_2 p_3 \leq p_1^{1/3} (p_1 p_2 p_3^3)^{1/3} \leq (x^{1/6+2\eta})^{1/3} D^{1/3} = x^{1/3(1/6+7/12+3\eta-50\eta)} < x^{1/4+\eta-16\delta}. \quad (3-10)$$

Case 2 ($p_2 p_3 > x^{1/4-3\eta}$ and $p_1 < x^{1/6+2\eta}$). The proof follows by Lemma 3.5 if $p_2 p_4 \in [x^{1/6+2\eta}, x^{1/4-3\eta}]$. Thus by definition of \mathcal{P}_4 , in this case we may assume

$$p_2 p_4 < x^{1/6+2\eta}. \quad (3-11)$$

Hence we have $p_4 < x^{12\eta+50\delta}$, since

$$p_2 > p_2 (p_1 p_2 p_3^3) / D > (p_2 p_3)^3 / D > x^{3((1/4)-3\eta)} / D = x^{1/6-10\eta+50\delta}. \quad (3-12)$$

If $p_1 p_4 > x^{1/6+2\eta}$, then the proof follows by Lemma 3.5 where $d_2 = p_1 p_4$ is $< x^{(1/6+2\eta)+12\eta+50\delta} < x^{1/4-3\eta}$, since $\eta < \frac{1}{204} - 3\delta$.

Else $p_1 p_4 < x^{1/6+2\eta}$. We shall apply Lemma 3.6 with $d_2 = p_1 p_4$.

If either $Nx^{-\delta}$ or $x^{1-2\delta}/DN$ is greater than $x^{1/4+\eta-16\delta} \geq p_2 p_3$, by (3-10), then Lemma 3.6 completes the proof with $(d_1, d_3) = (p_2 p_3, 1)$ or $(1, p_2 p_3)$, respectively. Otherwise, $Nx^{-\delta}, x^{1-2\delta}/DN \in [x^{1/6-2\eta-64\delta}, x^{1/4+\eta-16\delta}]$, since $x/D = x^{5/12-\eta+50\delta}$. But then, using $\eta < \frac{1}{108}$,

$$\begin{aligned} \max(Nx^{-\delta}, x^{1-2\delta}/DN) &\geq (x^{1-2\delta}/D)^{1/2} = x^{(5/2)/12-\eta/2+24\delta} > x^{1/6+2\eta} > p_2, \\ \min(Nx^{-\delta}, x^{1-2\delta}/DN) &\geq x^{1/6-2\eta-64\delta} > x^{1/8+\eta/2-8\delta} \geq (p_2 p_3)^{1/2} \geq p_3, \end{aligned} \quad (3-13)$$

by (3-10), which suffices again for Lemma 3.6. Note $p_6 < x^{12\eta+50\delta} < x^{1/12-5\eta}$ when $r \geq 6$, using $\eta < \frac{1}{204} - 3\delta$.

Case 3 ($p_2 p_3 < x^{1/6+2\eta}$ and $p_1 < x^{1/6+2\eta}$). By Lemma 3.5, it suffices to consider either $p_1 p_4 < x^{1/6+2\eta}$ or $p_1 p_4 > x^{1/4-3\eta}$.

Subcase 3.1 ($p_1 p_4 < x^{1/6+2\eta}$). Suppose we can show $p_6 < x^{1/12-5\eta}$ (when $r \geq 6$). Then since $x^{1-3\delta}/D > D^4/x^{2-6\delta}$, either $Nx^{-\delta}$ or $x^{1-2\delta}/DN$ is greater than $D^2/x^{1-3\delta}$. Thus Lemma 3.6 will complete the proof, with $(d_1, d_2, d_3) = (p_1 p_4, p_2 p_3, 1)$ or $(1, p_2 p_3, p_1 p_4)$.

If $p_1 p_3 > x^{1/4-3\eta}$ then in this subcase

$$x^{1/12+\eta} > (p_2 p_3)^{1/2} \geq p_3 = p_4 \frac{p_1 p_3}{p_1 p_4} > p_4 x^{1/12-5\eta},$$

so $p_4 < x^{6\eta}$. Hence $p_6 \leq p_4 < x^{1/12-5\eta}$ since $\eta < \frac{1}{108}$, which completes the proof.

Else $p_1 p_3 < x^{1/4-3\eta}$. By Lemma 3.5, it suffices $p_1 p_3 < x^{1/6+2\eta}$. Then we see $p_3 > x^{1/12-5\eta}$ implies $p_1 < x^{1/12+7\eta}$. If further $p_1 p_2 > x^{1/4-3\eta}$, then similarly

$$x^{1/12+7\eta} > p_1 \geq p_2 = p_3 \frac{p_1 p_2}{p_1 p_3} > p_3 x^{1/12-5\eta},$$

so $p_3 < x^{12\eta}$. Hence $p_6 \leq p_3 < x^{1/12-5\eta}$ since $\eta < \frac{1}{204}$, which completes the proof.

Else $p_1 p_2 < x^{1/4-3\eta}$. By Lemma 3.5, we may assume $p_1 p_2 < x^{1/6+2\eta}$.

Similarly, suppose $p_2 p_3 p_4 < x^{1/4-3\eta}$. By Lemma 3.5 we may assume $p_2 p_3 p_4 < x^{1/6+2\eta}$, and so

$$p_6 \leq (p_2 p_3 p_4)^{1/3} \leq x^{(2/3)/12+(2/3)\eta} \leq x^{1/12-5\eta},$$

using $\eta < \frac{1}{204}$, which completes the proof.

Thus we may assume $p_2 p_3 p_4 > x^{1/4-3\eta}$. But unless $p_6 < x^{1/12-5\eta}$, this subcase will contradict the definition of $\mathcal{P}_{6,1}$ in (3-6), hence completing the proof.

Subcase 3.2 ($p_1 p_4 > x^{1/4-3\eta}$). If $d_2 = p_2 p_3 p_4 < x^{1/6+2\eta}$, then Lemma 3.6 completes the proof with $(d_1, d_3) = (p_1, 1)$ or $(1, p_1)$, since

$$p_6 \leq (p_2 p_3 p_4)^{1/3} \leq x^{(1/3)(1/6+2\eta)} \leq x^{1/12-5\eta}$$

for $\eta < \frac{1}{204}$. And if $p_2 p_3 p_4 \in [x^{1/6+2\eta}, x^{1/4-3\eta}]$ the proof follows by Lemma 3.5.

Else $p_2 p_3 p_4 > x^{1/4-3\eta}$. Note $p_4 < x^{1/12+\eta}$ and $p_1 = p_1 p_4 / p_4 > x^{1/6-4\eta}$ and $p_2 p_3 p_4 < x^{1/4+3\eta}$. Also note we may factor $p_1 p_4 = d_1 d_3$ for $d_1 \leq N$, $d_3 \leq x^{1-2\delta}/DN$ (Indeed this follows if N or $x^{1-2\delta}/DN$ exceeds $x^{1/4+3\eta} \geq p_1 p_4$. Else $N, x^{1-2\delta}/DN \in [x^{1/6-4\eta-2\delta}, x^{1/4+3\eta}]$, which also works similarly as with (3-13), since $p_4 < x^{1/12+\eta} < x^{1/6-4\eta-2\delta}$ and $p_1 < x^{1/6+2\eta} < x^{1/2(5/12-\eta-\delta)}$ by $\eta < \frac{1}{60}$.)

If further $p_6 > x^{1/12-5\eta}$, then this subcase contradicts the definition of $\mathcal{P}_{6,2}$ in (3-6). Hence we have $p_6 \leq x^{1/12-5\eta}$, and so by the above paragraph Lemma 3.6 completes the proof with $d_2 = p_2 p_3$.

Combining all cases completes the proof of Proposition 3.3. □

3.1. Refined factorization of $\mathcal{D}^{\text{well}}$. Proposition 3.3 (programmably) factorizes each $d \in \mathcal{D}^* \subset \mathcal{D}^+(x^{7/12+\eta})$, and forms the key step to prove the weights λ^* are programmably factorable. With applications in mind to twin primes, we shall similarly (programmably) factorize certain subsets of the well-factorable support $\mathcal{D}^{\text{well}}$, as in (3-2).

In the following result, we factorize $d \in \mathcal{D}^{\text{well}}(D)$ for variable level $D \in (x^{4/7}, x^{3/5})$, depending on the anatomy of d . As $\mathcal{D}^{\pm} \subset \mathcal{D}^{\text{well}}$, this has implications to both upper and lower bounds for the standard linear sieve.

Proposition 3.7. *Let $\mathcal{D}^{\text{well}}(D)$ as in (3-2) for $D = x^{7/12+\eta-50\delta}$ and $-\frac{1}{84} < \eta < \frac{1}{60} - 30\delta$. Let $x^{1/4-3\eta} \geq p_1 \geq \dots \geq p_r$ be primes for which $d = p_1 \cdots p_r \in \mathcal{D}^{\text{well}}(D)$. Then d has factorization $d = abc$ satisfying (3-7) if $p_3 \leq x^{1/12-5\eta}$, or if*

$$d_2 \in [x^{1/6+2\eta}, x^{1/4-3\eta}] \quad \text{with } d_2 \mid p_1 p_2 p_3, \quad d_2 \neq p_3.$$

Proof. For $i = 1, 2, 3$, suppose $d_2 = p_1 \cdots p_i$ lies $[x^{1/6+2\eta}, x^{1/4-3\eta}]$, and let $A = Nx^{-\delta}$, $C = x^\delta D / Nd_2$. Since $p_1 \cdots p_{j-1} p_j^2 \leq D$ for all $i < j \leq r$, we get $p_j^2 \leq AC / (p_{i+1} \cdots p_{j-1})$ for $i < j \leq r$. As such, we may factor $p_{i+1} \cdots p_r = d_1 d_3$ for $d_1 \leq A$, $d_3 \leq C$. Hence by Lemma 3.4 $p_1 \cdots p_r = d_1 d_2 d_3$ satisfies (3-7).

Else, by assumption $p_1 < x^{1/(4)-3\eta}$ so we may assume further $p_1 < x^{1/(6)+2\eta}$. In particular this gives $p_1^2 \leq D/d_2$. For the remaining $d_2 \mid p_1 p_2 p_3$:

- If $d_2 = p_2 p_3$ then $p_1^2 \leq D/d_2 = AC$ implies $p_1 \leq A$ or $p_1 \leq C$.
- If $d_2 = p_1 p_3$ then $p_2^2 \leq D/d_2 = AC$ implies $p_2 \leq A$ or $p_2 \leq C$.
- If $d_2 = p_2$ then $p_3^2 \leq D/p_1 d_2$ implies a factorization $p_1 p_3 = d_1 d_3$ for $d_1 \leq A$, $d_3 \leq C$.

For each d_2 above, we factored $p_1 p_2 p_3 = d_1 d_2 d_3$ for $d_1 \leq A$, $d_3 \leq C$. Since $p_1 \cdots p_{j-1} p_j^2 \leq D$ for all $j \leq r$, by induction we may factor $p_1 \cdots p_r = d_1 d_2 d_3$ for $d_1 \leq A$, $d_3 \leq C$. By Lemma 3.4 $p_1 \cdots p_r = d_1 d_2 d_3$ satisfies (3-7).

Finally, suppose $p_3 \leq x^{1/12-5\eta}$ is less than the width of the interval $[x^{1/6+2\eta}, x^{1/4-3\eta}]$. Since $p_1 < x^{1/6+2\eta}$, we have $p_1 p_3 < x^{1/4-3\eta}$ so by the above argument we may assume $d_2 := p_1 p_3 < x^{1/6+2\eta}$. Then $p_2^3 \leq p_1 p_2^2 \leq D$ implies

$$p_2^2 \leq x^{(2/3)(7/12+\eta)} < x^{5/12-\eta+47\delta} = \frac{x^{1-3\delta}}{D},$$

since $\eta < \frac{1}{60} - 30\delta$. Thus $p_2 \leq Nx^{-\delta}$ or $p_2 \leq x^{1-2\delta}/DN$, so there is a factorization $p_1 p_2 p_3 = d_1 d_2 d_3$ satisfying (3-9). Hence the same greedy argument as in Lemma 3.6 completes the proof, with p_3 playing the role of p_6 . \square

Taking the maximum valid η as above, we may reexpress the above factorization of level x^θ , $\theta = \frac{7}{12} + \eta$, as follows. Note the maximum θ for which $t \in [\frac{1}{6} + 2\eta, \frac{1}{4} - 3\eta]$ is given by

$$\theta(t) = \begin{cases} \frac{2-t}{3} & \text{if } t > \frac{1}{5}, \\ \frac{1+t}{2} & \text{if } t \leq \frac{1}{5}. \end{cases} \quad (3-14)$$

Similarly the maximum $\theta = \frac{7}{12} + \eta$ for which $t \leq \frac{1}{12} - 5\eta$ is $\frac{1}{5}(3-t)$.

Corollary 3.8. *Let $p_1 \geq \dots \geq p_r$ be primes and write $p_i = x^{t_i}$. If $d = p_1 \cdots p_r \in \mathcal{D}^{\text{well}}(x^{\theta-50\delta})$, then there is a factorization $d = d_1 d_2 d_3$ satisfying (3-7) provided*

$$\theta \leq \theta(t_1),$$

for $\theta(t)$ as in (3-14). Moreover if $t_1 \leq \frac{1}{5}$, then it suffices that

$$\theta \leq \theta(t_1, t_2, t_3) := \max \left\{ \frac{3-t_3}{5}, \theta(t_1), \theta(t_2), \theta(t_1+t_2+t_3), \theta(t_1+t_2), \theta(t_1+t_3), \theta(t_2+t_3) \right\}. \quad (3-15)$$

4. Modification of the linear sieve

In this section we shall bound the modified linear sieve, analogous to the bounds for the linear sieve (sometimes called the Jurkat–Richert theorem). This bound will form the basis for our final result in the next section, in which we modify the construction of Iwaniec’s weights.

Proposition 4.1. *Let $\varepsilon > 0$ be sufficiently small. For $\eta \leq \frac{1}{204}$, the modified weights λ^* as in (3-4) of level $D = x^{7/12+\eta-\varepsilon}$ satisfy*

$$S(\mathcal{A}, z) \leq |\mathcal{A}| V(z) \left(F^* \left(\frac{\log D}{\log z} \right) + o(1) \right) + \sum_{d|P(z)} \lambda^*(d) r_{\mathcal{A}}(d),$$

where $F^* = F_{\eta}^*$ is a function satisfying $F^*(s) = F(s) + O(\eta^5)$ for F as in (2-5).

Remark 4.2. It suffices for our purposes to obtain qualitative error $o(1)$ in the factor accompanying F^* . Though as with the Jurkat–Richert theorem, with greater care one should obtain a quantitative refinement, e.g., $O((\log D)^{-1/6})$; see (12.4)–(12.8) in [Friedlander and Iwaniec 2010].

We now adapt the proof. Let $D = x^{7/12+\eta}$ and $D_0 = x^{7/12}$. For $n \geq 1$, primes $p_1 \geq \dots \geq p_n$, if $p_1 \cdots p_n \notin \mathcal{D}^+(D)$ then there exists a minimal index $l \leq n$ such that $p_1 \cdots p_l \notin \mathcal{D}^+(D)$. By definition such minimal l is odd. (Explicitly, this occurs when $p_1 \cdots p_{l-1} p_l^3 > D$ but $p_1 \cdots p_{m-1} p_m^3 \leq D$ for all odd $m < l$.) Similarly, if $p_1 \cdots p_n \notin \mathcal{D}^*$ there exists a minimal index $l \leq n$ such that $p_1 \cdots p_l \notin \mathcal{D}^*$, which is also odd.

Indeed, to show this let $l \leq n$ be the minimal index such that $p_1 \cdots p_l \notin \mathcal{D}^*$. If $(p_1, \dots, p_j) \notin \mathcal{P}_j$ for all $j \leq l$, $j \in \{4, 6\}$, then clearly $l > j$ must be odd, as with $\mathcal{D}^+(D)$. On the other hand, if $(p_1, \dots, p_j) \in \mathcal{P}_j$ for some $j \leq l$, $j \in \{4, 6\}$, a priori one might expect l could be even. However, the key point in this case is that $p_1 \cdots p_j \in \mathcal{D}^+(D_0) \subset \mathcal{D}^*$ (since $p_1 \cdots p_j \approx D_0^{6/7}$ by definition of \mathcal{P}_j). Thus $l > j$ is the minimal index such that $p_1 \cdots p_l \notin \mathcal{D}^+(D_0)$, and hence must be odd as claimed.

Using this minimal index, we show the following lemma.

Lemma 4.3. *Let h be a multiplicative function with $0 \leq h(p) \leq 1$ for all primes p . Then we have*

$$\prod_{p|n} (1 - h(p)) \leq \sum_{d|n} \lambda^*(d) h(d).$$

Proof. Note if $h = 1$ identically, we interpret the product as $\mathbf{1}_{n=1}$. Now by definition,

$$\sum_{d|n} \lambda^*(d)h(d) - \prod_{p|n} (1 - h(p)) = \sum_{\substack{d|n \\ d \in \mathcal{D}^*}} \mu(d)h(d) - \sum_{d|n} \mu(d)h(d) = - \sum_{\substack{d|n \\ d \notin \mathcal{D}^*}} \mu(d)h(d).$$

Then splitting up $d \notin \mathcal{D}^*$ by its minimal index,

$$- \sum_{\substack{d|n \\ d \notin \mathcal{D}^*}} \mu(d)h(d) = \sum_{\substack{\text{odd } l \\ p_1 \cdots p_l \notin \mathcal{D}^*}} \sum_{\substack{p_1 < \cdots < p_l < z \\ p_1 \cdots p_{l-1} \in \mathcal{D}^*}} h(p_1 \cdots p_l) \sum_{\substack{p_1 \cdots p_l b | n \\ b | P(p_l)}} \mu(b)h(b) \geq 0,$$

since $h \geq 0$ and the inner sum over b factors as $\prod_{p|(P(p_l), n)} (1 - h(p)) \geq 0$, since $h(p) \leq 1$. \square

By Lemma 4.3 with $h(d) = 1$, we have

$$\mathbf{1}_{n=1} \leq \sum_{d|n} \lambda^*(d), \quad (4-1)$$

in which case we obtain

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, z) &= \sum_{n \in \mathcal{A}} \mathbf{1}_{(n, P(z))=1} \leq \sum_{n \in \mathcal{A}} \sum_{d|n} \lambda^*(d) = \sum_{d|P(z)} \lambda^*(d) |\mathcal{A}_d| \\ &= X \sum_{d|P(z)} \lambda^*(d) g(d) + \sum_{d|P(z)} \lambda^*(d) r_{\mathcal{A}}(d) =: XV^*(D, z) + R_{\mathcal{A}}^*(D, z). \end{aligned} \quad (4-2)$$

Following Lemma 4.3 with $h = g$, we have the identity

$$V^*(D, z) := \sum_{\substack{d|P(z) \\ d \in \mathcal{D}^*}} \mu(d)g(d) = V(z) + \sum_{\substack{\text{odd } n \\ p_n < \cdots < p_1 < z \\ p_1 \cdots p_{n-1} \in \mathcal{D}^* \\ p_1 \cdots p_n \notin \mathcal{D}^*}} g(p_1 \cdots p_n) V(p_n), \quad (4-3)$$

and similarly

$$V^+(D, z) = V(z) + \sum_{\substack{\text{odd } n \\ p_n < \cdots < p_1 < z \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D) \\ p_1 \cdots p_n \notin \mathcal{D}^+(D)}} g(p_1 \cdots p_n) V(p_n) =: V(z) + \sum_{\text{odd } n} V_n(D, z). \quad (4-4)$$

Then the difference of V^* and V^+ is

$$V^*(D, z) - V^+(D, z) = \sum_{\text{odd } n} \sum_{p_n < \cdots < p_1 < z} g(p_1 \cdots p_n) V(p_n) \mathbf{\Delta}, \quad (4-5)$$

where $\mathbf{\Delta}$ is the difference of indicator functions,

$$\mathbf{\Delta} := \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^* \\ p_1 \cdots p_{n-1} \in \mathcal{D}^*}} - \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D)}} = \mathbf{1}_{p_1 \cdots p_n \in \mathcal{D}^+(D) \setminus \mathcal{D}^*} - \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D) \setminus \mathcal{D}^*}},$$

recalling $\mathcal{D}^* \subset \mathcal{D}^+(D)$. Note if a point is $(p_1, \dots, p_6) \in \mathcal{P}_6$ then its projection is $(p_1, \dots, p_4) \notin \mathcal{P}_4$. So by definitions of \mathcal{D}^* , $\mathcal{D}^+(D)$ from (3-5), (3-1), for odd n we have the identities

$$\mathbf{1}_{\substack{p_1 \cdots p_n \in \mathcal{D}^+(D) \setminus \mathcal{D}^* \\ p_1 \cdots p_{n-1} \in \mathcal{D}^*}} = \sum_{\substack{j \in \{4,6\} \\ j < n}} \mathbf{1}_{(p_1, \dots, p_j) \in \mathcal{P}_j} \cdot \mathbf{1}_{\substack{p_1 \cdots p_n \in \mathcal{D}^+(D) \setminus \mathcal{D}^+(D_0) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D_0)}}, \quad (4-6)$$

$$\mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D) \setminus \mathcal{D}^*}} = \sum_{\substack{j \in \{4,6\} \\ j < n}} \mathbf{1}_{(p_1, \dots, p_j) \in \mathcal{P}_j} \cdot \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D) \setminus \mathcal{D}^+(D_0)}}. \quad (4-7)$$

We may plug (4-6) and (4-7) into Δ . In addition, we strategically add and subtract the indicator function of $\{p_1 \cdots p_n \notin \mathcal{D}^+(D), p_1 \cdots p_{n-1} \in \mathcal{D}^+(D_0)\}$, which together give

$$\begin{aligned} \Delta &= \sum_{\substack{j \in \{4,6\} \\ j < n}} \mathbf{1}_{(p_1, \dots, p_j) \in \mathcal{P}_j} \cdot \left(\mathbf{1}_{\substack{p_1 \cdots p_n \in \mathcal{D}^+(D) \setminus \mathcal{D}^+(D_0) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D_0)}} - \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D) \setminus \mathcal{D}^+(D_0)}} \right) \\ &= \sum_{\substack{j \in \{4,6\} \\ j < n}} \mathbf{1}_{(p_1, \dots, p_j) \in \mathcal{P}_j} \cdot \left(\mathbf{1}_{\substack{p_1 \cdots p_n \in \mathcal{D}^+(D) \setminus \mathcal{D}^+(D_0) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D_0)}} + \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D_0)}} \right. \\ &\quad \left. - \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D) \setminus \mathcal{D}^+(D_0)}} - \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D_0)}} \right) \\ &= \sum_{\substack{j \in \{4,6\} \\ j < n}} \mathbf{1}_{(p_1, \dots, p_j) \in \mathcal{P}_j} \cdot \left(\mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D_0) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D_0)}} - \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D)}} \right). \end{aligned}$$

Thus plugging Δ back into (4-5) gives

$$\begin{aligned} V^*(D, z) - V^+(D, z) &= \sum_{j \in \{4,6\}} \sum_{\substack{p_j < \cdots < p_1 < z \\ (p_1, \dots, p_j) \in \mathcal{P}_j}} \sum_{\substack{\text{odd } n > j \\ p_n < \cdots < p_{j+1} < p_j}} g(p_1 \cdots p_n) V(p_n) \left(\mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D_0) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D_0)}} - \mathbf{1}_{\substack{p_1 \cdots p_n \notin \mathcal{D}^+(D) \\ p_1 \cdots p_{n-1} \in \mathcal{D}^+(D)}} \right). \end{aligned}$$

Recalling the definition of $V_n(D, z)$ in (4-4), since g is multiplicative we have

$$\begin{aligned} V^*(D, z) - V^+(D, z) &= \sum_{j \in \{4,6\}} \sum_{\substack{p_j < \cdots < p_1 < z \\ (p_1, \dots, p_j) \in \mathcal{P}_j}} g(p_1 \cdots p_j) \sum_{\text{odd } n > j} \left(V_{n-j} \left(\frac{D_0}{p_1 \cdots p_j}, p_j \right) - V_{n-j} \left(\frac{D}{p_1 \cdots p_j}, p_j \right) \right) \\ &= \sum_{j \in \{4,6\}} \sum_{\substack{p_j < \cdots < p_1 < z \\ (p_1, \dots, p_j) \in \mathcal{P}_j}} g(p_1 \cdots p_j) \left(V^+ \left(\frac{D_0}{p_1 \cdots p_j}, p_j \right) - V^+ \left(\frac{D}{p_1 \cdots p_j}, p_j \right) \right). \quad (4-8) \end{aligned}$$

as $V^+(D, z) - V^+(D', z) = \sum_{\text{odd } n} [V_n(D, z) - V_n(D', z)]$.

Now assuming the two-sided condition (2-3) for g , the proof of [Friedlander and Iwaniec 2010, Theorem 11.12] (see (12.4)–(12.8)) gives asymptotic equality,

$$V^+(D, z) = V(z) \left\{ F \left(\frac{\log D}{\log z} \right) + O((\log D)^{-1/6}) \right\} \quad (z \leq D), \quad (4-9)$$

so that (4-8) becomes

$$\begin{aligned}
 V^*(D, z) &= V(z) \left\{ F\left(\frac{\log D}{\log z}\right) + O((\log D)^{-1/6}) \right\} \\
 &\quad + \sum_{j \in \{4,6\}} \sum_{\substack{p_j < \dots < p_1 < z \\ (p_1, \dots, p_j) \in \mathcal{P}_j}} g(p_1 \cdots p_j) V(p_j) \\
 &\quad \times \left\{ F\left(\frac{\log D_0/p_1 \cdots p_j}{\log p_j}\right) - F\left(\frac{\log D/p_1 \cdots p_j}{\log p_j}\right) + O\left(\log\left(\frac{D}{p_1 \cdots p_j}\right)^{-1/6}\right) \right\}. \quad (4-10)
 \end{aligned}$$

By partial summation and the prime number theorem, for each j we have,

$$\begin{aligned}
 &\sum_{\substack{p_j < \dots < p_1 < z \\ (p_1, \dots, p_j) \in \mathcal{P}_j}} g(p_1 \cdots p_j) V(p_j) F\left(\frac{\log D_0/p_1 \cdots p_j}{\log p_j}\right) \\
 &= \left(\frac{7}{12} + \eta\right) \int_{(x_1, \dots, x_j) \in P_j} \frac{dx_1 \cdots dx_j}{x_1 \cdots x_{j-1} x_j^2} F\left(\frac{7/12 - x_1 - \cdots - x_j}{x_j}\right) + O((\log D)^{-1/6}).
 \end{aligned}$$

Here P_j is the polytope in Euclidean space \mathbb{R}^j corresponding to \mathcal{P}_j , as below.

Hence from (4-10), we obtain

$$V^*(D, z) = V(z) \left\{ F^*\left(\frac{\log D}{\log z}\right) + O((\log D)^{-1/6}) \right\} \quad (z \leq D), \quad (4-11)$$

where the function F^* satisfies

$$\begin{aligned}
 &sF^*(s) - sF(s) \\
 &= \left(\frac{7}{12} + \eta\right) \cdot \sum_{j \in \{4,6\}} \int_{(x_1, \dots, x_j) \in P_j} \frac{dx_1 \cdots dx_j}{x_1 \cdots x_{j-1} x_j^2} \left[F\left(\frac{7/12 - x_1 - \cdots - x_j}{x_j}\right) - F\left(\frac{7/12 + \eta - x_1 - \cdots - x_j}{x_j}\right) \right]. \quad (4-12)
 \end{aligned}$$

Namely, $P_4 \subset \mathbb{R}^4$ is given by

$$P_4 = \left\{ (x_1, \dots, x_4) \in D^+\left(\frac{7}{12} + \eta\right) : x_1 < \frac{1}{6} + 2\eta \text{ and } x_2 + x_4 > \frac{1}{4} - 3\eta \right\},$$

and $P_6 = P_{6,1} \cup P_{6,2} \subset \mathbb{R}^6$ is given by

$$\begin{aligned}
 P_{6,1} &= \left\{ (x_1, \dots, x_6) \in D^+\left(\frac{7}{12} + \eta\right) : x_1 + x_2 < \frac{1}{6} + 2\eta \text{ and } x_6 > \frac{1}{12} - 5\eta \text{ and } x_2 + x_3 + x_4 > \frac{1}{4} - 3\eta \right\}, \\
 P_{6,2} &= \left\{ (x_1, \dots, x_6) \in D^+\left(\frac{7}{12} + \eta\right) : x_1, x_2 + x_3 < \frac{1}{6} + 2\eta \right. \\
 &\quad \left. \text{and } x_6 > \frac{1}{12} - 5\eta \text{ and } x_1 + x_4, x_2 + x_3 + x_4 > \frac{1}{4} - 3\eta \right\}. \quad (4-13)
 \end{aligned}$$

Similarly, D^+ is the set in Euclidean space corresponding to \mathcal{D}^+ , namely,

$$D^+(\tau) = \{(x_1, \dots, x_r) : x_1 > \dots > x_r > 0 \text{ and } x_1 + \dots + x_{l-1} + 3x_l < \tau \text{ for each odd } 1 \leq l \leq r\}.$$

Hence Proposition 4.1 follows.

4.1. Sieve function computation. We now compute F^* in terms of F .

Proposition 4.4. *Let $\eta = \frac{1}{204}$. Then for $1 \leq s \leq 3$, we have*

$$F^*(s) \leq 1.000081F(s). \quad (4-14)$$

Proof. From (4-12) we have

$$sF^*(s) = sF(s) + \left(\frac{7}{12} + \eta\right) \cdot 2e^\gamma \eta (J_4 + J_6), \quad (4-15)$$

for integrals J_j , $j \in \{4, 6\}$,

$$\begin{aligned} J_j &:= \frac{1}{2e^\gamma \eta} \int_{(x_1, \dots, x_j) \in P_j} \frac{dx_1 \cdots dx_j}{x_1 \cdots x_{j-1} x_j^2} \left[F\left(\frac{7/12 - x_1 - \cdots - x_j}{x_j}\right) - F\left(\frac{7/12 + \eta - x_1 - \cdots - x_j}{x_j}\right) \right] \\ &= \int_{(x_1, \dots, x_j) \in P_j} \frac{dx_1 \cdots dx_j}{x_1 \cdots x_j} \left[\left(\frac{7}{12} - x_1 - \cdots - x_j\right) \left(\frac{7}{12} + \eta - x_1 - \cdots - x_j\right) \right]^{-1}, \end{aligned}$$

since $sF(s) = 2e^\gamma$ for $s \in [1, 3]$. In particular $|P_j| = O(\eta^j)$ implies $J_j = O(\eta^j)$, and so from (4-15) we obtain $F^*(s) = F(s) + O(\eta^5)$.

For $\eta = \frac{1}{204}$, we use Mathematica⁴ to compute that

$$J_4 \leq 0.016896. \quad (4-16)$$

Next we bound J_6 . For $(x_1, \dots, x_6) \in P_6$ we have $x_4 < \frac{1}{2}(x_2 + x_3) < \frac{1}{12} + \eta$ and $\frac{7}{12} + \eta - x_1 - \cdots - x_6 > x_5$ so

$$J_6 \leq \int_{\overline{P_6}} \frac{dx_1 dx_2 dx_3}{x_1 x_2 x_3} \int_{1/12 - 5\eta < x_6 < x_5 < x_4 < 1/12 + \eta} \frac{dx_4 dx_5 dx_6}{x_4 x_5^2 x_6 (x_5 - \eta)},$$

where $\overline{P_6} = \overline{P_{6,1}} \cup \overline{P_{6,2}}$ is the (3-dimensional) projection of P_6 , given explicitly by

$$\begin{aligned} \overline{P_{6,1}} &= \{(x_1, x_2, x_3) \in D^+(\tfrac{7}{12} + \eta) : x_1 + x_2 < \tfrac{1}{6} + 2\eta \text{ and } x_3 > \tfrac{1}{12} - 5\eta \text{ and } x_2 + 2x_3 > \tfrac{1}{4} - 3\eta\}, \\ \overline{P_{6,2}} &= \{(x_1, x_2, x_3) \in D^+(\tfrac{7}{12} + \eta) : x_1, x_2 + x_3 < \tfrac{1}{6} + 2\eta \text{ and } x_3 > \tfrac{1}{12} - 5\eta \text{ and } x_1 + x_3, x_2 + 2x_3 > \tfrac{1}{4} - 3\eta\}. \end{aligned}$$

For $\eta = \frac{1}{204}$, we compute $J_6 \leq (J_{6,1} + J_{6,2})J_{6,0}$ where

$$\begin{aligned} J_{6,0} &= \int_{1/12 - 5\eta < x_6 < x_5 < x_4 < 1/12 + \eta} \frac{dx_4 dx_5 dx_6}{x_4 x_5^2 x_6 (x_5 - \eta)} \leq 2.33838, \\ J_{6,1} &= \int_{\overline{P_{6,1}}} \frac{dx_1 dx_2 dx_3}{x_1 x_2 x_3} \leq 0.000806853, \\ J_{6,2} &= \int_{\overline{P_{6,2}}} \frac{dx_1 dx_2 dx_3}{x_1 x_2 x_3} \leq 0.00397946. \end{aligned}$$

Hence combining with (4-16), for $s \in [1, 3]$ we conclude

$$\frac{F^*(s)}{F(s)} \leq 1 + \left(\frac{7}{12} + \eta\right) \cdot \eta (J_4 + (J_{6,1} + J_{6,2})J_{6,0}) \leq 1.000081. \quad \square$$

⁴The Mathematica package and code are available at arxiv.org/abs/2109.02851.

5. Factorable remainder, after Iwaniec

In Theorem 2.10, Iwaniec constructed a well-factorable variant $\tilde{\lambda}^\pm$ of the weights λ^\pm from the (Jurkat–Richert) linear sieve. In this section, we prove Theorem 2.12 for the programmably factorable variant $\tilde{\lambda}^*$ by adapting Iwaniec’s construction, similarly building on the Jurkat–Richert type Proposition 4.1 that we obtained in the previous section. We shall also prove a technical variation on this result, with a variable level depending on anatomy of the moduli.

To set up the construction, we first adapt [Friedlander and Iwaniec 2010, Proposition 12.18]. Denote $P(z, u) = P(z)/P(u) = \prod_{u < p \leq z} p$.

Proposition 5.1. *Let $\eta > 0$, and $D = x^{(7/12+\eta)/(1+\varepsilon+\tau)}$ for $\varepsilon > 0$ sufficiently small. Let \mathbf{D}_r^* be defined by (5-8). Let $\lambda^{(r)}$ be the standard (upper and lower, for r odd and even, resp.) weights for the linear sieve of level D^ε . Then for $u = D^{\varepsilon^2}$, $\tau = \varepsilon^9$,*

$$S(\mathcal{A}, z) \leq |\mathcal{A}|V(z)\{F^*(s) + O(\varepsilon^5)\} + \sum_{0 \leq r \leq \varepsilon^{-2}} \sum_{(D_1, \dots, D_r) \in \mathbf{D}_r^*} \frac{(-1)^r}{\gamma(D_1, \dots, D_r)} \sum_{\substack{p_1 \cdots p_r | P(z, u) \\ D_j < p_j \leq D_j^{1+\tau}}} \sum_{\substack{b | P(u) \\ b \leq D^\varepsilon}} \lambda^{(r)}(b)r_{\mathcal{A}}(bp_1 \cdots p_r). \quad (5-1)$$

Proof. First we write

$$S(\mathcal{A}, z) \leq S^*(\mathcal{A}, z) - \sum_{\text{odd } n \leq N} S_n(\mathcal{A}, z) \quad (5-2)$$

for any $N \geq 1$, where

$$S^*(\mathcal{A}, z) := \sum_{0 \leq r \leq \varepsilon^{-2}} (-1)^r \sum_{\substack{u \leq p_r < \dots < p_1 < z \\ p_1 \cdots p_r \in \mathbf{D}^*}} |\mathcal{A}_{p_1 \cdots p_r}|, \quad S_n(\mathcal{A}, z) := \sum_{\substack{p_n < \dots < p_1 < z \\ p_1 \cdots p_{n-1} \in \mathbf{D}^* \\ p_1 \cdots p_n \notin \mathbf{D}^*}} S(\mathcal{A}_{p_1 \cdots p_n}, p_n).$$

We apply the inequality (5-2), not for $\mathcal{A} = (a_n)$ itself but rather for the subsequence $\tilde{\mathcal{A}} = (a_n \mathbf{1}_{(n, P(u))=1})$. Here we take $u = D^{\varepsilon^2}$, and then return to \mathcal{A} by means of the fundamental lemma.

Let $z = D^{1/s}$ with $2 \leq s \leq \varepsilon^{-1}$. Since $z > u$, the only change to the above bound (5-2) when passing to $\tilde{\mathcal{A}}$ is the term $S^*(\tilde{\mathcal{A}}, z)$, provided that N is not too large in terms of ε . Specifically, we require the lower bound for p_n (by induction, the linear sieve conditions imply $p_1 \cdots p_m < D^{1-2^{-m}}$)

$$p_n \geq (D/p_1 \cdots p_{n-1})^{1/3} \geq D^{2^{1-N}/3} \quad (5-3)$$

to be larger than $u = D^{\varepsilon^2}$, which certainly holds provided

$$N \leq \frac{1}{2} \log \frac{1}{\varepsilon}. \quad (5-4)$$

Now it remains to evaluate $S^*(\tilde{\mathcal{A}}, z)$,

$$S^*(\tilde{\mathcal{A}}, z) = \sum_{0 \leq r \leq \varepsilon^{-2}} (-1)^r \sum_{\substack{u \leq p_r < \dots < p_1 < z \\ p_1 \cdots p_r \in \mathbf{D}^*}} |\tilde{\mathcal{A}}_{p_1 \cdots p_r}|. \quad (5-5)$$

For each r , we break the range in the inner sum into boxes. Namely, we let D_1, \dots, D_r run over numbers of form

$$D^{\varepsilon^2(1+\tau)^j}, \quad j = 0, 1, 2, \dots, \quad (5-6)$$

with $\tau = \varepsilon^9$. We denote by $\mathbf{D}_r^+ = \mathbf{D}_r^+(D)$ the set of r -tuples (D_1, \dots, D_r) with $D_r \leq \dots \leq D_1 \leq \sqrt{D}$ such that

$$\mathbf{D}_r^+ = \begin{cases} \{(D_1, \dots, D_r) : D_1 \cdots D_{m-1} D_m^3 < D \text{ for all odd } m \leq r\} & \text{if } r \text{ even,} \\ \{(D_1, \dots, D_r) : D_1 \cdots D_{m-1} D_m^3 < D^{1/(1+\tau)} \text{ for all odd } m \leq r\} & \text{if } r \text{ odd.} \end{cases}$$

We note, for $\varepsilon > 0$ sufficiently small, the cardinalities of the \mathbf{D}_r^+ are bounded by

$$\sum_{0 \leq r \leq \varepsilon^{-2}} |\mathbf{D}_r^+| \leq \exp(\varepsilon^{-3}). \quad (5-7)$$

Hereafter let $D = x^{(7/12+\eta)/(1+\tau+\varepsilon)}$, and define

$$\mathbf{D}_r^* = \{(D_1, \dots, D_r) \in \mathbf{D}_r^+(D) : (D_1, \dots, D_i) \notin \mathbf{P}_{i,r} \text{ for } i \leq r, i \in \{4, 6\}\}. \quad (5-8)$$

where $\mathbf{P}_{4,r}, \mathbf{P}_{6,r}$ are (τ -enlarged, for even r) analogues of the polytopes $\mathcal{P}_4, \mathcal{P}_6$ in (3-6), e.g.,

$$\mathbf{P}_{4,r} = \begin{cases} \{(D_1, \dots, D_4) : D_1^{1+\tau} < x^{1/6+2\eta} \text{ and } (D_2 D_4)^{1/(1+\tau)} > x^{1/4-3\eta}\} & \text{if } r \text{ even,} \\ \{(D_1, \dots, D_4) : D_1 < x^{1/6+2\eta} \text{ and } D_2 D_4 > x^{1/4-3\eta}\} & \text{if } r \text{ odd.} \end{cases}$$

Observe each integer $p_1 \cdots p_r$ has a unique vector (D_1, \dots, D_r) such that $p_i \in (D_i, D_i^{1+\tau}]$ for all $i \leq r$, inducing a map $v : \mathbb{N} \rightarrow \bigcup_r \mathbf{D}_r^+$. As a convention $v(1) = ()$ is the empty vector. By construction, for even r if $p_1 \cdots p_r \notin \mathcal{P}_4$ then $v(p_1 \cdots p_r) \notin \mathbf{P}_{4,r}$, and if $(D_1, \dots, D_4) \notin \mathbf{P}_{4,r}$ then $v^{-1}(D_1, \dots, D_4) \cap \mathcal{P}_4 = \emptyset$ for odd r . Continuing this argument, we deduce

$$\begin{aligned} p_1 \cdots p_r \in \mathcal{D}^* &\implies v(p_1 \cdots p_r) \in \mathbf{D}_r^* && \text{if } r \text{ even,} \\ (D_1, \dots, D_r) \in \mathbf{D}_r^* &\implies v^{-1}(D_1, \dots, D_r) \subset \mathcal{D}^* && \text{if } r \text{ odd.} \end{aligned} \quad (5-9)$$

Without loss, we may restrict \mathbf{D}_r^* to vectors with nonempty preimage in \mathcal{D}^* . Hence by construction, (5-5) becomes⁵

$$S^*(\tilde{\mathcal{A}}, z) \leq \sum_{0 \leq r \leq \varepsilon^{-2}} \sum_{(D_1, \dots, D_r) \in \mathbf{D}_r^*} \frac{(-1)^r}{\gamma(D_1, \dots, D_r)} \sum_{\substack{p_1 \cdots p_r \mid P(z) \\ D_j < p_j \leq D_j^{1+\tau}}} |\tilde{\mathcal{A}}_{p_1 \cdots p_r}|, \quad (5-10)$$

where $\gamma(D_1, \dots, D_r) = k_1! \cdots k_\ell!$ for the corresponding multiplicities $k_i \geq 1$ (i.e., we have $r = k_1 + \cdots + k_\ell$ and $D_1 = \cdots = D_{k_1} < D_{k_1+1} = \cdots = D_{k_2} < \cdots = D_r$). Note the term $r = 0$ corresponds to $|\tilde{\mathcal{A}}|$ with $p_1 = \cdots = p_r = 1$.

⁵Indeed, we have reverse engineered the definition of \mathbf{D}_r^* just so that (5-10) holds.

Now by the fundamental lemma [Friedlander and Iwaniec 2010, Theorem 6.9], we have upper (and lower) bounds

$$|\tilde{\mathcal{A}}_{p_1 \cdots p_r}| = S(\mathcal{A}_{p_1 \cdots p_r}, u) \leq g(p_1 \cdots p_r) |\mathcal{A}| V(u) \{1 + O(e^{-1/\varepsilon})\} + \sum_{b \leq D^\varepsilon} \lambda^{(r)}(b) r_{\mathcal{A}}(bp_1 \cdots p_r),$$

(with \leq replaced by \geq for the lower bound) where $\lambda^{(r)}$ is the upper (lower) bound β -sieve of level D^ε when r is even (odd). For further details on the fundamental lemma and β -sieve, we refer the reader to [Friedlander and Iwaniec 2010, Sections 6 and 11].

Plugging back into (5-10), we get

$$\begin{aligned} S(\mathcal{A}, z) &\leq S^*(\tilde{\mathcal{A}}, z) \\ &\leq \sum_{0 \leq r \leq \varepsilon^{-2}} \sum_{(D_1, \dots, D_r) \in \mathcal{D}_r^*} \frac{(-1)^r}{\gamma(D_1, \dots, D_r)} \sum_{\substack{p_1 \cdots p_r | P(z) \\ D_j < p_j \leq D_j^{1+\tau}}} \\ &\quad \times \left\{ g(p_1 \cdots p_r) |\mathcal{A}| V(u) \{1 + O(e^{-1/\varepsilon})\} + \sum_{b \leq D^\varepsilon} \lambda^{(r)}(b) r_{\mathcal{A}}(bp_1 \cdots p_r) \right\}. \end{aligned} \quad (5-11)$$

We now compare the main term above to that of the modified linear sieve, as in (4-3)–(4-11) from the proof of Proposition 4.1, namely,

$$V^*(D, z) := \sum_{d | P(z)/P(u)} \lambda^*(d) g(d) = V(z) \{F^*(s) + o(1)\}.$$

The difference between these main terms is accounted for by those d with two close prime factors, within a ratio D^τ , and those d near the boundary. The former contribution is

$$\sum_{\substack{d | P(z)/P(u) \\ p \leq p' < pD^\tau \\ pp' | d}} g(d) \leq \sum_{\substack{u \leq p < z \\ p \leq p' < pD^\tau}} g(pp') \cdot \prod_{u \leq p < z} (1 + g(p)),$$

and the latter contribution is

$$\sum_r \sum_{\substack{u < p_r < \dots < p_1 < z \\ D^{1/(1+\tau)} < p_1 \cdots p_m^3 < D}} g(p_1 \cdots p_r), \quad (5-12)$$

where m is the first index ($m \leq r$) for which this occurs. Both contributions may be shown to be $O(\varepsilon^5)$, see [Friedlander and Iwaniec 2010, pages 254–255]. Hence the Proposition follows. \square

Remark 5.2. We make a minor technical point. Namely, at an admissible cost $O(\varepsilon^5)$ we may assume Equation (5-1) holds, where \mathcal{D}_r^* is further restricted to vectors (D_1, \dots, D_r) satisfying

$$\nu^{-1}(D_1^{1/(1+\tau)}, \dots, D_r^{1/(1+\tau)}) \subset \mathcal{D}^*, \quad (5-13)$$

regardless of parity of r . To show this, note by definitions of $\mathbf{P}_{4,r}, \mathbf{P}_{6,r}$ the integers $p_1 \cdots p_r$ with $p_j \in [D_j^{1/(1+\tau)}, D_j]$, $j \leq r$, that lie outside \mathcal{D}^* must satisfy

$$x^{1/6+2\eta}/p_1 \quad \text{or} \quad x^{1/12-5\eta}/p_6 \in [x^{-2\tau}, x^{2\tau}].$$

Then for $B_1 = x^{1/6+2\eta+2\tau}$, $B_6 = x^{1/12-5\eta+2\tau}$, we have

$$\sum_{B_i \geq p_i \geq \max(B/x^{4\tau}, u)} g(p_i) \ll \log \frac{\log B_i}{\log \max(B_i/x^{4\tau}, u)} \ll \frac{\log x^{4\tau}}{\log u} \ll \frac{\tau}{\varepsilon^2},$$

and so the contribution of such integers to the main term of (5-11) is

$$\ll \sum_r \sum_{\substack{u < p_r < \cdots < p_1 < z \\ B_i \geq p_i > \max(B_i/x^{4\tau}, u), i \in \{1, 6\}}} g(p_1 \cdots p_r) \ll \frac{\tau}{\varepsilon^2} \prod_{u < p < z} (1 + g(p)) \ll \frac{\tau \log z}{\varepsilon^2 \log u} \ll \varepsilon^5.$$

Hence (5-13) follows.

We now proceed to Theorem 2.12. For each vector $(D_1, \dots, D_r) \in \mathbf{D}_r$ we define the weight $\lambda_{(D_1, \dots, D_r)}$ supported on d in $\nu^{-1}(D_1, \dots, D_r)$, namely,

$$\lambda_{(D_1, \dots, D_r)}(d) := \mathbf{1}_{\substack{d = p_1 \cdots p_r \\ D_j < p_j \leq D_j^{1+\tau} \forall j}}. \quad (5-14)$$

Next, we may decompose an integer d into its u -smooth and rough components, $d = b(p_1 \cdots p_r)$. Recall $\mathcal{D}^{(r)} = \mathcal{D}^\pm$, $\lambda^{(r)} = \lambda^\pm$ (depending on the parity of r), and for $b \mid P(u)$ we have $\lambda^{(r)}(b) = \mu(b)$ if $b \in \mathcal{D}^{(r)}$, and $\lambda^{(r)}(b) = 0$ else. Thus we may define the convolution $\lambda_{(D_1, \dots, D_r)}^{(r)} := \lambda_{(D_1, \dots, D_r)} * \lambda^{(r)}$, i.e.,

$$\lambda_{(D_1, \dots, D_r)}^{(r)}(d) = \begin{cases} \mu(b) & \text{if } d = bp_1 \cdots p_r \text{ for } b \in \mathcal{D}^{(r)}(D^\varepsilon), b \mid P(D^{\varepsilon^2}), \text{ and } D_j < p_j \leq D_j^{1+\tau} \forall j \leq r, \\ 0 & \text{else.} \end{cases} \quad (5-15)$$

Hence the remainder in (5-1) equals

$$\sum_{0 \leq r \leq \varepsilon^{-2}} \sum_{(D_1, \dots, D_r) \in \mathbf{D}_r^*} \frac{(-1)^r}{\gamma(D_1, \dots, D_r)} \sum_{d \mid P(z)} \lambda_{(D_1, \dots, D_r)}^{(r)}(d) r_{\mathcal{A}}(d) = \sum_{d \mid P(z)} \tilde{\lambda}^*(d) r_{\mathcal{A}}(d)$$

for the weights

$$\tilde{\lambda}^* = \sum_{0 \leq r \leq \varepsilon^{-2}} \sum_{(D_1, \dots, D_r) \in \mathbf{D}_r^*} \frac{(-1)^r}{\gamma(D_1, \dots, D_r)} \lambda_{(D_1, \dots, D_r)}^{(r)}. \quad (5-16)$$

Recalling the cardinality of $\mathbf{D}_r^* \subset \mathbf{D}_r^+$ from (5-7), it suffices to show the weights $\lambda_{(D_1, \dots, D_r)}^{(r)}$ are programmably factorable for each vector in \mathbf{D}_r^* . To this we have the following.

Lemma 5.3. *For an integer d denote the vector $\nu(d) = (D_1, \dots, D_r)$ from (5-9). If d has a factorization as in (3-7) at level D , then the corresponding weights $\lambda_{\nu(d)}$ and $\lambda_{\nu(d)}^{(r)}$, as in (5-14) and (5-15), resp., are programmably factorable sequences of levels $D^{1+\tau}$ and $D^{1+\tau+\varepsilon} = x^{7/12+\eta}$, resp. (relative to $x, \varepsilon/50$).*

Proof. By assumption for each $N \in [1, x^{1/3}]$, there is a factorization $d = d_1 d_2 d_3$ satisfying the system (3-7). For $j = 1, 2, 3$, write $d_j = \prod_{i \in I_j} p_i$ for the induced partition of indices $\{1, \dots, r\} = I_1 \cup I_2 \cup I_3$. Thus letting $Q_j = \prod_{i \in I_j} D_i$, the factorization $D_1 \cdots D_r = Q_1 Q_2 Q_3$ satisfies (2-1), since $D_i < p_i$.

Further, writing the corresponding subvectors $(D_i)_{i \in I_j}$ for $j = 1, 2, 3$, the weights $\lambda_{(D_i)_{i \in I_j}}$ are 1-bounded, supported on $[1, Q_j^{1+\tau}]$, and give the desired triple convolution,

$$\lambda_{(D_1, \dots, D_r)} = \lambda_{(D_i)_{i \in I_1}} * \lambda_{(D_i)_{i \in I_2}} * \lambda_{(D_i)_{i \in I_3}}.$$

Hence $\lambda_{(D_1, \dots, D_r)}$ is programmably factorable of level $D^{1+\tau}$ as claimed. Similarly $\lambda_{(D_1, \dots, D_r)}^{(r)} = \lambda_{(D_i)_{i \in I_1}} * \lambda_{(D_i)_{i \in I_2}} * \lambda_{(D_i)_{i \in I_3}}^{(r)}$ is programmably factorable of level $D^{1+\tau+\varepsilon}$. \square

Now for each vector $(D_1, \dots, D_r) \in \mathbf{D}_r^*$, by (5-13) there exists $d = p_1 \cdots p_r \in \mathcal{D}^*$ for some primes $p_i \in (D_i^{1/(1+\tau)}, D_i]$. Then for all $N \in [1, x^{1/3}]$ Proposition 3.3 gives a factorization of d as in (3-7), and so by Lemma 5.3 $\lambda_{(D_1, \dots, D_r)}^{(r)}$ is programmably factorable of level $x^{7/12+\eta}$.

This completes the proof of Theorem 2.12.

5.1. Variable level of distribution for the linear sieve weights. We now return to Iwaniec's well-factorable weights $\tilde{\lambda}^\pm$ for the (upper and lower) linear sieve, given explicitly from (5-15) as the weighted sum,

$$\tilde{\lambda}^\pm = \sum_{0 \leq r \leq \varepsilon^{-2}} \sum_{(D_1, \dots, D_r) \in \mathbf{D}_r^\pm} \frac{(-1)^r}{\gamma(D_1, \dots, D_r)} \lambda_{(D_1, \dots, D_r)}^{(r)}. \quad (5-17)$$

We introduce the analogous set of well-factorable vectors $\mathbf{D}_r^{\text{well}} = \mathbf{D}_r^{\text{well}}(D)$,

$$\mathbf{D}_r^{\text{well}} = \{(D_1, \dots, D_r) : D_1 \cdots D_{m-1} D_m^2 < D \text{ for all } m \leq r\}. \quad (5-18)$$

Note $\mathbf{D}_r^\pm \subset \mathbf{D}_r^{\text{well}}$, having dropped parity conditions on the indices $m \leq r$.

We have the following technical variation on Theorem 2.12 for the original linear sieve.

Proposition 5.4. *Let $(D_1, \dots, D_r) \in \mathbf{D}_r^{\text{well}}(D)$ and write $D = x^\theta$, $D_i = x^{t_i}$ for $i \leq r$. If $\theta \leq \theta(t_1) - \varepsilon$ as in (3-14), then*

$$\sum_{\substack{b=p_1 \cdots p_r \\ D_i < p_i \leq D_i^{1+\tau}}} \sum_{\substack{d=bc \leq x^\theta \\ c \mid P(p_r) \\ (d,a)=1}} \tilde{\lambda}^\pm(d) \left(\pi(x; d, a) - \frac{\pi(x)}{\varphi(d)} \right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}. \quad (5-19)$$

Moreover if $t_1 \leq \frac{1}{5}$ and $r \geq 3$, then (5-19) holds provided that $\theta \leq \theta(t_1, t_2, t_3) - \varepsilon$ as in (3-15).

If $t_1 \leq \frac{1}{5}$ and $r \leq 2$, then provided $\theta \leq \frac{1}{5}(3 - u) - \varepsilon$,

$$\sum_{\substack{b=p_1 \cdots p_r \\ D_i < p_i \leq D_i^{1+\tau}}} \sum_{\substack{d=bc \leq x^\theta \\ c \mid P(x^u) \\ (d,a)=1}} \tilde{\lambda}^\pm(d) \left(\pi(x; d, a) - \frac{\pi(x)}{\varphi(d)} \right) \ll_{a,A,\varepsilon} \frac{x}{(\log x)^A}.$$

In particular for $r = 0$ (i.e., the empty vector), $\theta \leq \frac{1}{5}(3 - u) - \varepsilon$, this simplifies as

$$\sum_{\substack{d \leq x^\theta \\ d | P(x^u) \\ (d, a) = 1}} \tilde{\lambda}^\pm(d) \left(\pi(x; d, a) - \frac{\pi(x)}{\varphi(d)} \right) \ll_{a, A, \varepsilon} \frac{x}{(\log x)^A}.$$

Proof. Given (D_1, \dots, D_r) , take an integer $b = p_1 \cdots p_r$ with $D_i < p_i \leq D_i^{1+\tau}$. Then for all multiples d of b with $d/b | P(p_r)$, the weight $\lambda_{(D'_1, \dots, D'_s)}^{(s)}(d)$ vanishes unless the vector (D'_1, \dots, D'_s) extends (D_1, \dots, D_r) . That is, $D'_i = D_i$ for all $i \leq r$. Conversely, given such a vector (D'_1, \dots, D'_s) we have $\lambda_{(D'_1, \dots, D'_s)}^{(s)}(d') = 0$ unless the first s primes of d' are $p_1 \cdots p_s$ with $D_i < p_i \leq D_i^{1+\tau}$, $i \leq r$. So by the definition of $\tilde{\lambda}^\pm$ as in (5-17), we have

$$\sum_{\substack{b = p_1 \cdots p_r \\ D_i < p_i \leq D_i^{1+\tau}}} \sum_{\substack{d = bc \leq x^\theta \\ c | P(p_r) \\ (d, a) = 1}} \tilde{\lambda}^\pm(d) = \sum_{r \leq s \leq \varepsilon^{-2}} \sum_{\substack{(D'_1, \dots, D'_s) \in \mathbf{D}_s^\pm \\ D'_i = D_i, i \leq r}} \frac{(-1)^s}{\gamma(D'_1, \dots, D'_s)} \sum_{\substack{d \leq x^\theta \\ (d, a) = 1}} \lambda_{(D'_1, \dots, D'_s)}^{(s)}(d). \quad (5-20)$$

Here we have extended (by zero) the inner sum to all $d \leq x^\theta$, $(d, a) = 1$.

Next, take such a vector $(D'_1, \dots, D'_s) \in \mathbf{D}_s^\pm(x^\theta)$ with $D'_i = D_i$ for $i \leq r$. Each integer $d = p_1 \cdots p_s$ with $D'_i < p_i \leq (D'_i)^{1+\tau}$ lies in $d \in \mathcal{D}^{\text{well}}(x^{\theta+\tau})$. In particular $p_1 \leq D_1^{1+\tau} \leq x^{t_1+\tau}$ so by Corollary 3.8, d has a factorization as in (3-7) at level $x^{\theta(t_1+\tau)}$. Since $\theta(t)$ is continuous (in fact, piecewise linear), for $\tau > 0$ sufficiently small $\theta(t_1 + \tau) \geq \theta(t_1) - \varepsilon \geq \theta$. Thus by Lemma 5.3 the weights $\lambda_{(D'_1, \dots, D'_s)}^{(s)}$ are programmably factorable sequences of level x^θ . Hence for each such vector, by Theorem 2.5 we have

$$\sum_{\substack{d \leq x^\theta \\ (d, a) = 1}} \lambda_{(D'_1, \dots, D'_s)}^{(s)}(d) \ll_{a, A, \varepsilon} \frac{x}{(\log x)^A}. \quad (5-21)$$

Plugging (5-21) back into (5-20) gives the bound (5-19), as claimed.

Moreover, if $t_1 \leq \frac{1}{5}$ and $r \geq 3$ then proceeding as in the above paragraph, by Corollary 3.8 d will factor as in (3-7) to level $x^{\theta(t_1+\tau, t_2+\tau, t_3+\tau)}$. Again $\theta(t, u, v)$ is continuous, so for $\tau > 0$ sufficiently small

$$\theta(t_1 + \tau, t_2 + \tau, t_3 + \tau) \geq \theta(t_1, t_2, t_3) - \varepsilon \geq \theta.$$

Hence (5-19) also follows in this case.

Similarly if $t_1 \leq \frac{1}{5}$ and $r \leq 2$, proceeding as above with $d = p_1 \cdots p_s$, the assumption $d/b | P(x^u)$ implies $s \leq 2$ or $p_3 \leq D_3^{1+\tau} \leq x^{u+\tau}$. Thus by Corollary 3.8 d will factor as in (3-7) to level $x^{(3-u-\tau)/5} \geq x^\theta$. Hence (5-19) follows in this case as well. \square

5.2. Equidistribution for products of primes. We use an extension of Theorem 2.5 to products of k primes. This is the analogue in the programmably factorable setting of [Wu 1990, Lemma 7] extending Theorem 2.2 of Bombieri, Friedlander and Iwaniec.

Proposition 5.5. *Let $\varepsilon > 0$ and λ be a programmably factorable sequence of level $D \leq x^{3/5-\varepsilon}$ (relative to $x, \varepsilon/50$). Take real numbers $\varepsilon_1, \dots, \varepsilon_k \geq \varepsilon$ such that $\sum_{i \leq k} \varepsilon_i = 1$. Then for any fixed integer $a \in \mathbb{Z}$, $A, B > 0$, letting $\Delta = 1 + (\log x)^{-B}$ we have*

$$\sum_{\substack{d \leq D \\ (d,a)=1}} \lambda(d) \left(\sum_{\substack{p_1 \cdots p_k \equiv a \pmod{d} \\ x^{\varepsilon_i} / \Delta < p_i \leq x^{\varepsilon_i} \forall i \leq k}} 1 - \frac{1}{\varphi(d)} \sum_{\substack{(p_1 \cdots p_k, d)=1 \\ x^{\varepsilon_i} / \Delta < p_i \leq x^{\varepsilon_i} \forall i \leq k}} 1 \right) \ll_{a,\varepsilon,A,B} \frac{x}{(\log x)^A}. \quad (5-22)$$

Proof. This follows by the same proof method as in Theorem 2.5 (i.e., [Maynard 2020, Theorem 1.1]). Indeed, Maynard just uses the Heath-Brown identity to decompose the indicator function of primes into Type I/II sums. A similar decomposition holds for products of k primes, after which we may apply the same Type I/II estimates in Propositions 5.1 and 5.2 of [Maynard 2020]. \square

In addition, by replacing Theorem 2.5 with Proposition 5.5 in the proof, we obtain analogues of Proposition 5.4 for the linear sieve weights $\lambda = \tilde{\lambda}^\pm$ in the case of products of k primes.

Corollary 5.6. *Let $(D_1, \dots, D_r) \in \mathbf{D}_r^{\text{well}}(D)$ and write $D = x^\theta$, $D_i = x^{t_i}$ for $i \leq r$. Let $\varepsilon > 0$ and real numbers $\varepsilon_1, \dots, \varepsilon_k \geq \varepsilon$ such that $\sum_{i \leq k} \varepsilon_i = 1$. Fix an integer $a \in \mathbb{Z}$, $A, B > 0$, and let $\Delta = 1 + (\log x)^{-B}$. If $\theta \leq \theta(t_1) - \varepsilon$ as in (3-14),*

$$\sum_{\substack{b=p'_1 \cdots p'_r \\ D_i < p'_i \leq D_i^{1+\tau}}} \sum_{\substack{d=bc \leq x^\theta \\ c | P(p'_i) \\ (d,a)=1}} \tilde{\lambda}^\pm(d) \left(\sum_{\substack{p_1 \cdots p_k \equiv a \pmod{d} \\ x^{\varepsilon_i} / \Delta < p_i \leq x^{\varepsilon_i} \forall i \leq k}} 1 - \frac{1}{\varphi(d)} \sum_{\substack{(p_1 \cdots p_k, d)=1 \\ x^{\varepsilon_i} / \Delta < p_i \leq x^{\varepsilon_i} \forall i \leq k}} 1 \right) \ll_{a,\varepsilon,A,B} \frac{x}{(\log x)^A}. \quad (5-23)$$

Moreover if $t_1 \leq \frac{1}{5}$, $r \geq 3$, then (5-23) holds provided that $\theta \leq \theta(t_1, t_2, t_3) - \varepsilon$ as in (3-15).

In addition, if $r \leq 2$, $u \leq t_r$, $t_1 \leq \frac{1}{5}$, and $\theta \leq \frac{3-u}{5} - \varepsilon$, then

$$\sum_{\substack{b=p'_1 \cdots p'_r \\ D_i < p'_i \leq D_i^{1+\tau}}} \sum_{\substack{d=bc \leq x^\theta \\ c | P(x^u) \\ (d,a)=1}} \tilde{\lambda}^\pm(d) \left(\sum_{\substack{p_1 \cdots p_k \equiv a \pmod{d} \\ x^{\varepsilon_i} / \Delta < p_i \leq x^{\varepsilon_i} \forall i \leq k}} 1 - \frac{1}{\varphi(d)} \sum_{\substack{(p_1 \cdots p_k, d)=1 \\ x^{\varepsilon_i} / \Delta < p_i \leq x^{\varepsilon_i} \forall i \leq k}} 1 \right) \ll_{a,\varepsilon,A,B} \frac{x}{(\log x)^A}. \quad (5-24)$$

6. Upper bound for twin primes

Now we shall apply the modified sieve to the set of twin primes

$$\mathcal{A} := \{p + 2 : p \leq x\}.$$

In this case the sieve notation specializes as $\mathcal{P} = \{p > 2\}$ and $g(d) = 1/\varphi(d)$ for odd d , so that $V(z) = \prod_{2 < p < z} (1 - 1/\varphi(p))$. Recall $V(z) \sim \mathfrak{S}_2/e^\gamma \log z$ by Mertens theorem, for the Hardy–Littlewood constant $\mathfrak{S}_2 = 2 \prod_{p > 2} (1 - 2/p)/(1 - 1/p)^2$ appearing in $\Pi(x) = \mathfrak{S}_2 x / (\log x)^2$.

We begin in the spirit of Fouvry and Grupp [1986], and apply a weighted sieve inequality. To each nonswitched term, we apply the Buchstab identity in order to lower the sieve threshold down to $z = x^\epsilon$ for some tiny $\epsilon > 0$. By Proposition 5.4, such smooth sums will satisfy level of distribution $\frac{1}{3}(3 - \epsilon)$.

Combined with variations on a theme, which identify programmably factorable weights in certain cases, the consequent increase in level will be sufficient to obtain the bound in Theorem 1.2. For a final bit of savings, we use refinements obtained by Wu's iteration method [2004].

Before moving on, we note that sieve methods and the switching principle have also yielded progress on the Goldbach problem. Indeed, the upper bound in [Wu 2004, Theorem 3] for twin primes is obtained by using the same formulae as in [loc. cit., Theorem 1] for the Goldbach problem, except for altering the level from $\frac{1}{2}$ to $\frac{4}{7}$ (this amounts to replacing factors of 4 with $\frac{7}{2}$ in a few instances). Importantly, the quantitative upper bounds for twin primes are much stronger than those of the Goldbach problem. This is because the latter relies on level $\frac{1}{2}$ from Bombieri and Vinogradov for a growing residue $a = N$, while the former may appeal to level of distribution $\frac{4}{7}$ from Bombieri, Friedlander and Iwaniec for the fixed residue $a = 2$ (and now the subsequent improvements of Maynard). As such our methods have nothing new to say for the Goldbach problem.

6.1. Lemmas. We begin with a standard lemma for $x^{1/u}$ -rough numbers in terms of the Buchstab function $\omega(u) = (f(u) + F(u))/(2e^\gamma)$ for linear sieve functions f, F as in Theorem 2.10. Alternatively, ω is directly defined via

$$\begin{aligned}\omega(u) &= \frac{1}{u} && \text{for } 1 \leq u \leq 2, \\ (u\omega(u))' &= \omega(u-1) && \text{for } 2 \leq u.\end{aligned}$$

Lemma 6.1. *Let $x \geq 2$ and $y = x^{1/u}$. Then we have*

$$\sum_{\substack{n \leq x \\ p | n \Rightarrow p \geq y}} 1 = \omega(u) \frac{x}{\log y} + O\left(\frac{x}{(\log y)^2}\right).$$

Proof. This is [Wu 1990, Lemma 12]. □

The argument of Wu makes essential use of weighted sieve inequalities, as in [Wu 2004, Lemmas 4.1 and 4.2]. We shall employ the latter inequality in the special case $d = 1, \sigma = 1$.

Lemma 6.2. *For $\frac{3}{10} \geq \rho \geq \tau_3 > \tau_2 > \tau_1 \geq \rho' \geq \frac{1}{20}$, we have*

$$5S(\mathcal{A}, x^\rho) \lesssim \sum_{1 \leq n \leq 21} \Gamma_n,$$

where

$$\begin{aligned}\Gamma_1 &:= 4S(\mathcal{A}, x^{\rho'}) + S(\mathcal{A}, x^{\tau_1}), & \Gamma_{12} &:= \sum_{\substack{x^{\rho'} \leq p_1 < p_2 < x^{\tau_1} \\ x^{\tau_3} \leq p_3 < x^\rho}} \sum \sum S(\mathcal{A}_{p_1 p_2 p_3}, p_2), \\ \Gamma_2 &:= - \sum_{x^{\rho'} \leq p < x^\rho} S(\mathcal{A}_p, x^{\rho'}), & \Gamma_{13} &:= \sum_{\substack{x^{\rho'} \leq p_1 < x^{\tau_1} \leq p_2 < x^{\tau_2} \leq p_3 < x^\rho}} \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3}, p_2), \\ \Gamma_3 &:= - \sum_{x^{\rho'} \leq p < x^{\tau_2}} S(\mathcal{A}_p, x^{\rho'}), & \Gamma_{14} &:= \sum_{\substack{x^{\rho'} \leq p_1 < x^{\tau_1} \\ x^{\tau_2} \leq p_2 < p_3 < x^\rho}} \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3}, p_2), \\ \Gamma_4 &:= - \sum_{x^{\rho'} \leq p < x^{\tau_3}} S(\mathcal{A}_p, x^{\rho'}),\end{aligned}$$

$$\begin{aligned}
\Gamma_5 &:= \sum_{x^{\rho'} \leq p_1 < p_2 < x^{\tau_2}} \sum S(\mathcal{A}_{p_1 p_2}, x^{\rho'}), & \Gamma_{15} &:= \sum_{x^{\tau_1} \leq p_1 < x^{\tau_2} \leq p_2 < x^{\tau_3} \leq p_3 < x^{\rho}} \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3}, p_2), \\
\Gamma_6 &:= \sum_{\substack{x^{\rho'} \leq p_1 < x^{\tau_1} \\ x^{\tau_2} \leq p_2 < x^{\tau_3}}} \sum \sum S(\mathcal{A}_{p_1 p_2}, x^{\rho'}), & \Gamma_{16} &:= \sum_{x^{\tau_2} \leq p_1 < p_2 < p_3 < p_4 < x^{\tau_3}} \sum \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3 p_4}, p_3), \\
\Gamma_7 &:= \sum_{x^{\rho'} \leq p_1 < p_2 < x^{\tau_1}} \sum \sum S(\mathcal{A}_{p_1 p_2}, p_1), & \Gamma_{17} &:= \sum_{x^{\tau_2} \leq p_1 < p_2 < p_3 < x^{\tau_3} \leq p_4 < x^{\rho}} \sum \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3 p_4}, p_3), \\
\Gamma_8 &:= \sum_{x^{\rho'} \leq p_1 < x^{\tau_1} \leq p_2 < x^{\tau_2}} \sum \sum S(\mathcal{A}_{p_1 p_2}, p_1), & \Gamma_{18} &:= \sum_{x^{\tau_2} \leq p_1 < p_2 < x^{\tau_3} \leq p_3 < p_4 < x^{\rho}} \sum \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3 p_4}, p_3), \\
\Gamma_9 &:= \sum_{x^{\tau_1} \leq p_1 < p_2 < p_3 < x^{\tau_3}} \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3}, p_2), & \Gamma_{19} &:= \sum_{\substack{x^{\tau_1} \leq p_1 x^{\tau_2} \\ x^{\tau_3} \leq p_2 < p_3 < p_4 < x^{\rho}}} \sum \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3 p_4}, p_3), \\
\Gamma_{10} &:= \sum_{x^{\tau_1} \leq p_1 < p_2 < x^{\tau_2} \leq p_3 < x^{\rho}} \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3}, p_2), & \Gamma_{20} &:= \sum_{x^{\tau_2} \leq p_1 < x^{\tau_3} \leq p_2 < p_3 < p_4 < p_5 < x^{\rho}} \sum \sum \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3 p_4 p_5}, p_4), \\
\Gamma_{11} &:= \sum_{x^{\tau_1} \leq p_1 < x^{\tau_2} \leq p_2 < p_3 < x^{\tau_3}} \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3}, p_2), & \Gamma_{21} &:= \sum_{x^{\tau_3} \leq p_1 < p_2 < p_3 < p_4 < p_5 < p_6 < x^{\rho}} \sum \sum \sum \sum \sum \sum S(\mathcal{A}_{p_1 p_2 p_3 p_4 p_5 p_6}, p_5).
\end{aligned}$$

Proof. This is [Wu 2004, Lemma 4.2] with $d = 1$, $\sigma = 1$. Here we simplify Wu's notation slightly, using $(\underline{d}^{1/s}, \underline{d}^{1/\kappa_3}, \underline{d}^{1/\kappa_2}, \underline{d}^{1/\kappa_1}, \underline{d}^{1/s'}) = (x^{\rho}, x^{\tau_3}, x^{\tau_2}, x^{\tau_1}, x^{\rho'})$. The basic proof idea is to iterate the Buchstab identity and to strategically neglect some terms by positivity. \square

6.2. Computations. Given $0.1 \leq \rho' \leq \tau_1 < 0.2 \leq \tau_2 < \tau_3 \leq \rho \leq 0.3$., we define integrals $I_n = I_n(\rho, \rho', \tau_1, \tau_2, \tau_3)$ by

$$\begin{aligned}
I_n &= \int_{\mathbb{D}_n} \omega \left(\frac{1-t-u-v}{u} \right) \frac{dt du dv}{tu^2v} & (9 \leq n \leq 15), \\
I_n &= \int_{\mathbb{D}_n} \omega \left(\frac{1-t-u-v-w}{v} \right) \frac{dt du dv dw}{tuv^2w} & (16 \leq n \leq 19), \\
I_{20} &= \int_{\mathbb{D}_{20}} \omega \left(\frac{1-t-u-v-w-x}{w} \right) \frac{dt du dv dw dx}{tuvw^2x}, \\
I_{21} &= \int_{\mathbb{D}_{21}} \omega \left(\frac{1-t-u-v-w-x-y}{x} \right) \frac{dt du dv dw dx dy}{tuvwx^2y},
\end{aligned} \tag{6-1}$$

where ω is the Buchstab function, and where the domains \mathbb{D}_n are

$$\begin{aligned}
\mathbb{D}_9 &= \{(t, u, v) : \tau_1 < t < u < v < \tau_3\}, \\
\mathbb{D}_{10} &= \{(t, u, v) : \tau_1 < t < u < \tau_2 < v < \rho\}, \\
\mathbb{D}_{11} &= \{(t, u, v) : \tau_1 < t < \tau_2 < u < v < \tau_3\}, \\
\mathbb{D}_{12} &= \{(t, u, v) : \rho' < t < u < \tau_1, \tau_3 < v < \rho\}, \\
\mathbb{D}_{13} &= \{(t, u, v) : \rho' < t < \tau_1 < u < \tau_2 < v < \rho\}, \\
\mathbb{D}_{14} &= \{(t, u, v) : \rho' < t < \tau_1, \tau_2 < u < v < \rho\},
\end{aligned}$$

$$\begin{aligned}
\mathbb{D}_{15} &= \{(t, u, v) : \tau_1 < t < \tau_2 < u < \tau_3 < v < \rho\}, \\
\mathbb{D}_{16} &= \{(t, u, v, w) : \tau_2 < t < u < v < w < \tau_3\}, \\
\mathbb{D}_{17} &= \{(t, u, v, w) : \tau_2 < t < u < v < \tau_3 < w < \rho\}, \\
\mathbb{D}_{18} &= \{(t, u, v, w) : \tau_2 < t < u < \tau_3 < v < w < \rho\}, \\
\mathbb{D}_{19} &= \{(t, u, v, w) : \tau_1 < t < \tau_2, \tau_3 < u < v < w < \rho\}, \\
\mathbb{D}_{20} &= \{(t, u, v, w, x) : \tau_2 < t < \tau_3 < u < v < w < x < \rho\}, \\
\mathbb{D}_{21} &= \{(t, u, v, w, x, y) : \tau_3 < t < u < v < w < x < y < \rho\}.
\end{aligned}$$

Recall the definitions (3-14) and (3-15),

$$\theta(t) = \begin{cases} \frac{2-t}{3} & \text{if } t > \frac{1}{5}, \\ \frac{1+t}{2} & \text{if } t \leq \frac{1}{5}. \end{cases}$$

We let $\theta_\epsilon = \frac{3-\epsilon}{5}$ and

$$\theta(t, u, v) := \max\left\{\frac{3-v}{5}, \theta(t), \theta(u), \theta(t+u+v), \theta(t+u), \theta(t+v), \theta(u+v)\right\}.$$

We also define

$$\begin{aligned}
G_1 &= 4G(\rho') + G(\tau_1), & G_3 &= G_0 + \bar{G}(\tau_2), \\
G_2 &= G_0 + \bar{G}(\rho), & G_4 &= G_0 + \bar{G}(\tau_3),
\end{aligned} \tag{6-2}$$

where for $c \leq \frac{1}{5}$,

$$\begin{aligned}
G(c) &= \frac{1}{\epsilon} F(\theta_\epsilon/\epsilon) - \frac{1}{\epsilon} \int_\epsilon^c \frac{dt}{t} f((\theta_\epsilon - t)/\epsilon) + \frac{1}{\epsilon} \int_\epsilon^c \int_\epsilon^t \frac{dt \, du}{tu} F((\theta_\epsilon - t - u)/\epsilon) \\
&\quad - \int_\epsilon^c \int_\epsilon^t \int_\epsilon^u \frac{dt \, du \, dv}{tuv^2} f((\theta(t, u, v) - t - u - v)/v), \tag{6-3}
\end{aligned}$$

and for $c > \frac{1}{5}$,

$$\bar{G}(c) = -\frac{1}{\epsilon} \int_{1/5}^c \frac{dt}{t} f((\theta(t) - t)/\epsilon) + \int_{1/5}^c \int_\epsilon^{\rho'} \frac{dt \, du}{tu^2} F((\theta(t) - t - u)/u) \tag{6-4}$$

as well as

$$\begin{aligned}
G_0 &= -\frac{1}{\epsilon} \int_{\rho'}^{1/5} \frac{dt}{t} f((\theta_\epsilon - t)/\epsilon) + \frac{1}{\epsilon} \int_{\rho'}^{1/5} \int_\epsilon^{\rho'} \frac{dt \, du}{tu} F((\theta_\epsilon - t - u)/\epsilon) \\
&\quad - \int_{\rho'}^{1/5} \int_\epsilon^{\rho'} \int_\epsilon^u \frac{dt \, du \, dv}{tuv^2} f((\theta(t, u, v) - t - u - v)/v). \tag{6-5}
\end{aligned}$$

We similarly let

$$\begin{aligned}
 G_5 &= \frac{1}{\epsilon} \int_{\rho'}^{1/5} \int_{\rho'}^t \frac{dt du}{tu} F((\theta_\epsilon - t - u)/\epsilon) + \frac{1}{\rho'} \int_{1/5}^{\tau_2} \int_{\rho'}^t \frac{dt du}{tu} F((\theta(t) - t - u)/\rho') \\
 &\quad - \int_{\rho'}^{1/5} \int_{\rho'}^t \int_{\epsilon}^{\rho'} \frac{dt du dv}{tuv^2} f((\theta(t, u, v) - t - u - v)/v), \\
 G_6 &= \frac{1}{\rho'} \int_{\tau_2}^{\tau_3} \int_{\rho'}^{\tau_1} \frac{dt du}{tu} F((\theta(t) - t - u)/\rho'), \\
 G_7 &= \frac{1}{\epsilon} \int_{\rho'}^{\tau_1} \int_{\rho'}^t \frac{dt du}{tu} F((\theta_\epsilon - t - u)/\epsilon) - \int_{\rho'}^{\tau_1} \int_{\rho'}^t \int_{\epsilon}^u \frac{dt du dv}{tuv^2} f((\theta(t, u, v) - t - u - v)/v), \\
 G_8 &= \frac{1}{\epsilon} \int_{\tau_1}^{1/5} \int_{\rho'}^{\tau_1} \frac{dt du}{tu} F((\theta_\epsilon - t - u)/\epsilon) + \int_{1/5}^{\tau_2} \int_{\rho'}^{\tau_1} \frac{dt du}{tu^2} F((\theta(t) - t - u)/u) \\
 &\quad - \int_{\tau_1}^{1/5} \int_{\rho'}^{\tau_1} \int_{\epsilon}^u \frac{dt du dv}{tuv^2} f((\theta(t, u, v) - t - u - v)/v).
 \end{aligned} \tag{6-6}$$

Recall the sieve functions F , f satisfy $F(s) = 2e^\gamma/s$ for $s \in [1, 3]$, $f(s) = 2e^\gamma \log(s-1)/s$ for $s \in [2, 4]$ and $F(s) = 2e^\gamma/s \cdot [1 + \int_2^{s-1} f(t) dt]$ for all $s \geq 1$.

The main bound is the following.

Proposition 6.3. *Let $0 < \epsilon \leq 0.1 \leq \rho' \leq \tau_1 < 0.2 \leq \tau_2 < \tau_3 \leq \rho \leq 0.3$. Then for I_n , G_n , and $G(c)$ as in (6-1), (6-2), (6-6), and (6-3), we have*

$$S(\mathcal{A}, x^\rho) \lesssim \frac{\Pi(x)}{5e^\gamma} \left(\sum_{n=1}^8 G_n + G\left(\frac{1}{5}\right) \sum_{n=9}^{21} I_n \right). \tag{6-7}$$

Proof. We first bound $S(\mathcal{A}, x^c)$ for $c \in [\epsilon, \frac{1}{5}]$. By the Buchstab identity,

$$S(\mathcal{A}, x^c) = S(\mathcal{A}, x^\epsilon) - \sum_{x^\epsilon \leq p < x^c} S(\mathcal{A}_p, p).$$

Iterating twice more, we obtain

$$S(\mathcal{A}, x^c) = S(\mathcal{A}, x^\epsilon) - \sum_{x^\epsilon \leq p_1 < x^c} S(\mathcal{A}_{p_1}, x^\epsilon) + \sum_{x^\epsilon \leq p_2 < p_1 < x^c} S(\mathcal{A}_{p_1 p_2}, x^\epsilon) - \sum_{x^\epsilon \leq p_3 < p_2 < p_1 < x^c} S(\mathcal{A}_{p_1 p_2 p_3}, p_3). \tag{6-8}$$

To each term $S(\mathcal{A}_d, x^\epsilon)$ above, we apply the linear sieve of level x^{θ_ϵ} for $\theta_\epsilon = \frac{1}{5}(3 - \epsilon)$, as in Theorem 2.10. And to each term $S(\mathcal{A}_{p_1 p_2 p_3}, p_3)$, we apply the linear sieve of level x^θ for $\theta = \theta(t_1, t_2, t_3)$, where $p_i = x^{t_i}$.

To handle the corresponding error terms, note for primes $x^\epsilon \leq p_2 < p_1 < x^c \leq x^{1/5}$ and $d \in \{1, p_1, p_1 p_2\}$, the prime factors of q/d above are bounded by x^ϵ so that the sets \mathcal{A}_q are equidistributed to level θ_ϵ . Hence for each $x^\epsilon \leq p_2 < p_1 < x^{1/5}$, by Proposition 5.4 with $u = \epsilon$, $b = 1$, $p_1, p_1 p_2$, we have

$$\sum_{\substack{q \leq x^{\theta_\epsilon} \\ q | P(x^\epsilon)}} \tilde{\lambda}^+(q) \left(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\varphi(q)} \right) = \sum_{\substack{q \leq x^{\theta_\epsilon} \\ q | P(x^\epsilon)}} \tilde{\lambda}^+(q) \left(\pi(x; q, -2) - \frac{\pi(x)}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A}$$

and

$$\sum_{p_1} \sum_{\substack{q=p_1 m \leq x^{\theta_\epsilon} \\ m \mid P(x^\epsilon)}} \tilde{\lambda}^-(q) \left(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A}$$

and

$$\sum_{p_2, p_1} \sum_{\substack{q=p_1 p_2 m \leq x^{\theta_\epsilon} \\ m \mid P(x^\epsilon)}} \tilde{\lambda}^+(q) \left(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A}.$$

In addition for each $p_3 < p_2 < p_1 < x^{1/5}$, $p_i = x^{t_i}$, letting $\theta = \theta(t_1, t_2, t_3)$, by Proposition 5.4 with $b = p_1 p_2 p_3$,

$$\sum_{p_3, p_2, p_1} \sum_{\substack{q=p_1 p_2 p_3 m \leq x^\theta \\ m \mid P(p_3)}} \tilde{\lambda}^-(q) \left(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A}.$$

Thus for $c \leq \frac{1}{5}$, the linear sieve bounds give

$$S(\mathcal{A}, x^\epsilon) \lesssim |\mathcal{A}| V(x^\epsilon) F\left(\frac{\theta_\epsilon}{\epsilon}\right) \quad (6-9)$$

and

$$\sum_{x^\epsilon \leq p_1 < x^c} S(\mathcal{A}_{p_1}, x^\epsilon) \gtrsim \sum_{x^\epsilon \leq p_1 < x^c} |\mathcal{A}| g(p_1) V(x^\epsilon) f\left(\frac{\theta_\epsilon - t_1}{\epsilon}\right) \quad (6-10)$$

and

$$\sum_{x^\epsilon \leq p_2 < p_1 < x^c} S(\mathcal{A}_{p_1 p_2}, x^\epsilon) \lesssim \sum_{x^\epsilon \leq p_2 < p_1 < x^c} |\mathcal{A}| g(p_1 p_2) V(x^\epsilon) F\left(\frac{\theta_\epsilon - t_1 - t_2}{\epsilon}\right) \quad (6-11)$$

and

$$\sum_{x^\epsilon \leq p_3 < p_2 < p_1 < x^c} S(\mathcal{A}_{p_1 p_2 p_3}, p_3) \gtrsim \sum_{x^\epsilon \leq p_3 < p_2 < p_1 < x^c} |\mathcal{A}| g(p_1 p_2 p_3) V(p_3) f\left(\frac{\theta - t_1 - t_2 - t_3}{t_3}\right). \quad (6-12)$$

Hence by (6-9), (6-10), (6-11), (6-12), we observe that (6-8) becomes

$$\begin{aligned} S(\mathcal{A}, x^c) &\lesssim -|\mathcal{A}| \sum_{x^\epsilon \leq p_3 < p_2 < p_1 < x^c} g(p_1 p_2 p_3) V(p_3) f\left(\frac{\theta - t_1 - t_2 - t_3}{\epsilon}\right) \\ &\quad + |\mathcal{A}| V(x^\epsilon) \left(F\left(\frac{\theta_\epsilon}{\epsilon}\right) - \sum_{x^\epsilon \leq p_1 < x^c} g(p_1) f\left(\frac{\theta_\epsilon - t_1}{\epsilon}\right) + \sum_{x^\epsilon \leq p_2 < p_1 < x^c} g(p_1 p_2) F\left(\frac{\theta_\epsilon - t_1 - t_2}{\epsilon}\right) \right). \end{aligned} \quad (6-13)$$

Recall $V(z) \sim \mathfrak{S}_2 / e^\gamma \log z$ by Merten's theorem. Thus by partial summation and the prime number theorem, we obtain

$$S(\mathcal{A}, x^c) \lesssim \frac{\Pi(x)}{e^\gamma} G(c), \quad (6-14)$$

for $G(c)$ as in (6-3). Hence for $c = \rho'$, τ_1 we have $c \in [\epsilon, \frac{1}{5}]$, so we bound Γ_1 as

$$\Gamma_1 = 4S(\mathcal{A}, x^{\rho'}) + S(\mathcal{A}, x^{\tau_1}) \lesssim \frac{\Pi(x)}{e^\gamma} (4G(\rho') + G(\tau_1)) = \frac{\Pi(x)}{e^\gamma} G_1.$$

Now consider $c, c' \in [\frac{1}{5}, \frac{2}{7}]$. We shall apply the linear sieve of level $\theta(t_1)$, as in (3-14). In general, for $p_1 = x^{t_1}$ and $\theta(t_1)$ as in (3-14), Proposition 5.4 gives

$$\sum_{p_1} \sum_{\substack{p_1 | q, q \leq x^{\theta(t_1)} \\ q/p_1 | P(p_1)}} \tilde{\lambda}^-(q) \left(|\mathcal{A}_q| - \frac{|\mathcal{A}|}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A},$$

so that for $c, c' \in [\frac{1}{5}, \frac{2}{7}]$, the linear sieve of level $\theta(t_1)$ gives

$$\sum_{x^{c'} \leq p_1 < x^c} S(\mathcal{A}_{p_1}, p_1) \gtrsim \sum_{x^{c'} \leq p_1 < x^c} |\mathcal{A}| g(p_1) V(p_1) f\left(\frac{\theta(t_1) - t_1}{t_1}\right). \quad (6-15)$$

Thus by partial summation and the prime number theorem,

$$\Gamma_2 = \sum_{x^{\rho'} \leq p < x^\rho} S(\mathcal{A}_p, x^{\rho'}) \lesssim \frac{\Pi(x)}{e^\gamma} (G + \bar{G}(\rho)) = \frac{\Pi(x)}{e^\gamma} G_2.$$

Similarly, we obtain

$$\Gamma_n \lesssim \frac{\Pi(x)}{e^\gamma} G_n \quad \text{for } 1 \leq n \leq 8. \quad (6-16)$$

Finally, for the remaining Γ_n , we apply the switching principle. Namely, for Γ_9 we have

$$\Gamma_9 = \sum_{x^{\tau_1} \leq p_1 < p_2 < p_3 < x^{\tau_3}} \sum \sum S(\mathcal{A}_{p_1 p_2 p_3}, p_2) = S(\mathcal{B}, x^{1/2}) + O(x^{1/2}) \quad (6-17)$$

for the set

$$\mathcal{B} = \{p_1 p_2 p_3 m - 2 \leq x : x^{\tau_1} \leq p_1 < p_2 < p_3 < x^{\tau_3}, p' | m \Rightarrow p' \geq p_2\}. \quad (6-18)$$

Note since $p_2, p_1 > x^{\tau_1} > x^{0.1}$, each m above has at most 7 prime factors.

Now by a standard subdivision argument, \mathcal{B} is similarly equidistributed in arithmetic progressions as is \mathcal{A} . Indeed, the basic idea is to partition $\mathcal{B} = \bigcup_{r \leq 7} \mathcal{B}^{(r)}$, where $\mathcal{B}^{(r)}$ is the subset corresponding to integers m with r prime factors. Then we cover the prime tuples (p_1, \dots, p_r) into hypercubes of the form

$$[\Delta^{l_1}, \Delta^{l_1+1}) \times \dots \times [\Delta^{l_r}, \Delta^{l_r+1})$$

for $\Delta = 1 + (\log x)^{-B}$ with $B > 0$ sufficiently large, and apply Corollary 5.6 with $u = \epsilon$, $b = 1$. This gives

$$\sum_{\substack{q \leq x^{\theta \epsilon} \\ q | P(x^\epsilon)}} \tilde{\lambda}^+(q) \left(|\mathcal{B}_q| - \frac{|\mathcal{B}|}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A}.$$

Similarly for each $x^\epsilon < p'_3 < p'_2 < p'_1 < x^{1/5}$, by Corollary 5.6 with $u = \epsilon$ and $b = p'_1, p'_1 p'_2$, we have

$$\sum_{p'_1} \sum_{\substack{q = p'_1 m' \leq x^{\theta \epsilon} \\ m' | P(x^\epsilon)}} \tilde{\lambda}^-(q) \left(|\mathcal{B}_q| - \frac{|\mathcal{B}|}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A}$$

and

$$\sum_{p'_2, p'_1} \sum_{\substack{q=p'_1 p'_2 m' \leq x^{\theta \epsilon} \\ m' | P(x^\epsilon)}} \tilde{\lambda}^+(q) \left(|\mathcal{B}_q| - \frac{|\mathcal{B}|}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A}.$$

In addition for each $p'_3 < p'_2 < p'_1 < x^{1/5}$, $p'_i = x^{t_i}$, letting $\theta = \theta(t_1, t_2, t_3)$, by Corollary 5.6 with $b = p'_1 p'_2 p'_3$, we have

$$\sum_{p'_3, p'_2, p'_1} \sum_{\substack{q=p'_1 p'_2 p'_3 m' \leq x^\theta \\ m' | P(p'_3)}} \tilde{\lambda}^-(q) \left(|\mathcal{B}_q| - \frac{|\mathcal{B}|}{\varphi(q)} \right) \ll_A \frac{x}{(\log x)^A}.$$

Iterating the Buchstab identity, we have

$$\begin{aligned} S(\mathcal{B}, x^{1/2}) &\leq S(\mathcal{B}, x^{1/5}) \\ &= S(\mathcal{B}, x^\epsilon) - \sum_{x^\epsilon \leq p'_1 < x^{1/5}} S(\mathcal{B}_{p'_1}, x^\epsilon) + \sum_{x^\epsilon \leq p'_2 < p'_1 < x^{1/5}} S(\mathcal{B}_{p'_1 p'_2}, x^\epsilon) - \sum_{x^\epsilon \leq p'_3 < p'_2 < p'_1 < x^{1/5}} S(\mathcal{B}_{p'_1 p'_2 p'_3}, p'_3), \end{aligned}$$

and hence by the linear sieve bounds we obtain

$$\Gamma_9 \lesssim S(\mathcal{B}, x^{1/2}) \leq S(\mathcal{B}, x^{1/5}) \lesssim e^{-\gamma} \frac{G(\frac{1}{5})}{\log x} \mathfrak{S}_2 |\mathcal{B}|. \quad (6-19)$$

Now to compute $|\mathcal{B}|$ in (6-18), by Lemma 6.1 we have

$$|\mathcal{B}| \sim \frac{x}{\log x} \int_{\tau_1 < t_1 < t_2 < t_3 < \tau_2} \frac{dt_1 dt_2 dt_3}{t_1 t_2^2 t_3} \omega \left(\frac{1 - t_1 - t_2 - t_3}{t_2} \right) = \frac{x}{\log x} \cdot I_9.$$

Thus we have $\Gamma_9 \lesssim e^{-\gamma} G(\frac{1}{5}) \Pi(x) I_9$. Similarly, we obtain

$$\Gamma_n \lesssim G(\frac{1}{5}) \frac{\Pi(x)}{e^\gamma} I_n \quad (\text{for } 9 \leq n \leq 21). \quad (6-20)$$

Hence plugging (6-16), (6-20) into Lemma 6.2 completes the proof. \square

Let

$$\begin{aligned} \rho &= 0.27195, & \tau_3 &= 0.24589, \\ \rho' &= 0.12313, & \tau_2 &= 0.20867, \\ \epsilon &= 0.002, & \tau_1 &= 0.16288. \end{aligned} \quad (6-21)$$

For such choices of parameters, we compute the following integrals from Proposition 6.3,

$$\sum_{9 \leq n \leq 21} I_n \leq 0.174404, \quad \sum_{1 \leq n \leq 8} G_n \leq 28.34581, \quad G(\frac{1}{5}) \leq 5.99237.$$

Thus by Proposition 6.3, we obtain the bound

$$\pi_2(x) \lesssim S(\mathcal{A}, x^\rho) \lesssim \frac{\Pi(x)}{5e^\gamma} \left(\sum_{n=1}^8 G_n + G\left(\frac{1}{5}\right) \sum_{n=9}^{21} I_n \right) \lesssim 3.30042\Pi(x). \quad (6-22)$$

We also record the individual bounds (see table at end of paper).

6.3. Completing the proof of Theorem 1.2. We shall refine our argument in certain cases for which Wu's iteration method [2004] applies directly without modification. As such we have chosen simplicity over full optimization.

In the lemma below, we consider the cases of sets $\mathcal{A}_{p_1 p_2}$, where p_1, p_2 lie in a prescribed range, and such that for all multiples $bp_1 p_2$, the sets $\mathcal{A}_{bp_1 p_2}$ are equidistributed to level x^θ (we also need level x^θ for corresponding switched sets \mathcal{B} of integers $mbp_1 p_2 - 2$).

Lemma 6.4. *Let $\theta \in [\frac{1}{2}, 1]$, $s \in [2, 3]$, and $\mathcal{A} = \{p+2 : p \leq x\}$. There is a function $H_\theta(s)$, monotonically increasing in θ for fixed s and decreasing in s for fixed θ , such that the following holds: For each $(D_1, D_2) \in \mathbf{D}_2^{\text{well}}(x^\theta)$, we have*

$$\sum_{\substack{D_1 < p_1 < D_1^{1+\tau} \\ D_2 < p_2 < D_2^{1+\tau}}} S(\mathcal{A}_{p_1 p_2}, z) \lesssim \frac{\Pi(x)}{e^\gamma} \sum_{\substack{D_1 < p_1 < D_1^{1+\tau} \\ D_2 < p_2 < D_2^{1+\tau}}} \frac{\log x}{\varphi(p_1 p_2) \log z} \left(F(s) - \frac{2e^\gamma}{s} H_\theta(s) \right), \quad (6-23)$$

where $z^s = x^\theta / p_1 p_2$, provided (5-22) holds for $\lambda = \tilde{\lambda}^+$ at level $D = x^\theta$ with $(x^{\varepsilon_1}, x^{\varepsilon_2}) = (D_1, D_2)$, and provided for all vectors $(D_1, \dots, D_r) \in \mathbf{D}_r^{\text{well}}(x^\theta)$ extending (D_1, D_2) ,

$$\sum_{\substack{b=p_1 \cdots p_r \\ D_i < p_i \leq D_i^{1+\tau}}} \sum_{\substack{b|d, d \leq x^\theta \\ (d,a)=1}} \tilde{\lambda}^\pm(d) \left(\pi(x; d, a) - \frac{\pi(x)}{\varphi(d)} \right) \ll_{a,\varepsilon,A} \frac{x}{(\log x)^A}.$$

Proof. Wu iterates the weighted sieve inequality [Wu 2004, Lemmas 4.1 and 4.2] on the subset of (nonswitched) terms whose sieving parameter s lies in the interval $s \in [2, 3]$. This yields a recurrence relation for a function $H_\theta(s)$, which encodes the percent savings over the (normalized) linear sieve $sF(s)/(2e^\gamma)$.

Starting from a term $S(\mathcal{A}_{p_1 p_2}, z)$, each successive iteration of the weighted sieve inequality is composed of terms of the form $S(\mathcal{A}_{bp_1 p_2}, z')$ for some multiple $bp_1 p_2$ corresponding to some vector extending (D_1, D_2) . By assumption, all such sets are equidistributed to level x^θ , when weighted by the upper/lower linear sieve. Similarly, the switched sets are also equidistributed to level x^θ (here we only need the upper bound weights $\tilde{\lambda}^+$). Finally, the savings function $H_\theta(s)$ inherits the stated monotonicity properties by construction of the iteration. \square

The function H_θ depends on the known level of distribution x^θ (i.e., Wu used $\theta = \frac{1}{2}$ for Goldbach, and $\theta = \frac{4}{7}$ for twin primes). For parameters as in Tables 1 and 2 of [Wu 2004, pages 30–32],

$$H_{1/2}(t) \geq \begin{cases} 0.0223939 & \text{if } 2.0 < t \leq 2.2, \\ 0.0217196 & \text{if } 2.2 < t \leq 2.3, \\ 0.0202876 & \text{if } 2.3 < t \leq 2.4, \\ 0.0181433 & \text{if } 2.4 < t \leq 2.5, \\ 0.0158644 & \text{if } 2.5 < t \leq 2.6, \\ 0.0129923 & \text{if } 2.6 < t \leq 2.7, \\ 0.0100686 & \text{if } 2.7 < t \leq 2.8, \\ 0.0078162 & \text{if } 2.8 < t \leq 2.9, \\ 0.0072943 & \text{if } 2.9 < t \leq 3.0, \\ 0 & \text{else,} \end{cases}$$

and

$$H_{4/7}(t) \geq \begin{cases} 0.0287118 & \text{if } 2.0 \leq t \leq 2.1, \\ 0.0280509 & \text{if } 2.1 < t \leq 2.2, \\ 0.0264697 & \text{if } 2.2 < t \leq 2.3, \\ 0.0241936 & \text{if } 2.3 < t \leq 2.4, \\ 0.0214619 & \text{if } 2.4 < t \leq 2.5, \\ 0.0183875 & \text{if } 2.5 < t \leq 2.6, \\ 0.0149960 & \text{if } 2.6 < t \leq 2.7, \\ 0.0117724 & \text{if } 2.7 < t \leq 2.8, \\ 0.0094724 & \text{if } 2.8 < t \leq 2.9, \\ 0.0090024 & \text{if } 2.9 < t \leq 3.0, \\ 0 & \text{else.} \end{cases}$$

As such, Wu [2004, Theorem 3] obtained $\pi_2(x)/\Pi(x) \lesssim \frac{7}{2}(1 - H_{4/7}(2.1)) \leq 3.39951$.

To complete our proof of Theorem 1.2 we apply Lemma 6.4, now valid up to level $x^{7/12}$ by Corollary 2.7 and Proposition 5.5 for $\tilde{\lambda}^+$. Note when the largest integration variable is $t \geq \frac{1}{5}$, we have $\theta(t) = \frac{1}{3}(2-t) \leq \frac{7}{12}$ if and only if $t \geq \frac{1}{4}$. A key feature we use to satisfy the conditions of Lemma 6.4 is that the level $x^{\theta(t)}$ persists, since the largest prime $p_1 = x^{t_1}$ is preserved through successive iterations.

Thus, in practice, Lemma 6.4 simply amounts to modifying the integral in G_2 by substituting $F(s) - \frac{2e^\gamma}{s}H_\theta(s)$ in for $F(s)$, $s = (\theta(t) - t - u)/u$, when $t \geq \frac{1}{4}$ (the only parameter $\geq \frac{1}{4}$ is ρ , so we only refine G_2). Denote this as G_2^{Wu} . For ease we also use $H_\theta(s) \geq H_{4/7}(s)$, by monotonicity in θ . Doing so, with the same parameter choices (6-21), we obtain $G_2^{\text{Wu}} \leq -5.598667$ and hence

$$\pi_2(x) \lesssim \frac{\Pi(x)}{5e^\gamma} \left(G_2^{\text{Wu}} + \sum_{1 \leq n \neq 2 \leq 8} G_n + G\left(\frac{1}{5}\right) \sum_{9 \leq n \leq 21} I_n \right) \lesssim 3.299552\Pi(x). \quad (6-24)$$

This completes the proof of Theorem 1.2.

n	G_n	n	I_n	n	I_n
1	39.00163	9	0.0332157	17	0.000315
2	-5.591009	10	0.0228322	18	0.000269
3	-3.986553	11	0.0092564	19	0.000164
4	-5.060499	12	0.0150101	20	$\leq 2.70 \cdot 10^{-6}$
5	1.864133	13	0.0547244	21	$\leq 5.50 \cdot 10^{-9}$
6	0.741181	14	0.0260202		
7	0.453663	15	0.0124636		
8	0.923736	16	0.0001314		

Table 2. Values of G_1, \dots, G_8 and I_9, \dots, I_{21} .

For slight numerical gains, one may compute $H_\theta(s)$ when $\theta \in [\frac{4}{7}, \frac{7}{12}]$, by tweaking the formulae in [Wu 2004]. More substantially, Wu defined a lower bound savings $h_\theta(s)$, for a substitution of $f(s)$ by $f(s) + \frac{2e^\gamma}{s} h_\theta(s)$. But in practice, to compute h would require derivations (analogous to H) of as yet undetermined formulae. We leave these to the reader.

Acknowledgments

The author is grateful to James Maynard for suggesting the problem and for many valuable discussions. The author also thanks Carl Pomerance, Kyle Pratt, and the anonymous referee for helpful feedback. The author was supported by a Clarendon Scholarship at the University of Oxford.

References

- [Bombieri and Davenport 1966] E. Bombieri and H. Davenport, “Small differences between prime numbers”, *Proc. Roy. Soc. Lond. Ser. A* **293** (1966), 1–18. MR Zbl
- [Bombieri et al. 1986] E. Bombieri, J. B. Friedlander, and H. Iwaniec, “Primes in arithmetic progressions to large moduli”, *Acta Math.* **156**:3-4 (1986), 203–251. MR Zbl
- [Brun 1919] V. Brun, “La série $1/5 + 1/7 + 1/11 + 1/13 + 1/17 + 1/19 + 1/29 + 1/31 + 1/41 + 1/43 + 1/59 + 1/61 + \dots$ où les dénominateurs sont nombres premiers jumeaux est convergente ou finie”, *Bull. Sci. Math.* **43** (1919), 100–104, 124–128. Zbl
- [Cai and Lu 2003] Y. Cai and M. Lu, “On the upper bound for $\pi_2(x)$ ”, *Acta Arith.* **110**:3 (2003), 275–298. MR Zbl
- [Chen 1973] J. R. Chen, “On the representation of a larger even integer as the sum of a prime and the product of at most two primes”, *Sci. Sinica* **16** (1973), 157–176. MR Zbl
- [Chen 1978] J. R. Chen, “On the Goldbach’s problem and the sieve methods”, *Sci. Sinica* **21**:6 (1978), 701–739. MR Zbl
- [Elliott and Halberstam 1970] P. D. T. A. Elliott and H. Halberstam, “A conjecture in prime number theory”, pp. 59–72 in *Symposia Mathematica, IV* (Rome, 1968/1969), Academic Press, London, 1970. MR Zbl
- [Fouvry 1984] É. Fouvry, “Autour du théorème de Bombieri–Vinogradov”, *Acta Math.* **152**:3-4 (1984), 219–244. MR Zbl
- [Fouvry and Grupp 1986] É. Fouvry and F. Grupp, “On the switching principle in sieve theory”, *J. Reine Angew. Math.* **370** (1986), 101–126. MR Zbl
- [Fouvry and Iwaniec 1983] É. Fouvry and H. Iwaniec, “Primes in arithmetic progressions”, *Acta Arith.* **42**:2 (1983), 197–218. MR Zbl
- [Friedlander and Iwaniec 2010] J. Friedlander and H. Iwaniec, *Opera de cribro*, Amer. Math. Soc. Colloq. Publ. **57**, Amer. Math. Soc., Providence, RI, 2010. MR Zbl

- [Hardy and Littlewood 1923] 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**:1 (1923), 1–70. MR Zbl
- [Iwaniec 1980] H. Iwaniec, “A new form of the error term in the linear sieve”, *Acta Arith.* **37** (1980), 307–320. MR Zbl
- [Maynard 2020] J. Maynard, “Primes in arithmetic progressions to large moduli, II: Well-factorable estimates”, preprint, 2020. arXiv 2006.07088
- [Pan 1964] C.-d. Pan, “A new application of the Yu. V. Linnik large sieve method”, *Acta Math. Sinica* **14** (1964), 597–606. In Chinese, translated in *Chinese Math. Acta* **5** (1964), 642–652. MR Zbl
- [Riesel and Vaughan 1983] H. Riesel and R. C. Vaughan, “On sums of primes”, *Ark. Mat.* **21**:1 (1983), 46–74. MR Zbl
- [Selberg 1952] A. Selberg, “On elementary methods in prime number theory and their limitations”, pp. 13–22 in *Den 11te Skandinaviske Matematikerkongress* (Trondheim, Norway, 1949), Tanum, Oslo, 1952. MR Zbl
- [Siebert 1976] H. Siebert, “Montgomery’s weighted sieve for dimension two”, *Monatsh. Math.* **82**:4 (1976), 327–336. MR Zbl
- [Wu 1990] J. Wu, “Sur la suite des nombres premiers jumeaux”, *Acta Arith.* **55**:4 (1990), 365–394. MR Zbl
- [Wu 2004] J. Wu, “Chen’s double sieve, Goldbach’s conjecture and the twin prime problem”, *Acta Arith.* **114**:3 (2004), 215–273. MR Zbl

Communicated by Andrew Granville

Received 2022-05-08 Revised 2024-01-09 Accepted 2024-02-13

jared.d.lichtman@gmail.com

Mathematical Institute, University of Oxford, Oxford, United Kingdom

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR
Antoine Chambert-Loir
Université Paris-Diderot
France

EDITORIAL BOARD CHAIR
David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Jason P. Bell	University of Waterloo, Canada	Philippe Michel	École Polytechnique Fédérale de Lausanne
Bhargav Bhatt	University of Michigan, USA	Martin Olsson	University of California, Berkeley, USA
Frank Calegari	University of Chicago, USA	Irena Peeva	Cornell University, USA
J.-L. Colliot-Thélène	CNRS, Université Paris-Saclay, France	Jonathan Pila	University of Oxford, UK
Brian D. Conrad	Stanford University, USA	Anand Pillay	University of Notre Dame, USA
Samit Dasgupta	Duke University, USA	Bjorn Poonen	Massachusetts Institute of Technology, USA
Hélène Esnault	Freie Universität Berlin, Germany	Victor Reiner	University of Minnesota, USA
Gavril Farkas	Humboldt Universität zu Berlin, Germany	Peter Sarnak	Princeton University, USA
Sergey Fomin	University of Michigan, USA	Michael Singer	North Carolina State University, USA
Edward Frenkel	University of California, Berkeley, USA	Vasudevan Srinivas	SUNY Buffalo, USA
Wee Teck Gan	National University of Singapore	Shunsuke Takagi	University of Tokyo, Japan
Andrew Granville	Université de Montréal, Canada	Pham Huu Tiep	Rutgers University, USA
Ben J. Green	University of Oxford, UK	Ravi Vakil	Stanford University, USA
Christopher Hacon	University of Utah, USA	Akshay Venkatesh	Institute for Advanced Study, USA
Roger Heath-Brown	Oxford University, UK	Melanie Matchett Wood	Harvard University, USA
János Kollár	Princeton University, USA	Shou-Wu Zhang	Princeton University, USA
Michael J. Larsen	Indiana University Bloomington, USA		

PRODUCTION

production@msp.org
Silvio Levy, Scientific Editor

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2025 is US \$565/year for the electronic version, and \$820/year (+\$70, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW[®] from MSP.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2025 Mathematical Sciences Publishers

Algebra & Number Theory

Volume 19 No. 1 2025

A modification of the linear sieve, and the count of twin primes JARED DUKER LICHTMAN	1
Ranks of abelian varieties in cyclotomic twist families ARI SHNIDMAN and ARIEL WEISS	39
Picard rank jumps for K3 surfaces with bad reduction SALIM TAYOU	77
Curves with few bad primes over cyclotomic \mathbb{Z}_ℓ -extensions SAMIR SIKSEK and ROBIN VISSER	113
Vanishing results for the coherent cohomology of automorphic vector bundles over the Siegel variety in positive characteristic THIBAUT ALEXANDRE	143
Super-Hölder vectors and the field of norms LAURENT BERGER and SANDRA ROZENSZTAJN	195