The Quarterly Journal of Mathematics, 2026, 00, 1–25
https://doi.org/10.1093/qmath/haag010
Published: 13 May 2026
Article

OXFORD

EXISTENCE OF SMOOTH NUMBERS IN SHORT INTERVALS

by SARVAGYA JAIN *

Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland

*Corresponding author. E-mail: sajain@utu.fi

ABSTRACT

Let $X \geq y \geq 2$, and let $u = \frac{\log X}{\log y}$. We say a number is *y-smooth* if all of its prime factors are less than or equal to y. In this paper, we study the distribution of y-smooth numbers in short intervals. In particular, for $y \geq \exp\left((\log X)^{2/3+\varepsilon}\right)$, we show that the interval $[x, x+h]$ contains a y-smooth number for almost all $x \in [X, 2X]$, provided $h \geq \exp\left((1+\varepsilon)\left(\frac{11}{8}u\log u + 4\log\log X\right)\right)$, and X is sufficiently large depending on ε. This result improves upon an earlier result by Matomäki. Additionally, we provide the corresponding 'all intervals' type result. Our approach relies on a strategically factorized Dirichlet polynomial, much like the earlier work of Matomäki. The improvement in our results stems from the integration of ideas introduced in the breakthrough work of Matomäki and Radziwiłł.

1. INTRODUCTION

Let $P(n)$ denote the largest prime factor of n. A number n is said to be y-smooth (or y-friable) if $P(n) \leq y$. We denote the set of all y-smooth numbers by $\mathbf{S}(y)$. Let $\Psi(x, y)$ denote the number of y-smooth numbers up to x, that is, $|\mathbf{S}(y) \cap [1, x]|$. It is known that $\Psi(x, y) \sim \rho(u)x$ holds for a wide range of the smoothness parameter y, where $u = \frac{\log x}{\log y}$ and $\rho(u)$ is the Dickman function (see Equation 3.1 below for its definition). Nevertheless, our understanding of the distribution of smooth numbers in short intervals is still far from complete.

One might expect that for a wide range of parameters h and y, an asymptotic of the form

$$\Psi(x+h, y) - \Psi(x, y) \sim \rho(u)h$$

holds. In a breakthrough paper, Matomäki and Radziwiłł proved that for any fixed $u \geq 1$ and any function $\psi(x) \to \infty$ as $x \to \infty$, the above asymptotic is valid for almost all $x \in [X, 2X]$, when $y = X^{1/u}$ and $h = \psi(X)$ (see [10]).

Goudout [3, Théorème 4] extended this result by showing that the asymptotic holds when

$$(1+\rho(u)^{-1})^{\psi(X)} \leq h \leq X \quad \text{and} \quad 1 \leq u \leq (\log X)^{1/6-o(1)},$$

which is equivalent to $y \geq \exp\left((\log X)^{5/6+o(1)}\right)$.

Received: 7 April 2025. **Revised:** 2 February 2026. **Accepted:** 16 March 2026

A natural question is whether one can at least guarantee the existence of a smooth number in a short interval, that is, whether

$$\Psi\left(x+h,y\right)-\Psi\left(x,y\right)\geq 1,$$

holds for a broader range of parameters h and y. Since $\rho(u)=\frac{1}{u^{u+o(u)}}$, probabilistic heuristics suggest that one should expect a y-smooth number to appear in almost every interval of length

$$h\geq \exp\Big((1+o(1))\,u\log u\Big).$$

However, establishing such a result remains a significant challenge.

In our work, we follow the approach of Matomäki and Radziwiłł (also used by Goudout) to advance this relaxed problem. For readers familiar with Goudout's work, see Remarks 1.1 and 1.2 for a discussion of the key differences in our method.

At a high level, our approach integrates the use of factorized Dirichlet polynomials for detecting smooth numbers—drawing on the work of Soundararajan [13] and Matomäki [9]—with the ideas developed by Matomäki and Radziwiłł.

THEOREM 1.1. *For any $\varepsilon>0$, there exists a positive constant $C=C(\varepsilon)$ such that the following holds. Let $X\geq 2$ be large enough depending on ε. If*

$$\exp\left(C(\log X)^{2/3}(\log\log X)^{4/3}\right)\leq y\leq X^{\frac{1}{C}}$$

and

$$h\geq \exp\left((1+\varepsilon)\left(\frac{11}{8}u\log u+4\log\log X\right)\right),$$

for $u=\frac{\log X}{\log y}$, then the interval $[x,x+h]$ contains a y-smooth number for almost all $x\in[X,2X]$.

In particular, we improve upon an earlier result of Matomäki [9, Theorem 1.3], where the corresponding conditions were

$$y\geq \exp\left((\log X)^{2/3}(\log\log X)^{4/3+\varepsilon}\right)$$

and

$$h\geq \exp\left(\left(\frac{14}{3}+\varepsilon\right)(4u\log u+\log\log X)\right).$$

In this paper, we focus on the regime where y is small (or equivalently, u is large). This permits a loss of logarithmic factors, which is acceptable in this context. (Note that for small u, Goudout's results yield a shorter interval length. In fact, by adapting Goudout's methods, one can remove the logarithmic factors entirely to show that smooth numbers exist in almost every interval of length

$$\exp((c+o(1))\,u\log u),$$

for some absolute constant $c>0$ worse than $11/8$, even when u is small. We do not pursue this refinement here.)

One can also obtain results about existence in all short intervals rather than in almost all short intervals. In this direction, Matomäki and Radziwiłł showed that for every fixed $u\geq 1$, there exists a sufficiently large constant $C_u>0$ such that every interval $[x,x+h]$ contains an $x^{1/u}$-smooth number, provided that $h\geq C_u\sqrt{x}$ (see [10, Corollary 1]). This unconditional result improves upon earlier work by Soundararajan [13], who obtained a similar conclusion under the assumption of the Riemann hypothesis.

Again, Goudout [3, Théorème 5] extended their result, showing that one can take $1 \leq u \leq (\log X)^{1/6-\varepsilon}$, provided that the length of the short interval satisfies $h \geq \sqrt{x}\rho(u)^{-3-\varepsilon}$.[1]

In the same spirit as these works, we also prove a result concerning the existence of a y-smooth number in all intervals of the form $[x, x + x^{1/2+o(1)}]$ for sufficiently large x.

THEOREM 1.2. *For any $\varepsilon > 0$, there exists a positive constant $C = C(\varepsilon)$ such that the following holds. If x is large in terms of ε,*

$$\exp\left(C(\log x)^{2/3}(\log\log x)^{4/3}\right) \leq y \leq x^{\frac{1}{C}},$$

and

$$h \geq \sqrt{x}\exp\left((1+\varepsilon)\left(\frac{11}{16}\tilde{u}\log\tilde{u} + 2\log\log x\right)\right),$$

where $\tilde{u} = \frac{\log x}{\log y}$, then the interval $[x, x+h]$ contains a y-smooth number.

The above result can be contrasted with Theorem 1.1 in [9]. Matomäki proved in this theorem that, for sufficiently large x, one could take

$$h \geq \sqrt{x}\exp\left(\left(\frac{7}{3}+\varepsilon\right)\left(4\tilde{u}\log\tilde{u} + \log\log x\right)\right),$$

provided that $y \geq \exp\left((\log x)^{2/3}(\log\log x)^{4/3+\varepsilon}\right)$.

For asymptotic formulas in short intervals when y is small, the known results are even more restrictive. For instance, Younis [15, Theorem 1.1] established that an asymptotic formula holds in all intervals under the conditions

$$x^{17/30+o(1)} \leq h \leq x \quad \text{and} \quad \exp\Big((\log x)^{2/3+o(1)}\Big) \leq y \leq 2x.$$

There have been various other results concerning the distribution of smooth numbers in short intervals. For example, Hildebrand and Tenenbaum [7] proved an asymptotic formula for the number of smooth numbers in almost all short intervals. However, the intervals they consider are significantly longer than those considered here. Building on their work, Friedlander and Granville [2] obtained the corresponding 'all intervals' type result. We refer the reader to [4, 7] for more detailed surveys of results on smooth numbers and their applications to other problems in analytic number theory and cryptography.

1.1. Sketch of the argument

To orient the reader and highlight the key differences from previous works, we now present a brief outline of the proof of Theorem 1.1. In this sketch, we use the imprecise notations $\lesssim$, $\gtrsim$, and $\approx$, to conveniently suppress the factors of $\rho(u)^\varepsilon$ and $\log X$.

Suppose that we want to show for some $h \geq 2$, the interval $[x, x+h]$ contains a y-smooth number for almost all $x \in [X, 2X]$. Choose weights $\{w_n\}_n$ such that $w_n \geq 0$ if $n \in \mathbf{S}(y)$ and 0 otherwise. Let H be such that $\sum_{x\leq n\leq x+H} w_n > 0$ for all $x \in [X, 2X]$. It turns out that, for our choice of weights, we can take $H = Xy^{-3/8}$ (see Lemma 4.2 below). If we show that

$$\left|\frac{1}{h}\sum_{x\leq n\leq x+h} w_n - \frac{1}{H}\sum_{x\leq n\leq x+H} w_n\right| = o\left(\frac{1}{H}\sum_{x\leq n\leq x+H} w_n\right)$$

for almost all $x \in [X, 2X]$, then we will have our result. By a routine Chebyshev-style argument (see Section 7 below), we can reduce it further to showing that

$$\frac{1}{X}\int_X^{2X}\left|\frac{1}{h}\sum_{x\leq n\leq x+h} w_n - \frac{1}{H}\sum_{x\leq n\leq x+H} w_n\right|^2 dx = o\left(\left(\frac{1}{X}\sum_{X\leq n\leq 2X} w_n\right)^2\right).$$

[1] By combining the remark following [3, Théorème 5] with Lemma 3.3 below, one can in fact take the improved range $h \geq \sqrt{x}\rho(u)^{-23/16-\varepsilon}$.

With our choice of w_n, this reduces to essentially showing that the left-hand side is $o(\rho(u)^2)$. Now, letting $F(s) = \sum_n w_n/n^s$, we bound the second moment on the left-hand side by using a Parseval-style bound (see Section 5 below). Doing this, we essentially get the bound

$$\frac{1}{X}\int_X^{2X} \left| \frac{1}{h} \sum_{x \le n \le x+h} w_n - \frac{1}{H} \sum_{x \le n \le x+H} w_n \right|^2 dx \lesssim \int_{y^{1/8}}^{\frac{X}{h}} |F(1+it)|^2 dt.$$

We choose our weights so that

$$F(s) = P_1(s)P_2(s)P_3(s)^J M(s)$$

for

$$P_j(s) = \sum_{p_j \sim P_j} 1/p^s \quad \text{and} \quad M(s) = \sum_{m \in \mathcal{M}} 1/m^s$$

for some $\mathcal{M} \subseteq \mathbf{S}(y)$ and $P_1 \le P_2 \le P_3 \le y$.

REMARK 1.1. *Unlike Goudout's method, since we are focusing only on the existence of y-smooth numbers, we may use any convenient Dirichlet polynomial supported on y-smooth numbers. In particular, our polynomials P_j are supported on primes from a full dyadic intervals (in contrast with [3, 10]). This choice eliminates the need for executing finer than dyadic division and gives a better short interval length when y is small.*

If we carefully choose the parameters P_1, P_2, P_3, J and the set $\mathcal{M}$, we can exploit the factorization of our polynomial to obtain non-trivial savings in its mean value estimate using the techniques introduced in [10]. Our goal is to get enough savings so that we have the bound

$$\int_{y^{1/8}}^{\frac{X}{h}} |F(1+it)|^2 dt = o\left(\rho(u)^2\right).$$

The details of how one can get such a saving are given in Section 6. Here, we content ourselves with a very informal sketch. We begin by choosing $0 < \alpha_1 < \alpha_2 < 1/4$ with $\alpha_1, \alpha_2 \approx 1/4$ and then partitioning the interval $[y^{1/8}, X/h]$ into $\mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_3$, where $t \in \mathcal{T}_1$ if $|P_1(1+it)| \le P_1^{-\alpha_1}$, $t \in \mathcal{T}_2$ if $|P_1(1+it)| > P_1^{-\alpha_1}$, $|P_2(1+it)| \le P_2^{-\alpha_2}$ and $\mathcal{T}_3$ contains the remaining values of t.

To bound the integral over $\mathcal{T}_1$, we note that

$$\int_{\mathcal{T}_1} |F(1+it)|^2 dt \lesssim P_1^{-2\alpha_1} \int_{-X/h}^{X/h} |P_2(1+it)P_3(1+it)^J M(1+it)|^2 dt.$$

At this point, it may seem tempting to use the standard mean value theorem (see Lemma 2.1 below) to bound the integral above. But the approach mentioned above turns out to be lossy as it does not incorporate the sparsity of the coefficients of the Dirichlet polynomial $P_2(s)P_3(s)^J M(s)$, say $\{a_n\}_n$, that comes from being supported on y-smooth numbers. To exploit this sparsity better, one instead uses a variant of the standard mean value theorem that preserves some non-diagonal terms (see Lemma 2.2 below). This naturally leads us to upper bounding sums essentially of the form $\sum_{n \sim N} \mathbf{1}_{\mathbf{S}(y)}(n)\mathbf{1}_{\mathbf{S}(y)}(n+k)$ (see Lemma 6.2 below). These kinds of sums have been considered in earlier works, and bounds of the form $O(\rho(u)^{\varphi}N)$ for some $1 < \varphi < 2$ are known. We use the best existing bound, that is, the largest available φ (see Lemma 3.3 below). This essentially results in an upper bound of the form

$$\int_{\mathcal{T}_1} |F(1+it)|^2 \lesssim \frac{P_1^{1/2}\rho(u)}{h} + \frac{\rho(u)^{\varphi}}{P_1^{1/2}}.$$

To bound the integral over $\mathcal{T}_2$, we note that

$$\int_{\mathcal{T}_2} |F(1+it)|^2 dt \lesssim P_2^{-2\alpha_2} \int_{-X/h}^{X/h} |P_1(1+it)P_3(1+it)^J M(1+it)|^2 dt.$$

To bound the integral above, we exploit $|P_1^{\alpha_1} P_1(1+it)| \geq 1$ to increase the length of Dirichlet polynomial $|P_1(1+it)P_3(1+it)^J M(1+it)|^2$ by multiplying it with $|P_1^{\alpha_1} P_1(1+it)|^{2\ell}$ for some appropriately chosen $\ell = \left\lceil \frac{\log P_2}{\log P_1} \right\rceil$ (see Lemma 6.3 below). This results in an upper bound of the form

$$\int_{\mathcal{T}_2} |F(1+it)|^2 dt \lesssim P_2^{-2(\alpha_2-\alpha_1)} \int_{-X/h}^{X/h} |P_1(1+it)^{\ell+1} P_3(1+it)^J M(1+it)|^2 dt.$$

The motivation behind these manipulations is the well-known fact that the standard mean value theorem is the most efficient when the length of the Dirichlet polynomial involved and the range of integration are of similar order (see [**8**, Chapter 9] for further discussion). Applying the standard mean value theorem and choosing P_2 appropriately in terms of P_1, we have that

$$\int_{\mathcal{T}_2} |F(1+it)|^2 \lesssim \frac{P_1^{1/2}\rho(u)}{h} + \frac{\rho(u)^{\varphi}}{P_1^{1/2}}.$$

Finally, to bound the integral over $\mathcal{T}_3$, we first note that $|P_3(1+it)| \lesssim P_3^{-\sigma_0}$, where $\sigma_0 \asymp (\log X)^{-2/3}(\log\log X)^{-1/3}$ for $y^{1/8} \leq t \leq X$ and hence for all $t \in \mathcal{T}_3$. This bound is derived from Perron's formula combined with the Vinogradov–Korobov zero-free region (see Lemma 2.6 below).

REMARK 1.2. *Unlike Goudout's method, we amplify the savings from the cancellations in the polynomial $P_3(s)$ by raising it to the power J (in his work, the polynomial $P_\infty(s)$ plays a similar role as $P_3(s)$ here). This boost lets us work with smaller values of y.*

To maximize this saving coming from $P_3(s)$, it is natural to choose P_3 as large as possible, that is, $P_3 \approx y$. Consequently, for a non-trivial saving, we require $y \geq \exp\left(\sigma_0^{-1}\right)$. After further bookkeeping, we obtain a slightly stronger restriction on y.

Substituting the bound for $P_3(s)$, we arrive at the estimate

$$\int_{\mathcal{T}_3} |F(1+it)|^2 dt \lesssim y^{-2J\sigma_0} \int_{\mathcal{T}_3} |P_1(1+it)P_2(1+it)M(1+it)|^2 dt.$$

Now we can find a well-spaced set $\mathcal{T} \subseteq \mathcal{T}_3$ so that we have the bound

$$\int_{\mathcal{T}_3} |F(1+it)|^2 dt \lesssim y^{-2J\sigma_0} \sum_{t\in\mathcal{T}} |P_1(1+it)P_2(1+it)M(1+it)|^2.$$

By using Halász–Montgomery inequality (see Lemma 2.5 below), we can bound the above by

$$\int_{\mathcal{T}_3} |F(1+it)|^2 dt \lesssim y^{-2J\sigma_0} \rho(u) \left(1 + \frac{|\mathcal{T}|y^J}{\sqrt{X}}\right).$$

It turns out that one can use $|P_2(1+it)| > P_2^{-\alpha_2}$ for all $t \in \mathcal{T}$ to show that $|\mathcal{T}|y^J \lesssim X^{1/2-\delta}$ for some $\delta > 0$ (using Lemma 2.4 below). This step is the primary reason for choosing $\alpha_2 < 1/4$. Additionally, by choosing J appropriately, we have the bound $y^{-2J\sigma_0} \lesssim \rho(u)^{24}$. Thus, we have the bound

$$\int_{\mathcal{T}_3} |F(1+it)|^2 dt \lesssim \rho(u)^{25}.$$

In conclusion, we have the bound

$$\int_{y^{1/8}}^{X/h} |F(1+it)|^2 dt \lesssim \frac{P_1^{1/2}\rho(u)}{h} + \frac{\rho(u)^{\varphi}}{P_1^{1/2}} + \rho(u)^{25}.$$

Since we want the right-hand side to be $o(\rho(u))^2$, we are forced to choose $P_1 \gtrsim \rho(u)^{2\varphi-4}$. Finally, if we set $h \gtrsim P_1^{1/2}\rho(u)^{-1} \gtrsim \rho(u)^{\varphi-3}$, the right-hand side will be $o(\rho(u)^2)$. By our earlier remarks, we will have a y-smooth number in almost all short intervals of size $\gtrsim \rho(u)^{\varphi-3}$ (the additional $\log X$ factors arise from careful bookkeeping).

REMARK 1.3. *As can be seen from the above sketch, improvements in Lemmas 2.6 and 3.3 are expected to lead to refinements in the smoothness parameter and short interval length, respectively.*

1.2. Notations

For the reader's convenience, we summarize the notations and conventions used throughout this article. The symbols p, q, p_i and q_i are reserved for primes, while d, k, ℓ, m and n denote positive integers. We write $\sum^{\ddagger}$ to indicate that the summation variable runs over powers of 2. The notation $n \sim X$ in a sum restricts the variable to the interval $X < n \leq 2X$.

The parameter $\varepsilon > 0$ denotes an arbitrarily small, fixed constant, and $C_1, C_2, \ldots$ represent positive absolute constants. We use the standard asymptotic notation $o(\cdot), O(\cdot), \ll$, and $\gg$, writing $X \asymp Y$ to denote $X \ll Y \ll X$. Let $\mathbf{1}_S$ denote the indicator function of a set S.

We denote the set of y-smooth numbers by $\mathbf{S}(y)$, and let $\Psi(X, y)$ be the count of y-smooth numbers not exceeding X. Finally, $\rho(u)$ denotes the Dickmann function; we shall frequently employ the estimate $\rho(u) = u^{-u-o(u)}$ without further explicit mention.

2. DIRICHLET POLYNOMIAL ESTIMATES

In this section, we recall various estimates from the literature for Dirichlet polynomials that we will need. Throughout this section, we let $X \geq 1$ and $T > 0$. Additionally, for $a_n \in \mathbb{C}$, we define the Dirichlet polynomial

$$G(s) := \sum_{n \sim X} \frac{a_n}{n^s}.$$

The following result is a standard mean value estimate that provides bounds for the second moment of a Dirichlet polynomial.

LEMMA 2.1. *One has the estimate*

$$\int_{-T}^{T} |G(it)|^2 \, dt \ll (T + X) \sum_{n \sim X} |a_n|^2.$$

Proof. For a proof, see [8, Theorem 9.1]. □

The mean value estimate from Lemma 2.1 loses some of the information about the sparsity of the Dirichlet polynomial due to the term $X \sum_n |a_n|^2$. In our proof, we will use a Dirichlet polynomial whose coefficients are supported on a subset of $\mathbf{S}(y)$ and are therefore quite sparse. The following variant of the mean value theorem will allow us to exploit the sparsity of the Dirichlet polynomial to some extent. This variant has been used in works on almost primes in short intervals [11, 14].

LEMMA 2.2. *One has the estimate*

$$\int_{-T}^{T} |G(it)|^2 \, dt \ll T \sum_{n \sim X} |a_n|^2 + T \sum_{1 \leq k \leq \frac{2X}{T}} \sum_{n \sim X} |a_n||a_{n+k}|.$$

Proof. For a proof, see [8, Lemma 7.1]. See also [14, Lemma 4]. □

Later, we will also amplify the length of a Dirichlet polynomial by introducing a moment of a prime-supported polynomial. The following lemma will allow us to control the mean value with these augmentations.

LEMMA 2.3. *Let $X \geq P_2 \geq P_1 \geq 2$ be parameters and set $\ell = \lceil \frac{\log P_2}{\log P_1} \rceil$. For*

$$P(s) = \sum_{p \sim P_1} \frac{1}{p^s} \quad \text{and} \quad A(s) = \sum_{n \sim \frac{X}{P_2}} \frac{a_n}{n^s},$$

we have

$$\int_{-T}^{T} |P(1+it)^{\ell} A(1+it)|^2 \, dt \ll \left(\frac{T}{X} + 2^{\ell} P_1 \right) (\ell+1)!^2 \max_n |a_n|^2.$$

Proof. This follows immediately from [10, Lemma 13]. □

A set $\mathcal{T} \subseteq \mathbb{R}$ is said to be *well-spaced* if for all $t, u \in \mathcal{T}$ with $t \neq u$, we have $|t-u| \geq 1$. At some point in our argument, we will encounter a prime-supported Dirichlet polynomial that is large on a well-spaced set. The following lemma will allow us to use this information to obtain a power-saving bound for the size of the well-spaced set.

LEMMA 2.4. *Let*

$$P(s) = \sum_{p \sim P} \frac{a_p}{p^s} \quad \text{with } |a_p| \leq 1.$$

Let $\mathcal{T} \subset [-T, T]$ be a sequence of well-spaced points such that $|P(1+it)| \geq V^{-1}$ for every $t \in \mathcal{T}$. Then

$$|\mathcal{T}| \ll T^{2\frac{\log V}{\log P}} V^2 \exp\left(2 \frac{\log T}{\log P} \log\log T \right).$$

Proof. For a proof, see [10, Lemma 8]. □

Having produced a relatively small well-spaced set, the following discrete mean value theorem, also known as the Halász–Montgomery inequality, will allow us to upper bound the mean square of a Dirichlet polynomial over such sets.

LEMMA 2.5. *Let $\mathcal{T} \subset [-T, T]$ be a well-spaced set. Then,*

$$\sum_{t \in \mathcal{T}} |G(it)|^2 \ll (X + |\mathcal{T}| T^{\frac{1}{2}})(\log T) \sum_{n \sim X} |a_n|^2.$$

Proof. For a proof, see [8, Theorem 9.6]. □

Finally, we need the following pointwise bound, which is a consequence of Perron's formula and the Vinogradov–Korobov zero-free region. For a complex number s, let $\sigma := \Re(s)$ and $\tau := |\Im(s)| + 2$. Then, for some constant $A_{vk} > 0$, the region

$$\sigma \geq 1 - \frac{A_{vk}}{(\log \tau)^{2/3} (\log\log \tau)^{1/3}}$$

denotes the Vinogradov–Korobov zero-free region. Let

$$\sigma_0 := \frac{A_{vk}}{(\log X)^{2/3} (\log\log X)^{1/3}}.$$

LEMMA 2.6. *For* $P \geq \exp\left((\log X)^{2/3}(\log\log X)^{4/3}\right)$, *let*

$$P(s) = \sum_{p \sim P} \frac{1}{p^s}.$$

Then, for all $|t| \leq X$,

$$|P(1+it)| \ll P^{-\sigma_0} + \frac{(\log X)^3}{1+|t|}.$$

Proof. The result follows from standard techniques used in [5, Section 1.4]. The stronger bound above can be derived using the same methods, instead of the weaker one mentioned in [5, Lemma 1.5]. □

3. SOME RESULTS ABOUT SMOOTH NUMBERS

In this section, we recall results about smooth numbers that will be useful for our analysis. We first state some results concerning the distribution of smooth numbers.

The global distribution of smooth numbers for a wide range of smoothness parameters is governed by the Dickman function $\rho : \mathbb{R}_{\geq 0} \to \mathbb{R}$. It is defined as the continuous solution to the system

$$\rho(u) = 1 \quad \text{for} \quad 0 \leq u \leq 1,$$
$$-u\rho'(u) = \rho(u-1) \quad \text{for} \quad u > 1. \tag{3.1}$$

As mentioned in the introduction, it is well known that $\rho(u) = \frac{1}{u^{u+o(u)}}$ (see, for example, [4, (1.6)]). We will make repeated use of this estimate throughout our analysis without further reference.

In our analysis, we consider y-smooth numbers where some of their prime factors lie in certain dyadic ranges. Counting these numbers will involve controlling the ratio $\rho(u-v)/\rho(u)$ when v is much smaller than u. To accomplish this, we use the following lemma.

LEMMA 3.1. *For* $u > 2$ *and* $|v| \leq u/2$, *we have*

$$\rho(u-v) = \rho(u)\exp\left(v\xi(u) + O\left(\frac{1+v^2}{u}\right)\right),$$

where $\xi(u)$ *satisfies the asymptotic*

$$\xi(u) = \log u + \log\log(u+2) + O\left(\frac{\log\log(u+2)}{\log(u+2)}\right).$$

Proof. See [7, Corollary 2.4 and Lemma 2.2]. □

The following lemma gives us asymptotics for the number of smooth numbers in long intervals and moderately short intervals.

LEMMA 3.2. *Fix* $\varepsilon > 0$. *Let* $X \geq 2$, $y \geq \exp\left((\log\log X)^{5/3+\varepsilon}\right)$, *and write* $u = \frac{\log X}{\log y}$.

(i) We have the estimate

$$\Psi(X, y) = X\rho(u)\left(1 + O_\varepsilon\left(\frac{\log(u+1)}{\log y}\right)\right).$$

(ii) For $Xy^{-\frac{5}{12}} \le h \le X$, we have

$$\Psi\left(X+h, y\right) - \Psi\left(X, y\right) = h\rho(u)\left(1 + O_\varepsilon\left(\frac{\log(u+1)}{\log y}\right)\right).$$

Proof. For (i) and (ii), see [6, Theorem 1, Theorem 3], respectively. □

Next, we state an upper bound for the number of integers $n \in [X, 2X]$ such that both n and $an+b$ lie in $\mathbf{S}(y)$. This will allow us to bound the off-diagonal contribution in Lemma 2.2 when applying it to a Dirichlet polynomial supported on $[X, 2X] \cap \mathbf{S}(y)$.

LEMMA 3.3. *For $\varepsilon > 0$, there exist positive constants $C, \delta > 0$ such that for X large enough, $(\log X)^C \le y \le X$ and $1 \le a, |b| \le X^\delta$, we have*

$$\sum_{n \sim X} \mathbf{1}_{\mathbf{S}(y)}(n)\mathbf{1}_{\mathbf{S}(y)}(an+b) \ll_\varepsilon X\rho(u)^{\varphi-\varepsilon},$$

where $u := \frac{\log X}{\log y}$ and $\varphi := \frac{13}{8}$.

Proof. The result follows from [1, Corollaire 4.2] upon noticing that one can replace the $\frac{3}{5}$ in [1, Théorème 2.1] by $\frac{5}{8}$. See [12, Corollary 7.2]. □

4. DEFINING THE DIRICHLET POLYNOMIAL

Let $X \ge 2$ be large, and let y be the smoothness parameter such that

$$\exp\left(C\frac{\log\log X}{\sigma_0}\right) = \exp\left(\frac{C}{A_{\nu k}}(\log X)^{2/3}(\log\log X)^{4/3}\right) \le y \le X^{\frac{1}{C}}, \tag{4.1}$$

where $C \ge 1$ is a large but fixed constant to be chosen later. Additionally, define

$$u := \frac{\log X}{\log y}. \tag{4.2}$$

Note that

$$C \le u \le \frac{A_{\nu k}(\log X)^{1/3}}{C(\log\log X)^{4/3}} = \frac{\sigma_0 \log X}{C\log\log X}. \tag{4.3}$$

From Equations (4.1) and (4.3), it follows that for arbitrarily large A, we have

$$\left(\frac{\log X}{\rho(u)}\right)^A \ll_A y. \tag{4.4}$$

As discussed in the introduction, we will prove Theorem 1.1 by showing that for a set of weights $\{w_n\}_n$ supported on $\mathbf{S}(y)$, we have

$$\sum_{x \le n \le x+h} w_n > 0 \quad \text{for almost all } x \in [X, 2X],$$

where h is the length of the short interval. To define the weights, we introduce some notation.

Let

$$J := \left\lceil \frac{200u\log u}{\sigma_0 \log y}\right\rceil = \left\lceil \frac{200(\log X)^{2/3}(\log\log X)^{1/3}u\log u}{A_{\nu k}\log y}\right\rceil \le \frac{200}{C}u + 1 \tag{4.5}$$

and

$$v := \frac{J\log\left(\frac{y}{2}\right)}{\log y}. \tag{4.6}$$

REMARK 4.1. *To get a rough idea of the size of the quantities above, consider* $y = \exp\big((\log X)^\tau\big)$ *for some* $\tau \in (\frac{2}{3}, 1)$. *Then, we have* $u = (\log X)^{1-\tau}$ *and*

$$v \approx J \asymp 1 + (\log X)^{5/3-2\tau}(\log\log X)^{4/3}.$$

We will always assume that C is large enough so that $u > 3v$ holds. For $\eta > 0$ small enough to be chosen later, let P_1, P_2, and P_3 be parameters such that

$$\frac{(\log X)^4}{\rho(u-v)^{2/3}} \le P_1 \le \frac{(\log X)^{20}}{\rho(u-v)^{14/3}}, \quad P_2 := \big(P_1(\log X)^J\big)^{\frac{1}{\eta}}, \quad \text{and} \quad P_3 := \frac{y}{2}. \tag{4.7}$$

We will optimize the choice of P_1 later. It will turn out that, in the course of proving Theorems 1.1 and 1.2, we will choose

$$P_1 \approx (\log X)^4 \rho(u-v)^{-3/4},$$

for Theorem 1.1, and

$$P_1 \approx (\log X)^4 \rho(u-v)^{-3/2},$$

for Theorem 1.2.

The parameter P_1 essentially determines the length of the short intervals in the statements of these theorems. For example, the short interval length in Theorem 1.1 is approximately

$$\rho(u-v)^{-5/8}P_1 \approx (\log X)^4 \rho(u-v)^{-11/8}.$$

Define

$$\mathcal{M} := \mathbf{S}(y) \cap \left[\frac{X}{2^{J+5}P_1P_2P_3^J}, \frac{8X}{P_1P_2P_3^J}\right]. \tag{4.8}$$

With the above choice of parameters, we define our weights to be

$$w_n := \sum_{\substack{n = q_1q_2p_1\ldots p_Jm \\ q_1\sim P_1, q_2\sim P_2 \\ P_3 < p_1,\ldots,p_J \le 2P_3 \\ m\in\mathcal{M}}} 1. \tag{4.9}$$

As a consequence of Lemma 4.1 (ii) below, Equations (4.7) and (4.4), we have that

$$P_1 \le P_2 \le P_3 \le y/2.$$

Hence, the above weights are supported on $\mathbf{S}(y) \cap \left[\frac{X}{2^{J+5}}, 2^{J+5}X\right]$.

Now, let $F(s)$ be the Dirichlet polynomial with the above weights as coefficients:

$$F(s) := \sum_n \frac{w_n}{n^s}.$$

It follows immediately from Equation (4.9) that

$$F(s) = P_1(s)P_2(s)P_3(s)^J M(s), \tag{4.10}$$

where

$$P_j(s) := \sum_{p\sim P_j} \frac{1}{p^s} \quad \text{for } j \in \{1,2,3\},$$

and

$$M(s) := \sum_{m\in\mathcal{M}} \frac{1}{m^s}.$$

We will use this notation going forward without further mention.

The presence of multiple prime-supported polynomials can make the expressions notationally cumbersome. In the next lemma, we collect some simple observations that will help simplify the expressions during our analysis.

LEMMA 4.1. *For the choice of parameters above, the following statements hold:*

(i) For any $\varepsilon > 0$, there exists a positive constant $C_0(\varepsilon)$ such that

$$\rho(u) \le \rho(u-v) \ll \rho(u)^{1-\varepsilon},$$

provided that $C \ge C_0(\varepsilon)$.

(ii) Let $A > 0$ be a fixed positive constant. Then, for any $\varepsilon > 0$, there exists a positive constant $C_1(\varepsilon)$ such that

$$(A\log X)^J \ll_A \frac{\log X}{\rho(u-v)^{\varepsilon}},$$

provided that $C \ge C_1(\varepsilon)$.

(iii) Let $A > 0$ and κ be fixed constants. Let τ be such that $|\tau| \le 1$. Then, we have

$$\rho\left(u - \tau v - \kappa\frac{\log(A^J P_1 P_2)}{\log y}\right) \asymp_{A,\kappa,\eta} \rho(u-\tau v).$$

Proof. We begin by proving part (i). The first inequality follows from the fact that ρ is a decreasing function. For the other inequality, observe that since $u > 3v$, by applying Lemma 3.1, we get

$$\rho(u-v) \le \rho(u)\exp\left(2J\log u\right).$$

Now, substituting the upper bound for J from Equation (4.5), and for large enough C in terms of ε, we see that

$$\rho(u-v) \le u^2\rho(u)\exp\left(\frac{400}{C}u\log u\right) \ll \rho(u)^{1-\varepsilon}.$$

For part (ii), observe that by substituting the value of J from Equation (4.5), we obtain

$$(A\log X)^J = \exp\left(J(\log A + \log\log X)\right) \ll_A \log X \exp\left(\frac{200u\log u(\log\log X + \log A)}{\sigma_0 \log y}\right).$$

For large enough X in terms of A, we have $\log\log X + \log A \le 2\log\log X$, so

$$(A\log X)^J \ll_A \log X \exp\left(\frac{400u\log u\log\log X}{\sigma_0\log y}\right).$$

Using the lower bound for y in Equation (4.1), we get the following bound:

$$(A\log X)^J \ll_A \log X \exp\left(\frac{400u\log u}{C}\right).$$

Finally, by using part (i) and choosing C large enough in terms of ε, we obtain

$$(A\log X)^J \ll_A \frac{\log X}{\rho(u)^{\varepsilon/2}} \ll \frac{\log X}{\rho(u-v)^{\varepsilon}}.$$

By part (ii), Equations (4.7) and (4.4), we have

$$A^J P_1 P_2 \ll_A \left(\frac{\rho(u-v)}{\log X}\right)^{20+\frac{100}{\eta}} \le (y/2)^J.$$

Hence, recalling the definition of v and the fact that $u > 3v$, it follows that

$$\frac{\log(A^J P_1 P_2)}{\log y} \le v \le \frac{u-v}{2}.$$

Thus, by Lemma 3.1, we have

$$\rho\left(u-\tau v-\kappa\frac{\log(A^J P_1P_2)}{\log y}\right)=\rho(u-\tau v)\exp\left((1+o(1))\kappa\frac{\log(A^J P_1P_2)}{\log y}\log(u-\tau v)\right).$$

To bound the right-hand side, we note that

$$\frac{\log(A^J P_1P_2)}{\log y}\log(u-\tau v)\ll_A\frac{\log P_2\log u}{\log y}\ll_{A,\eta}\frac{\log\log X\log u+u(\log u)^2}{\log y}.$$

Using Equation (4.3), we see that

$$\frac{\log\log X\log u+u(\log u)^2}{\log y}=\frac{u\log u\log\log X}{\log X}+\frac{u^2(\log u)^2}{\log X}=O(1).$$

Therefore, we conclude that

$$\rho\left(u-\tau v-\kappa\frac{\log(A^J P_1P_2)}{\log y}\right)=\rho(u-\tau v)\exp(O_{A,\kappa,\eta}(1))\asymp_{A,\kappa,\eta}\rho(u-\tau v).$$

□

Part (i) of the above lemma allows us to interchange $\rho(u-v)$ and $\rho(u)$ freely. Part (ii) essentially states that, by sacrificing a factor of $\log X$, we can ignore the $(\log X)^J$ term that arises from multiplicity in the counting, due to the presence of multiple prime-supported polynomials. The loss of the $\log X$ factor is only significant for larger values of y. Finally, part (iii) enables us to express the number of y-smooth numbers of size $\asymp\frac{X}{(AP_3)^J P_1P_2}$ in a convenient form.

From Equation (4.9) and Lemma 4.1 (ii), for large enough C in terms of ε, we conclude that

$$w_n\le\frac{(\log n)^{J+2}}{\log P_1\log P_2(\log P_3)^J}\ll\frac{(2\log X)^{J+2}}{\log P_1\log P_2(\log P_3)^J}\ll\frac{(\log X)^3}{\rho(u-v)^\varepsilon}.\tag{4.11}$$

In the following lemma, we derive a lower bound for the averages of our weights over moderately short intervals.

LEMMA 4.2. *Let $\varepsilon>0$. There exists a positive constant $C(\varepsilon)$ such that for $x\in[X,2X]$ and $2Xy^{-5/12}\le h\le X$,*

$$\frac{1}{h}\sum_{x\le n\le x+h}w_n\gg_\eta\frac{\rho(u-v)}{\log P_1\log P_2(2\log P_3)^J}\gg_\eta\left(\frac{\rho(u-v)}{\log X}\right)^{1+\varepsilon}$$

holds uniformly in x, provided $C\ge C(\varepsilon)$.

Proof. Using Equation (4.9), we have that

$$\frac{1}{h}\sum_{x\le n\le x+h}w_n=\frac{1}{h}\sum_{q_1\sim P_1}\sum_{q_2\sim P_2}\sum_{P_3<p_1,\dots,p_J\le 2P_3}\sum_{\substack{\frac{x}{q_1q_2p_1\cdots p_J}\le m\le\frac{x+h}{q_1q_2p_1\cdots p_J}\\ m\in\mathbf{S}(y)}}1.\tag{4.12}$$

The innermost summation is over smooth numbers in a short interval. From our hypothesis, it follows that

$$\frac{x}{q_1q_2p_1\dots p_J}y^{-5/12}\le\frac{h}{q_1q_2p_1\dots p_J}\le\frac{x}{q_1q_2p_1\dots p_J}.$$

Therefore, we can apply Hildebrand's short interval estimate from Lemma 3.2 (ii). This gives

$$\sum_{\substack{\frac{x}{q_1q_2p_1\cdots p_J}\le m\le\frac{x+h}{q_1q_2p_1\cdots p_J}\\ m\in\mathbf{S}(y)}}1\gg\frac{h}{q_1q_2p_1\dots p_J}\rho\left(\frac{\log\frac{x}{q_1q_2p_1\cdots p_J}}{\log y}\right).\tag{4.13}$$

Combining Equations (4.12) and (4.13), we obtain

$$\frac{1}{h}\sum_{x\le n\le x+h} w_n \gg \sum_{q_1\sim P_1}\sum_{q_2\sim P_2}\sum_{P_3<p_1,\dots,p_J\le 2P_3} \frac{1}{q_1q_2p_1\dots p_J}\rho\left(\frac{\log\frac{x}{q_1q_2p_1\cdots p_J}}{\log y}\right)$$

$$\gg \rho\left(\frac{\log\frac{X}{P_1P_2P_3^J}}{\log y}\right)\frac{1}{\log P_1\log P_2(2\log P_3)^J}.$$

Finally, applying Lemma 4.1 (iii) and (ii), and using Equation (4.7), we conclude that

$$\frac{1}{h}\sum_{x\le n\le x+h} w_n \gg_\eta \frac{\rho(u-v)}{\log P_1\log P_2(2\log P_3)^J}\ge\frac{\rho(u-v)}{\log P_1\log P_2(2\log X)^J}.$$

From Equation (4.7) and Lemma 4.1 (ii), we have

$$P_2\ll_\varepsilon\left(\frac{P_1\log X}{\rho(u-v)^\varepsilon}\right)^{\frac{1}{\eta}}.$$

Substituting the above bound for P_2 and the upper bound for P_1 from Equation (4.7), together with an application of Lemma 4.1, we arrive at the inequality

$$\frac{\rho(u-v)}{\log P_1\log P_2(2\log P_3)^J}\gg_\eta\left(\frac{\rho(u-v)}{\log X}\right)^{1+\varepsilon}. \tag{4.14}$$

□

5. PARSEVAL REDUCTION

In this section, we recall a Parseval-type bound that will help reduce the problem of counting smooth numbers in almost all short intervals to establishing a non-trivial upper bound for the mean value of the Dirichlet polynomial $F(s)$.

LEMMA 5.1. *Suppose $X, T_0>0$ are such that $X\ge 2T_0^3$, and let J be a positive integer. Let $\{a_n\}_n$ be a sequence of complex numbers supported on $\mathcal{I}=\left[\frac{X}{2^{J+5}}, 2^{J+5}X\right]$. For $x\in[X,2X]$ and $2\le h_1\le h_2\le\frac{X}{T_0^3}$, define*

$$S_{h_j}(x):=\frac{1}{h_j}\sum_{x\le n\le x+h_j} a_n \quad \text{for} \quad j=1,2.$$

Finally, let

$$G(s):=\sum_n\frac{a_n}{n^s}.$$

Then, we have the following bound:

$$\frac{1}{X}\int_X^{2X}\left|\frac{1}{h_1}S_{h_1}(x)-\frac{1}{h_2}S_{h_2}(x)\right|^2dx\ll\frac{(J\max_{n\in\mathcal{I}}|a_n|)^2}{T_0}+\int_{T_0}^{\frac{X}{h_1}}|G(1+it)|^2\,dt$$
$$+\max_{T\ge\frac{X}{h_1}}\frac{X}{Th_1}\int_T^{2T}|G(1+it)|^2\,dt.$$

Proof. Notice that

$$|G(1+it)|=\left|\sum_{n\in\mathcal{I}}\frac{a_n}{n^{1+it}}\right|\ll\max_{n\in\mathcal{I}}|a_n|\sum_{n\in\mathcal{I}}\frac{1}{n}\ll J\max_{n\in\mathcal{I}}|a_n|. \tag{5.1}$$

The proof is essentially the same as in [10, Lemma 14], with two key differences. First, we do not specify T_0, and second, instead of the bound $|G(1+it)| = O(1)$, we use the estimate in Equation (5.1). See also [14, Lemma 1]. □

Let $S_{h_j}(x) := \sum_{x \le n \le x+h_j} w_n$ for $j \in \{1,2\}$, where $h_1 \ge 2$ will be chosen later, and $h_2 := Xy^{-3/8}$. By Lemma 5.1, for $T_0 = y^{1/8}$, we have

$$\frac{1}{X}\int_X^{2X} \left| \frac{1}{h_1} S_{h_1}(x) - \frac{1}{h_2} S_{h_2}(x) \right|^2 dx \ll \frac{(J \max_n w_n)^2}{y^{1/8}} + \int_{y^{1/8}}^{X/h_1} |F(1+it)|^2 dt$$
$$+ \max_{T \ge \frac{X}{h_1}} \frac{X}{T h_1} \int_T^{2T} |F(1+it)|^2 dt. \tag{5.2}$$

To apply the Chebyshev-style argument, we need to bound the right-hand side of the above inequality. The main challenge lies in bounding the mean values of $F(s)$. In the next section, we will establish a non-trivial upper bound for the mean value of $F(s)$, and in the section following that, we will apply this result to complete the proof of Theorem 1.1.

6. BOUNDING THE MEAN VALUE

In this section, we will show that $F(s)$ satisfies the following mean value estimate by exploiting the factorization in Equation (4.10).

PROPOSITION 6.1. *Let $\varepsilon > 0$ be small, and let $\varphi := \frac{13}{8}$. Let $F(s)$ be as defined in* Equation (4.10)*. Then, there exist positive constants $X_0(\eta)$ and $C(\varepsilon, \eta)$ such that for all $X \ge X_0(\eta)$, any fixed $C \ge C(\varepsilon, \eta)$, and any $T \ge 1$, the following bound holds:*

$$\int_{y^{1/8}}^{T} |F(1+it)|^2 \, dt \ll_{\varepsilon,\eta} P_1^{-\frac{1}{2}+64\eta} \left(\frac{P_1 T}{X} \rho(u-v)^{1-\varepsilon} + \rho(u-v)^{\varphi-\varepsilon} \right) \frac{(\log X)^2}{(\log P_2 (\log P_3)^J)^2}$$
$$+ \left(\frac{\rho(u-v)}{(\log X)^J} \right)^{25}.$$

Before proving the proposition, we will establish several intermediate lemmas. Throughout this section, $C_1, C_2, \dots$ will denote large positive constants that are independent of the parameters defined in Section 4, specifically $X, y, P_1, P_2, P_3, \varepsilon,$ and η. Additionally, $\sum^{\ddagger}$ indicates that the summation variable runs over powers of 2.

The first lemma will assist us in obtaining an upper bound for the contribution from the range of integration where the Dirichlet polynomial $P_1(s)$ is small in magnitude.

LEMMA 6.2. *Let $\varepsilon > 0$ be small, and define $\varphi := \frac{13}{8}$. Let $A(s) := P_2(s)P_3(s)^J M(s)$. Then, there exists a positive constant $C(\varepsilon)$ such that for any fixed $C \ge C(\varepsilon)$ and any $T \ge 1$, we have the following bound:*

$$\int_{-T}^{T} |A(1+it)|^2 \, dt \ll_{\varepsilon,\eta} \left(\frac{P_1 T}{X} \rho(u-v)^{1-\varepsilon} + \rho(u-v)^{\varphi-\varepsilon} \right) \frac{(\log X)^2}{(\log P_2 (\log P_3)^J)^2}.$$

Proof. From the definition of $A(s)$, it is easy to see that it can be written as $\sum_n \frac{a_n}{n^s}$ for

$$a_n := \sum_{\substack{n = q_2 p_1 \dots p_J m \\ q_2 \sim P_2 \\ P_3 < p_1, \dots, p_J \le 2P_3 \\ m \in \mathcal{M}}} 1 \le \frac{(\log n)^{J+1}}{\log P_2 (\log P_3)^J}. \tag{6.1}$$

Note that $\{a_n\}_n$ is supported on $n \in \left[\frac{X}{2^{J+5} P_1}, \frac{2^{J+4} X}{P_1} \right]$. Let $\mathcal{I} := \left[\frac{X}{2^{J+6} P_1}, \frac{2^{J+5} X}{P_1} \right]$.

We perform dyadic division and then apply the Cauchy–Schwarz inequality to obtain

$$\int_{-T}^{T} |A(1+it)|^2\, dt \ll \int_{-T}^{T} \left| \sum_{M\in\mathcal{I}}^{\ddagger} \sum_{n\sim M} \frac{a_n}{n^{1+it}} \right|^2 dt$$

$$\ll J \sum_{M\in\mathcal{I}}^{\ddagger} \int_{-T}^{T} \left| \sum_{n\sim M} \frac{a_n}{n^{1+it}} \right|^2 dt.$$

Note that the result follows immediately when $T \le X^{1-\varepsilon}$ by trivially bounding the integral, so we may assume $T > X^{1-\varepsilon}$. Now, we are in a position to apply Lemma 2.2 to the integral. By doing so, we get

$$\int_{-T}^{T} |A(1+it)|^2\, dt \ll JT \sum_{M\in\mathcal{I}}^{\ddagger} \frac{1}{M^2} \left(\sum_{n\sim M} |a_n|^2 + \sum_{k\le \frac{2M}{T}} \sum_{n\sim M} |a_n a_{n+k}| \right)$$

$$\ll JT \sum_{M\in\mathcal{I}}^{\ddagger} \frac{1}{M^2} \left(\sum_{n\sim M} |a_n|^2 + \sum_{k\le \frac{2M}{T}} \sum_{\substack{n\sim M\\ n+k\in \mathbf{S}(y)}} |a_n| \right) \max_{M\le n\le 4M} |a_n|. \tag{6.2}$$

To bound the diagonal contribution (that is, the first term of the outer summation above), we observe that by using Equation (6.1), Lemma 3.2 (ii), and Lemma 4.1 (iii), we get

$$\sum_{n\sim M} |a_n| \ll \sum_{q_2\sim P_2} \sum_{P_3<p_1,\dots,p_J\le 2P_3} \sum_{\substack{m\in\mathcal{M}\\ m\sim \frac{M}{q_2p_1\cdots p_J}}} 1$$

$$\ll_{\eta} C_1^J \rho(u-v) \frac{M}{\log P_2 (\log P_3)^J}. \tag{6.3}$$

For the off-diagonal contribution (that is, the remaining term), we first note that by Equation (6.1) and the fact that $\mathcal{M} \subseteq \mathbf{S}(y)$, it follows that

$$\sum_{\substack{n\sim M\\ n+k\in\mathbf{S}(y)}} |a_n| \ll \sum_{q_2\sim P_2} \sum_{P_3<p_1,\dots,p_J\le 2P_3} \sum_{\substack{m\sim \frac{M}{q_2p_1\cdots p_J}\\ m(q_2p_1\dots p_J m+k)\in\mathbf{S}(y)}} 1.$$

Note that $|k| \le X^{\varepsilon}$. Additionally, from Equations (4.7) and (4.5), it follows that $P_2(2P_3)^J \le y^{J+1} \le y^2 \exp\left(200\frac{u\log u}{\sigma_0}\right) \le X^{\frac{202}{C}}$. Hence, for large enough C in terms of ε, we may apply Lemma 3.3 to the innermost summation and simplify using Lemma 4.1 (iii) to obtain

$$\sum_{\substack{n\sim M\\ n+k\in\mathbf{S}(y)}} |a_n| \ll_{\varepsilon,\eta} C_2^J \rho(u-v)^{\varphi-\frac{\varepsilon}{4}} \frac{M}{\log P_2(\log P_3)^J}. \tag{6.4}$$

Substituting the bounds in Equations (6.1), (6.3) and (6.4) into Equation (6.2), we obtain

$$\int_{-T}^{T} |A(1+it)|^2\, dt \ll_{\varepsilon,\eta} \left(\frac{P_1T}{X}\rho(u-v) + \rho(u-v)^{\varphi-\frac{\varepsilon}{4}} \right) \frac{(C_3\log X)^{J+1}}{(\log P_2(\log P_3)^J)^2}.$$

Finally, applying Lemma 4.1 (ii) by choosing C even larger in terms of ε if needed, we obtain the desired bound

$$\int_{-T}^{T} |A(1+it)|^2\, dt \ll_{\varepsilon,\eta} \left(\frac{P_1T}{X}\rho(u-v)^{1-\frac{\varepsilon}{4}} + \rho(u-v)^{\varphi-\frac{\varepsilon}{2}} \right) \frac{(\log X)^2}{(\log P_2(\log P_3)^J)^2}$$

$$\ll_{\varepsilon,\eta} \left(\frac{P_1T}{X}\rho(u-v)^{1-\varepsilon} + \rho(u-v)^{\varphi-\varepsilon} \right) \frac{(\log X)^2}{(\log P_2(\log P_3)^J)^2}. \qquad \square$$

The next lemma provides an upper bound for the contribution from the region of integration where $P_2(s)$ is small, but $P_1(s)$ is not too small. The key idea is to amplify the length of the Dirichlet polynomial by introducing a moment of $P_1(s)$.

LEMMA 6.3. *Let $A(s) := P_1(s)P_3(s)^J M(s)$. Additionally, for $\alpha > 0$ and $1 \le T \le X$, let $\mathcal{T} \subseteq [-T, T]$ be a set such that $|P_1(1+it)| \ge P_1^{-\alpha}$ for all $t \in \mathcal{T}$. Then, for $\ell := \left\lceil \frac{\log P_2}{\log P_1} \right\rceil$, we have the following bound:*

$$\int_{\mathcal{T}} |A(1+it)|^2\, dt \ll P_1^{2\alpha\ell} \left(\frac{T}{X} + P_1\right) \left(\frac{(C_4 \log X)^{J+1} \log P_2}{\log P_1 (\log P_3)^J}\right)^2 \times \exp\left(2\frac{\log P_2}{\log P_1} \log\log P_2\right).$$

Proof. We begin by performing similar manipulations as in the proof of the previous lemma. We can write $A(s) = \sum_n \frac{b_n}{n^s}$ for

$$b_n := \sum_{\substack{n = q_1 p_1 \dots p_J m \\ q_1 \sim P_1 \\ P_3 < p_1, \dots, p_J \le 2P_3 \\ m \in \mathcal{M}}} 1 \le \frac{(\log n)^{J+1}}{\log P_1 (\log P_3)^J}. \tag{6.5}$$

Note that $\{b_n\}_n$ is supported on $n \in \left[\frac{X}{2^{J+5}P_2}, \frac{2^{J+4}X}{P_2}\right]$. Let $\mathcal{I} := \left[\frac{X}{2^{J+6}}, 2^{J+5}X\right]$.

Using dyadic division and the Cauchy–Schwarz inequality, we obtain

$$\int_{\mathcal{T}} |A(1+it)|^2\, dt = \int_{\mathcal{T}} \left| \sum_{M \in \mathcal{I}}^{\ddagger} \sum_{n \sim \frac{M}{P_2}} \frac{b_n}{n^{1+it}} \right|^2 dt \ll J \sum_{M \in \mathcal{I}}^{\ddagger} \int_{\mathcal{T}} \left| \sum_{n \sim \frac{M}{P_2}} \frac{b_n}{n^{1+it}} \right|^2 dt.$$

Now we use the lower bound on $P_1(s)$ to increase the length of the Dirichlet polynomial inside the integral. Observe that from $|P_1(1+it)| > P_1^{-\alpha}$, it follows that

$$|P_1(1+it)P_1^{\alpha}|^{2\ell} \ge 1 \quad \text{for all} \quad \ell \in \mathbb{N}.$$

In particular, for $\ell = \left\lceil \frac{\log P_2}{\log P_1} \right\rceil$, we have

$$\begin{aligned} \int_{\mathcal{T}} |A(1+it)|^2\, dt &\ll J P_1^{2\alpha\ell} \sum_{M \in \mathcal{I}}^{\ddagger} \int_{\mathcal{T}} \left| P_1(1+it)^{\ell} \sum_{n \sim \frac{M}{P_2}} \frac{b_n}{n^{1+it}} \right|^2 dt \\ &\le J P_1^{2\alpha\ell} \sum_{M \in \mathcal{I}}^{\ddagger} \int_{-T}^{T} \left| P_1(1+it)^{\ell} \sum_{n \sim \frac{M}{P_2}} \frac{b_n}{n^{1+it}} \right|^2 dt. \end{aligned}$$

Next, applying Lemma 2.3 followed by the bound in Equation (6.5), we get

$$\begin{aligned} \int_{\mathcal{T}} |A(1+it)|^2\, dt &\ll J P_1^{2\alpha\ell} \sum_{M \in \mathcal{I}}^{\ddagger} \left(\frac{T}{M} + 2^{\ell} P_1\right) (\ell+1)!^2 \max_{n \sim M} |b_n|^2 \\ &\ll P_1^{2\alpha\ell} \left(\frac{T}{X} + 2^{\ell} P_1\right) (\ell+1)!^2 \left(\frac{(C_4 \log X)^{J+1}}{\log P_1 (\log P_3)^J}\right)^2. \end{aligned} \tag{6.6}$$

To simplify the above expression, we use Stirling's approximation, which gives $2^{\ell}(\ell+1)!^2 \ll \exp(2\ell \log \ell)$. Furthermore, by the mean value theorem

$$(x+1)\log(x+1) \le x \log x + \log x + 2 \quad \text{for large enough } x.$$

So it follows that

$$\ell \log \ell \le \frac{\log P_2}{\log P_1} \log \frac{\log P_2}{\log P_1} + \log \frac{\log P_2}{\log P_1} + 2.$$

Thus, we obtain the bound

$$2^{\ell}(\ell+1)! \ll (\log P_2)^2 \exp\left(2\frac{\log P_2}{\log P_1} \log\log P_2\right).$$

Substituting this bound into Equation (6.6), we get the desired estimate:

$$\int_{\mathcal{T}} |A(1+it)|^2\, dt \ll P_1^{2\alpha\ell}\left(\frac{T}{X}+P_1\right)\left(\frac{(C_4\log X)^{J+1}\log P_2}{\log P_1(\log P_3)^J}\right)^2 \times \exp\left(2\frac{\log P_2}{\log P_1}\log\log P_2\right). \qquad \square$$

Finally, we will need the following lemma to upper bound the contribution from the remaining range of integration. In this range, we bound the polynomial $P_3(s)$ using the pointwise bound provided by Lemma 2.6.

LEMMA 6.4. *Let $A(s) := P_1(s)P_2(s)M(s)$. For $T \geq 1$, let $\mathcal{T} \subseteq [-T, T]$ be a set. Then there exists a well-spaced set $\mathcal{U} \subseteq \mathcal{T}$ such that*

$$\int_{\mathcal{T}} |A(1+it)|^2\, dt \ll_\eta J^2 \log T\left(1+|\mathcal{U}|\sqrt{T}\frac{(2P_3)^J}{X}\right)\left(\frac{\log X}{\log P_1\log P_2}\right)^2 \rho(u-v).$$

Proof. We can write $A(s) = \sum_n \frac{c_n}{n^s}$ for

$$c_n := \sum_{\substack{n=q_1q_2m\\ q_j\sim P_j\\ m\in\mathcal{M}}} 1 \leq \frac{(\log n)^2}{\log P_1\log P_2}. \tag{6.7}$$

Note that the sequence $\{c_n\}_n$ is supported on $n \in \left[\frac{X}{2^{J+5}P_3^J}, \frac{32X}{P_3^J}\right]$.

We can now find a well-spaced set $\mathcal{U} \subseteq \mathcal{T}$ such that

$$\int_{\mathcal{T}} |A(1+it)|^2\, dt \ll \sum_{t\in\mathcal{U}} |A(1+it)|^2.$$

Let $\mathcal{I} := \left[\frac{X}{2^{J+6}P_3^J}, \frac{64X}{P_3^J}\right]$. Using dyadic division and the Cauchy–Schwarz inequality, we have

$$\int_{\mathcal{T}} |A(1+it)|^2\, dt \ll \sum_{t\in\mathcal{U}}\left|\sum_{M\in\mathcal{I}}^{\ddagger}\sum_{n\sim M}\frac{c_n}{n^{1+it}}\right|^2$$

$$\ll J\sum_{M\in\mathcal{I}}^{\ddagger}\sum_{t\in\mathcal{U}}\left|\sum_{n\sim M}\frac{c_n}{n^{1+it}}\right|^2.$$

Applying Lemma 2.5, we get

$$\int_{\mathcal{T}} |A(1+it)|^2\, dt \ll J\log T\sum_{M\in\mathcal{I}}^{\ddagger}\left(M+|\mathcal{U}|\sqrt{T}\right)\sum_{n\sim M}\left|\frac{c_n}{n}\right|^2$$

$$\ll J\log T\sum_{M\in\mathcal{I}}^{\ddagger}\left(M+|\mathcal{U}|\sqrt{T}\right)\frac{\max_{n\sim M} c_n}{M^2}\sum_{n\sim M} c_n.$$

Additionally, using Lemmas 3.2 (i) and 4.1 (iii), we obtain

$$\sum_{n\sim M} c_n = \sum_{q_1\sim P_1}\sum_{q_2\sim P_2}\sum_{\substack{m\in\mathcal{M}\\ m\sim\frac{M}{q_1q_2}}} 1 \ll \sum_{q_1\sim P_1}\sum_{q_2\sim P_2}\frac{M}{q_1q_2}\rho\left(\frac{\log\frac{M}{q_1q_2}}{\log y}\right) \ll_\eta \frac{M}{\log P_1\log P_2}\rho(u-v).$$

Finally, combining these estimates with Equation (6.7), we get

$$\int_{\mathcal{T}} |A(1+it)|^2\, dt \ll_\eta J^2\log T\left(1+|\mathcal{U}|\sqrt{T}\frac{(2P_3)^J}{X}\right)\left(\frac{\log X}{\log P_1\log P_2}\right)^2\rho(u-v). \qquad \square$$

Now we turn to the proof of Proposition 6.1.

Proof of Proposition 6.1 By Lemma 2.1, we have the bound

$$\int_{-T}^{T} |F(1+it)|^2\, dt \ll (T+X)\sum_n\left(\frac{w_n}{n}\right)^2.$$

Write $\mathcal{I} := \left[\frac{X}{2^{J+6}}, 2^{J+5}X\right]$. From Equations (4.9) and (4.11), we obtain

$$\sum_{n}\left(\frac{w_n}{n}\right)^2 \ll \sum_{M\in\mathcal{I}}^{\ddagger}\frac{1}{M^2}\left(\sum_{q_1\sim P_1}\sum_{q_2\sim P_2}\sum_{P_3<p_1,\dots,p_J\le 2P_3}\sum_{\substack{m\in\mathcal{M}\\ m\sim\frac{M}{q_1q_2p_1\cdots p_J}}}1\right)\frac{(2\log X)^{J+2}}{\log P_1\log P_2(\log P_3)^J}.$$

Using Lemmas 3.2 (ii) and 4.1 (ii) and (iii), we conclude that

$$\sum_{n}\left(\frac{w_n}{n}\right)^2 \ll_\eta \frac{\rho(u-v)^{1-\varepsilon}(\log X)^3}{X(\log P_1\log P_2(\log P_3)^J)^2}.$$

Hence, by combining the results, we get

$$\int_{-T}^{T}|F(1+it)|^2\,dt \ll_\eta \left(\frac{T}{X}+1\right)\frac{\rho(u-v)^{1-\varepsilon}(\log X)^3}{(\log P_1\log P_2(\log P_3)^J)^2}.$$

Using the lower bound on P_1 from Equation (4.7), we observe that this bound is sufficient when $T > X$. Therefore, we may assume that $T \le X$ going forward.

Set $\alpha_1 = \frac{1}{4} - 4\beta$, $\alpha_2 = \frac{1}{4} - 2\beta$ for $\beta = 8\eta$, and write the interval $[y^{1/8}, T]$ as a union of three sets:

$$\mathcal{T}_1 = \{t\in[y^{1/8},T] : |P_1(1+it)| \le P_1^{-\alpha_1}\},$$
$$\mathcal{T}_2 = \{t\in[y^{1/8},T] : |P_2(1+it)| \le P_2^{-\alpha_2}\}\setminus\mathcal{T}_1,$$
$$\mathcal{T}_3 = [y^{1/8},T]\setminus(\mathcal{T}_1\cup\mathcal{T}_2).$$

Thus, we can write the integral as

$$\int_{y^{1/8}}^{T} = \int_{\mathcal{T}_1} + \int_{\mathcal{T}_2} + \int_{\mathcal{T}_3}.$$

We will now show that the contribution from each of the $\mathcal{T}_j$ is bounded by the desired quantity. □

6.1. Contribution from $\mathcal{T}_1$

Let $A_1(s) = P_2(s)P_3(s)^J M(s)$. As $|P_1(1+it)| \le P_1^{-\alpha_1}$ for $t\in\mathcal{T}_1$, we have

$$\int_{\mathcal{T}_1}|F(1+it)|^2\,dt \ll P_1^{-2\alpha_1}\int_{-T}^{T}|A_1(1+it)|^2\,dt.$$

Applying Lemma 6.2 to the integral on the right-hand side, we obtain

$$\int_{\mathcal{T}_1}|F(1+it)|^2\,dt \ll_{\varepsilon,\eta} P_1^{-2\alpha_1}\left(\frac{P_1T}{X}\rho(u-v)^{1-\varepsilon}+\rho(u-v)^{\varphi-\varepsilon}\right)\frac{(\log X)^2}{(\log P_2(\log P_3)^J)^2}.$$

6.2. Contribution from $\mathcal{T}_2$

Let $A_2(s) := P_1(s)P_3(s)^J M(s)$. As $|P_2(1+it)| \le P_2^{-\alpha_2}$ for $t\in\mathcal{T}_2$, we have

$$\int_{\mathcal{T}_2}|F(1+it)|^2\,dt \ll P_2^{-2\alpha_2}\int_{\mathcal{T}_2}|A_2(1+it)|^2\,dt.$$

By Lemma 6.3, for $\ell = \left\lceil\frac{\log P_2}{\log P_1}\right\rceil$, we have

$$\int_{\mathcal{T}_2}|F(1+it)|^2dt \ll P_2^{-2\alpha_2}P_1^{2\alpha_1\ell}\left(\frac{T}{X}+P_1\right)\left(\frac{(C_4\log X)^{J+1}\log P_2}{\log P_1(\log P_3)^J}\right)^2$$
$$\times\exp\left(2\frac{\log P_2}{\log P_1}\log\log P_2\right).$$

Simplifying this expression, we obtain

$$\int_{\mathcal{T}_2} |F(1+it)|^2 dt \ll \left(\frac{T}{X}+1\right)\left(\frac{(C_5\log X)^{J+1}P_1\log P_2}{\log P_1(\log P_3)^J}\right)^2 \times \exp\left(2\left(\alpha_1-\alpha_2+\frac{\log\log P_2}{\log P_1}\right)\log P_2\right).$$

Since $\alpha_2-\alpha_1=2\beta$ and $\frac{\log\log P_2}{\log P_1}\le\beta$ for large enough X depending on η, we can simplify further to obtain

$$\int_{\mathcal{T}_2} |F(1+it)|^2\, dt \ll \left(\frac{T}{X}+1\right)\left(\frac{(C_5\log X)^{J+1}P_1\log P_2}{\log P_1(\log P_3)^J}\right)^2 P_2^{-\beta}.$$

Substituting $P_2=\left(P_1(\log X)^J\right)^{\frac{1}{\eta}}$ and using $P_1\ge\rho(u-v)^{-2/3}$ from Equation (4.7), for small enough η, we conclude that

$$\int_{\mathcal{T}_2} |F(1+it)|^2\, dt \ll_{\varepsilon,\eta} P_1^{-2\alpha_1}\left(\frac{P_1T}{X}\rho(u-v)^{1-\varepsilon}+\rho(u-v)^{\varphi-\varepsilon}\right)\frac{(\log X)^2}{(\log P_2(\log P_3)^J)^2}.$$

6.3. Contribution from $\mathcal{T}_3$

Let $A_3(s)=P_1(s)P_2(s)M(s)$. By Lemma 2.6, $|P_3(1+it)|\ll P_3^{-\sigma_0}$ for $t\ge y^{1/8}$. Hence, we have

$$\int_{\mathcal{T}_3} |F(1+it)|^2\, dt \ll P_3^{-2J\sigma_0}\int_{\mathcal{T}_3}|A_3(1+it)|^2\, dt.$$

Using Lemma 6.4, we can find a well-spaced set $\mathcal{T}\subseteq\mathcal{T}_3$ such that

$$\int_{\mathcal{T}_3} |F(1+it)|^2\, dt \ll_\eta J^2P_3^{-2J\sigma_0}\log T\left(1+|\mathcal{T}|\sqrt{T}\frac{(2P_3)^J}{X}\right)\left(\frac{\log X}{\log P_1\log P_2}\right)^2\rho(u-v).$$

Note that by the definition of P_2 and the fact that $P_1\ge\log X$ from Equation (4.7), it follows that

$$\frac{\log\log T}{\log P_2}\le\frac{\eta\log\log T}{\log P_1}\le\eta.$$

Since $|P_2(1+it)|>P_2^{-\alpha_2}$ and $\frac{\log\log T}{\log P_2}\le\eta$, by Lemma 2.4 we have

$$|\mathcal{T}|\ll T^{2\alpha_2}P_2^{2\alpha_2}\exp\left(2\frac{\log T}{\log P_2}\log\log T\right)\ll T^{\frac12-4\beta+2\eta}P_2^{2\alpha_2}\ll T^{\frac12-30\eta}P_2^{2\alpha_2}.$$

Using upper bound in Equation (4.5), for large enough C in terms of η, we have

$$|\mathcal{T}|\sqrt{T}(2P_3)^J=|\mathcal{T}|\sqrt{T}y^J\le|\mathcal{T}|\sqrt{T}yX^{\frac{400}{C}}\ll X^{1-\delta}$$

for some $\delta=\delta(\eta)>0$. Thus, we conclude that

$$\int_{\mathcal{T}_3} |F(1+it)|^2\, dt \ll_\eta J^2P_3^{-2J\sigma_0}\log T\left(\frac{\log X}{\log P_1\log P_2}\right)^2\rho(u-v).$$

Finally, observe that for large enough X, we have

$$100u\log u<\frac12 J\sigma_0\log y \quad\text{and}\quad 100J\log\log X<\frac12 J\sigma_0\log y.$$

Thus, we obtain

$$100(u\log u+J\log\log X)<J\sigma_0\log y.$$

Since $P_3=\frac{y}{2}$, using the above inequality, we get

$$P_3^{-2J\sigma_0}\le\exp\left(-J\sigma_0\log y\right)<\left(\frac{\rho(u-v)}{(\log X)^J}\right)^{50}.$$

Therefore, we conclude that

$$\int_{\mathcal{T}_3} |F(1+it)|^2\, dt \ll_\eta \left(\frac{\rho(u-v)}{(\log X)^J}\right)^{25}.$$

This completes the proof.

REMARK 6.1. *It might be possible to cut down or even remove the extra* $\log X$ *factors from the counting argument in* Equation (6.1) *by using a variant of* [10, *Lemma 12*]. *This would require controlling how the prime factors of a y-smooth number are distributed in a given interval. In fact, Goudout achieved this when*

$$y \geq \exp\Big((\log X)^{5/6+o(1)}\Big)$$

(see [3, *Lemme* 23]). *We do not take this approach here since our main focus is on small values of* y.

7. FINISHING THE PROOF OF THEOREM 1.1

In this section, we complete the proof of Theorem 1.1.

Proof of Theorem 1.1 Set $\eta := \frac{\varepsilon}{64}$ and $\varphi := \frac{13}{8}$. We begin by bounding the right-hand side of the Equation (5.2). For the first term, it follows from Equation (4.11) that

$$\frac{(J \max_n w_n)^2}{y^{1/8}} \ll_\varepsilon \frac{J^2(\log X)^6}{\rho(u-v)^{2\varepsilon} y^{1/8}}.$$

Using Equation (4.4) and Lemma 4.1 (ii), we can bound the right-hand side above by

$$\ll \left(\frac{\rho(u-v)}{\log X}\right)^{\varepsilon} \left(\frac{\rho(u-v)}{\log P_2 (\log P_3)^J}\right)^2. \tag{7.1}$$

Using Proposition 6.1 and setting $h_1 = P_1 \rho(u-v)^{1-\varphi}$, the remaining terms in Equation (5.2) are

$$\ll_\varepsilon P_1^{-\frac{1}{2}+\varepsilon} \rho(u-v)^{\varphi-\varepsilon} \left(\frac{\log X}{\log P_2 (\log P_3)^J}\right)^2 + \left(\frac{\rho(u-v)}{(\log X)^J}\right)^{25}. \tag{7.2}$$

Next, if we set

$$P_1 := \left(\frac{(\log X)^{2+\varepsilon}}{\rho(u-v)^{2-\varphi+2\varepsilon}}\right)^{\frac{2}{1-2\varepsilon}},$$

we obtain the following bound for Equation (7.2):

$$\ll_\varepsilon \left(\frac{\rho(u-v)}{\log X}\right)^{\varepsilon} \left(\frac{\rho(u-v)}{\log P_2 (\log P_3)^J}\right)^2.$$

Combining these results, we conclude that

$$\frac{1}{X}\int_X^{2X} \left|\frac{1}{h_1} S_{h_1}(x) - \frac{1}{h_2} S_{h_2}(x)\right|^2 dx \ll_\varepsilon \left(\frac{\rho(u-v)}{\log X}\right)^{\varepsilon} \left(\frac{\rho(u-v)}{\log P_2 (\log P_3)^J}\right)^2.$$

Now, let $\mathcal{E}$ be the set of $x \in [X, 2X]$ such that

$$\left|\frac{1}{h_1} S_{h_1}(x) - \frac{1}{h_2} S_{h_2}(x)\right| \geq \frac{\rho(u-v)}{(\log X)^{\varepsilon/4} \log P_1 \log P_2 (2\log P_3)^J},$$

then

$$\frac{|\mathcal{E}|}{X} \left(\frac{\rho(u-v)}{(\log X)^{\varepsilon/4} \log P_1 \log P_2 (2\log P_3)^J}\right)^2 \ll_\varepsilon \left(\frac{\rho(u-v)}{\log X}\right)^{\varepsilon} \left(\frac{\rho(u-v)}{\log P_2 (\log P_3)^J}\right)^2.$$

By rearranging the above expression and noting that $\log P_1 \ll \left(\frac{\log X}{\rho(u-v)}\right)^{\varepsilon/8}$ and $2^J \ll \rho(u-v)^{-\varepsilon/4}$, we obtain

$$|\mathcal{E}| \ll_\varepsilon X \left(\frac{\rho(u-v)}{\log X}\right)^{\varepsilon/4} = o(X).$$

Finally, for $x \in [X, 2X] \setminus \mathcal{E}$, Lemma 4.2 implies that

$$\frac{1}{h_1} S_{h_1}(x) \gg \frac{\rho(u-v)}{\log P_1 \log P_2 (2\log P_3)^J} \gg_\varepsilon \left(\frac{\rho(u-v)}{\log X}\right)^{1+\varepsilon}.$$

Here, the implicit constant is independent of x. From this, it follows that the interval $[x, x+h]$ contains a y-smooth number for $h \geq h_1$ whenever $x \in [X, 2X] \setminus \mathcal{E}$. In particular, we may take the length of the interval as specified in the hypothesis of Theorem 1.1 by using Lemma 4.1 (i) and the fact that ε is arbitrary. Hence, the proof is complete. □

8. EXISTENCE IN ALL INTERVALS

In this section, we prove Theorem 1.2. The proof follows a similar approach to that of Theorem 1.1, but we adjust the choice of weights to optimize the log-factors.

Let $x \geq 2$ be large enough, $\varphi := \frac{13}{8}$, and let C be a sufficiently large but fixed positive constant. The precise choice of C will become clear in the course of the argument.

For

$$\exp\left(C(\log x)^{2/3}(\log\log x)^{4/3}\right) \leq y \leq x^{\frac{1}{C}},$$

define

$$u_0 := \frac{\log x}{2\log y},$$

and set

$$P_1 := \left(\frac{(\log x)^{2+2\varepsilon}}{\rho(u_0)^{4-2\varphi+4\varepsilon}}\right)^{\frac{2}{1-4\varepsilon}}.$$

Additionally, we define

$$X := \sqrt{xP_1}.$$

The parameters u, J, v, P_2, P_3 and $\mathcal{M}$ are given by the respective definitions in Equations (4.2) and (4.5)–(4.8).

Note that for an appropriately chosen C, the value of y satisfies Equation (4.1), and P_1 meets the condition in Equation (4.7) by applying Lemma 4.1.

Furthermore, define

$$R := \sqrt{\frac{x}{P_1}},$$

and set

$$\mathcal{R} := (R, 2R] \cap \mathbf{S}(y).$$

We define our new set of weights as

$$\widetilde{w}_n := \sum_{\substack{n = q_1 q_2 p_1 \ldots p_J m r \\ q_1 \sim P_1,\, q_2 \sim P_2 \\ P_3 < p_1, \ldots, p_J \leq 2P_3 \\ m \in \mathcal{M},\, r \in \mathcal{R}}} 1. \tag{8.1}$$

The above weights are supported on $\mathbf{S}(y) \cap \left[\frac{x}{2^{J+5}}, 2^{J+6}x\right]$.

Let $G(s)$ be a Dirichlet polynomial with coefficients $\widetilde{w}_n$, defined as

$$G(s) := \sum_n \frac{\widetilde{w}_n}{n^s}. \tag{8.2}$$

It immediately follows that

$$G(s) = F(s)R(s) = P_1(s)P_2(s)P_3(s)^J M(s)R(s),$$

where $F(s)$, $P_1(s)$, $P_2(s)$, $P_3(s)$ and $M(s)$ are as defined in Section 4, and

$$R(s) := \sum_{r \in \mathcal{R}} \frac{1}{r^s}.$$

Similarly to Lemma 4.2, we establish a lower bound for a moderately long average of the weights $\widetilde{w}_n$.

LEMMA 8.1. *Let $\varepsilon > 0$. There exists a positive constant $C(\varepsilon)$ such that for $x \geq 2$ and $xy^{-5/12} \leq h \leq x$,*

$$\frac{1}{h}\sum_{x\leq n\leq x+h} \widetilde{w}_n \gg_\varepsilon \frac{\rho(u_0)^{2+\varepsilon/2}}{\log P_1 \log P_2 (2\log P_3)^J}$$

holds, provided that $C \geq C(\varepsilon)$.

Proof. Using Equation (8.1), we have

$$\frac{1}{h}\sum_{x\leq n\leq x+h} \widetilde{w}_n = \frac{1}{h}\sum_{q_1\sim P_1}\sum_{q_2\sim P_2}\sum_{P_3<p_1,\ldots,p_J\leq 2P_3}\sum_{r\in\mathcal{R}}\sum_{\substack{\frac{x}{q_1q_2p_1\cdots p_Jr}\leq m\leq \frac{x+h}{q_1q_2p_1\cdots p_Jr}\\ m\in \mathbf{S}(y)}} 1. \tag{8.3}$$

The innermost summation is over smooth numbers in a short interval. From our hypothesis, it follows that

$$\frac{x}{q_1q_2p_1\ldots p_Jr}y^{-5/12} \leq \frac{h}{q_1q_2p_1\ldots p_Jr} \leq \frac{x}{q_1q_2p_1\ldots p_Jr}.$$

Therefore, we can apply Hildebrand's short interval estimate from Lemma 3.2 (ii). This gives

$$\sum_{\substack{\frac{x}{q_1q_2p_1\cdots p_Jr}\leq m\leq \frac{x+h}{q_1q_2p_1\cdots p_Jr}\\ m\in \mathbf{S}(y)}} 1 \gg \frac{h}{q_1q_2p_1\ldots p_Jr}\rho\left(\frac{\log \frac{x}{q_1q_2p_1\cdots p_Jr}}{\log y}\right). \tag{8.4}$$

Combining Equations (8.3) and (8.4), we obtain

$$\frac{1}{h}\sum_{x\leq n\leq x+h} \widetilde{w}_n \gg \sum_{q_1\sim P_1}\sum_{q_2\sim P_2}\sum_{P_3<p_1,\ldots,p_J\leq 2P_3}\sum_{r\in\mathcal{R}}\frac{1}{q_1q_2p_1\ldots p_Jr}\rho\left(\frac{\log\frac{x}{q_1q_2p_1\cdots p_Jr}}{\log y}\right)$$
$$\gg \rho\left(\frac{\log\frac{x}{P_1P_2P_3^JR}}{\log y}\right)\rho\left(\frac{\log(2R)}{\log y}\right)\frac{1}{\log P_1\log P_2(2\log P_3)^J}.$$

Finally, applying Lemma 4.1 (i) and (iii), we conclude that

$$\frac{1}{h}\sum_{x\leq n\leq x+h} \widetilde{w}_n \gg_\varepsilon \frac{\rho(u_0)^{2+\varepsilon/2}}{\log P_1 \log P_2 (2\log P_3)^J}.$$

□

We now prove the following mean value estimate for $G(s)$, which is a quick consequence of Proposition 6.1.

LEMMA 8.2. *Let $\varepsilon > 0$ be small, and let $\varphi := \frac{13}{8}$. Let $G(s)$ be as in* Equation (8.2)*. Then, there exist positive constants $x_0(\eta)$ and $C(\varepsilon,\eta)$ such that for all $x \geq x_0(\eta)$, any fixed $C \geq C(\varepsilon,\eta)$, and any $T \geq 1$, the following bound holds:*

$$\int_{y^{1/8}}^{T} |G(1+it)|\,dt$$
$$\ll_{\varepsilon,\eta} P_1^{-\frac{1}{4}+32\eta}\left(\sqrt{\frac{P_1}{x}}T\rho(u_0)^{1-\varepsilon}+\rho(u_0)^{\varphi-\varepsilon}\right)\frac{\log X}{\log P_2(\log P_3)^J}$$
$$+\left(\frac{\rho(u_0)}{(\log X)^J}\right)^{10}\left(\sqrt{\frac{P_1}{x}}T\rho(u_0)+\rho(u_0)^{\varphi-\varepsilon}\right)^{1/2}.$$

Proof. Using the Cauchy–Schwarz inequality, we have

$$\int_{y^{1/8}}^{T} |G(1+it)|\,dt \le \left(\int_{y^{1/8}}^{T} |F(1+it)|^2\,dt\right)^{1/2} \left(\int_{-T}^{T} |R(1+it)|^2\,dt\right)^{1/2}. \tag{8.5}$$

By Proposition 6.1, we estimate the first integral as follows:

$$\int_{y^{1/8}}^{T} |F(1+it)|^2\,dt \ll P_1^{-\frac{1}{2}+64\eta}\left(\frac{P_1 T}{X}\rho(u-v)^{1-\varepsilon} + \rho(u-v)^{\varphi-\varepsilon}\right)\frac{(\log X)^2}{(\log P_2(\log P_3)^J)^2} + \left(\frac{\rho(u-v)}{(\log X)^J}\right)^{25}. \tag{8.6}$$

For the second integral, we use Lemmas 2.2 and 3.3 to obtain

$$\int_{-T}^{T} |R(1+it)|^2\,dt \ll_\varepsilon \frac{T}{R}\rho\left(\frac{\log R}{\log y}\right) + \rho\left(\frac{\log R}{\log y}\right)^{\varphi-\varepsilon}. \tag{8.7}$$

Combining Equations (8.5)–(8.7), and applying Lemma 4.1 (i) and (iii), we obtain the desired result. □

We will need the following smoothing function in the proof of Theorem 1.2. Let

$$\eta_{\xi,\kappa}(z) = \begin{cases} 1 & \text{if } 1-\kappa \le z \le 1+\kappa, \\ \frac{1+\kappa+\xi-z}{\xi} & \text{if } 1+\kappa \le z \le 1+\xi+\kappa, \\ \frac{z+\kappa+\xi-1}{\xi} & \text{if } 1-\xi-\kappa \le z \le 1-\kappa, \\ 0 & \text{otherwise.} \end{cases}$$

We define $\widetilde{\eta}_{\xi,\kappa}$ as the Mellin transform of $\eta_{\xi,\kappa}$. That is,

$$\widetilde{\eta}_{\xi,\kappa}(s) := \int_0^\infty t^{s-1}\eta_{\xi,\kappa}(t)\,dt.$$

By Mellin inversion, we have the following representation:

$$\eta_{\xi,\kappa}(z) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} z^{-s}\cdot\widetilde{\eta}_{\xi,\kappa}(s)\,ds. \tag{8.8}$$

Next, observe that

$$\begin{aligned} s\cdot\widetilde{\eta}_{\xi,\kappa}(s) &= -\int_0^\infty t^s\,d\eta_{\xi,\kappa}(t) \\ &= -\int_{1-\kappa-\xi}^{1-\kappa}\frac{t^s}{\xi}\,dt + \int_{1+\kappa}^{1+\kappa+\xi}\frac{t^s}{\xi}\,dt \\ &= \frac{(1+\xi+\kappa)^{s+1}-(1+\kappa)^{s+1}}{\xi(s+1)} - \frac{(1-\kappa)^{s+1}-(1-\xi-\kappa)^{s+1}}{\xi(s+1)}. \end{aligned}$$

Proof of Theorem 1.2 Set $\eta := \frac{\varepsilon}{32}$. Define h_1 and h_2 as follows:

$$h_1 := h_0\sqrt{x}, \quad \text{where } h_0 := \sqrt{P_1}\rho(u_0)^{1-\varphi},$$
$$h_2 := xy^{-3/8}.$$

For $j = 1, 2$, define $\kappa_j := \frac{h_j}{x}$, $\xi_j := \frac{h_j}{x}$, and set $\eta_j := \eta_{\xi_j,\kappa_j}$. Finally, let

$$S_j := \sum_n \widetilde{w}_n\eta_j\left(\frac{n}{x}\right) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} x^s\cdot G(s)\widetilde{\eta}_j(s)\,ds. \tag{8.9}$$

We relate the short-interval distribution to the long-interval distribution by showing that

$$\left|\frac{S_1}{h_1} - \frac{S_2}{h_2}\right| \tag{8.10}$$

is small.

Decompose the integral on the right-hand side of Equation (8.9) as

$$\int_{1-i\infty}^{1+i\infty} x^s \cdot G(s)\widetilde{\eta}_j(s)\, ds = U_j + V_j,$$

where

$$U_j := \int_{|t|\le y^{1/8}} x^s \cdot G(s)\widetilde{\eta}_j(s)\, ds,$$

$$V_j := \int_{|t|> y^{1/8}} x^s \cdot G(s)\widetilde{\eta}_j(s)\, ds.$$

By a straightforward modification of the proof of [10, Theorem 4], incorporating Equation (5.1), Lemma 4.1 (iii) and Equation (7.1), we obtain

$$\begin{aligned}\left|\frac{U_1}{h_1} - \frac{U_2}{h_2}\right| &\ll \frac{|G(1)|y^{1/4}h_2}{x}\\ &= \frac{|F(1)R(1)|y^{1/4}h_2}{x}\\ &\ll J\frac{\max_n |w_n|\rho(u_0)}{y^{1/8}}\\ &\ll \left(\frac{\rho(u_0)}{\log X}\right)^{\varepsilon} \frac{\rho(u_0)^2}{\log P_2(\log P_3)^J}.\end{aligned}$$

and

$$\left|\frac{V_1}{h_1} - \frac{V_2}{h_2}\right| \ll \sum_{j=1}^{2}\left(\int_{y^{1/8}}^{x/h_j} |G(1+it)|\, dt + \frac{x}{h_j}\max_{T>\frac{x}{h_j}} \frac{1}{T}\int_T^{2T} |G(1+it)|\, dt\right).$$

For sufficiently large x depending on ε, from Lemma 8.2, we obtain

$$\begin{aligned}\left|\frac{V_1}{h_1} - \frac{V_2}{h_2}\right| &\ll_{\varepsilon} P_1^{-\frac{1}{4}+\varepsilon}\left(\frac{\sqrt{P_1}}{h_0}\rho(u_0)^{1-\varepsilon} + \rho(u_0)^{\varphi-\varepsilon}\right)\frac{\log X}{\log P_2(\log P_3)^J}\\ &\quad + \left(\frac{\rho(u_0)}{(\log X)^J}\right)^{10}\left(\frac{\sqrt{P_1}}{h_0}\rho(u_0)^{1-\varepsilon} + \rho(u_0)^{\varphi-\varepsilon}\right)^{1/2}.\end{aligned}$$

Substituting the expressions for P_1 and h_0, we get

$$\begin{aligned}\left|\frac{V_1}{h_1} - \frac{V_2}{h_2}\right| &\ll_{\varepsilon} P_1^{-\frac{1}{4}+\varepsilon}\rho(u_0)^{\varphi-\varepsilon}\frac{\log X}{\log P_2(\log P_3)^J} + \left(\frac{\rho(u_0)}{(\log X)^J}\right)^{10}\\ &\ll_{\varepsilon} \left(\frac{\rho(u_0)}{\log X}\right)^{\varepsilon}\frac{\rho(u_0)^2}{\log P_2(\log P_3)^J}.\end{aligned}$$

Thus, we conclude that

$$\left|\frac{S_1}{h_1} - \frac{S_2}{h_2}\right| \ll_{\varepsilon} \left(\frac{\rho(u_0)}{\log X}\right)^{\varepsilon}\frac{\rho(u_0)^2}{\log P_2(\log P_3)^J}. \tag{8.11}$$

From this and the definition of η_2, we conclude that

$$\frac{S_1}{h_1} \ge \frac{1}{h_2}\sum_{x-h_2\le n\le x+h_2} \widetilde{w}_n + O_{\varepsilon}\left(\left(\frac{\rho(u_0)}{\log X}\right)^{\varepsilon}\frac{\rho(u_0)^2}{\log P_2(\log P_3)^J}\right). \tag{8.12}$$

By Lemma 8.1, we have

$$\frac{1}{h_2}\sum_{x-h_2\le n\le x+h_2} \widetilde{w}_n \gg_{\varepsilon} \frac{\rho(u_0)^{2+\varepsilon/2}}{\log P_1 \log P_2(2\log P_3)^J}.$$

Substituting this into Equation (8.12) and using the definition of η_1, we obtain

$$\sum_{x-2h_1\le n\le x+2h_1} \widetilde{w}_n \gg_{\varepsilon} h_1\frac{\rho(u_0)^{2+\varepsilon/2}}{\log P_1 \log P_2(2\log P_3)^J}.$$

In particular, this shows that every interval of length at least $4h_1$ contains a y-smooth number. Noting that $\rho(u_0) = \rho\left(\frac{1}{2}\frac{\log x}{\log y}\right)$, we can take the interval length h as specified in the hypothesis of Theorem 1.2, using the fact that ε is arbitrary. Hence, we conclude the result. □

ACKNOWLEDGEMENTS

The author would like to thank his advisor, Kaisa Matomäki, for suggesting the problem and for her guidance throughout the project. Her insights not only led to a clearer exposition but also contributed to improving the quality of both the smoothness parameter and the short interval length. The author would also like to thank the anonymous referee for carefully reading the paper and providing useful suggestions.

FUNDING

The author was supported by the Academy of Finland, Centre of Excellence (Grant No. 346307), and the University of Turku Graduate School Exactus fellowship while working on this project.

REFERENCES

1. R. de la Bretèche and S. Drappeau, Niveau de répartition des polynômes quadratiques et crible majorant pour les entiers friables, *J. Eur. Math. Soc.* **22**(2020), 1577–1624.
2. J. B. Friedlander and A. Granville, Smoothing “smooth” numbers, *Philos. Trans. Roy. Soc. London Ser. A* **345**(1993), 339–347.
3. É. Goudout, Lois locales de la fonction ω dans presque tous les petits intervalles, *Proc. Lon. Math. Soc.* **115**(2017), 599–637.
4. A. Granville, Smooth numbers: computational number theory and beyond, in *Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography*, vol. **44** of MSRI Publications, Cambridge University Press, Cambridge, (2008), 267–323.
5. G. Harman, *Prime-Detecting Sieves*, vol. **33** of London Mathematical Society Monographs Series, Princeton University Press, Princeton, NJ, (2007).
6. A. Hildebrand, On the number of positive integers $\leq x$ and free of prime factors $> y$, *J. Number Theory* **22**(1986), 289–307.
7. A. Hildebrand and G. Tenenbaum, Integers without large prime factors, *J. Théor. Nr. Bordx* **5**(1993), 411–484.
8. H. Iwaniec and E. Kowalski, *Analytic Number Theory* vol. **53** of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, (2004).
9. K. Matomäki, Another note on smooth numbers in short intervals, *Int. J. Number Theory* **12**(2016), 323–340.
10. K. Matomäki and M. Radziwiłł, Multiplicative functions in short intervals, *Ann. Math. (2)* **183**(2016), 1015–1056.
11. K. Matomäki and J. Teräväinen, Almost primes in almost all short intervals II, *Trans. Am. Math. Soc.* **376**(2023), 5433–5459.
12. A. Pascadi, On the exponents of distribution of primes and smooth numbers. Preprint available at http://arxiv.org. as arXiv:2505.00653
13. K. Soundararajan, Smooth numbers in short intervals. Preprint available at: http://arxiv.org. as arXiv:1009.1591v1

14. J. Teräväinen, Almost primes in almost all short intervals, *Math. Proc. Cambridge Philos. Soc.* **161**(2016), 247–281.
15. K. Younis, Asymptotics for smooth numbers in short intervals. Preprint available at: http://arxiv.org/ as arXiv:2409.05761v1

Received: 7 April 2025. **Revised:** 2 February 2026. **Accepted:** 16 March 2026