NOTES ON UPPER AND LOWER BOUNDING $t(N)$

TERENCE TAO

1. Basics

The symbol p will always denote a prime. The primes $2, 3$ will play a special role here and will be referred to as *tiny primes*.

We use $\nu_p(a/b) = \nu_p(a) - \nu_p(b)$ to denote the p-adic valuation of a positive natural number a/b, that is to say the number of times p divides the numerator a, minus the number of times p divides the denominator b. For instance, $\nu_2(32/27) = 5$ and $\nu_3(32/27) = -3$. If one applies a logarithm to the fundamental theorem of arithmetic, one obtains the identity

$$\sum_p \nu_p(r) \log p = \log r \tag{1.1}$$

for any positive rational r.

Asymptotically, it is known that

$$\frac{1}{e} - \frac{O(1)}{\log N} \leq \frac{t(N)}{N} \leq \frac{1}{e} - \frac{c_0 + o(1)}{\log N}$$

where

$$\begin{aligned} c_0 &:= \frac{1}{e} \int_0^1 \left\lfloor \frac{1}{x} \right\rfloor \log\left(ex \left\lceil \frac{1}{ex} \right\rceil \right)\, dx \\ &= \frac{1}{e} \int_1^\infty \lfloor y \rfloor \log \frac{\lceil y/e \rceil}{y/e}\, \frac{dy}{y^2} \\ &= 0.3044190\dots \end{aligned} \tag{1.2}$$

We recall Legendre's formula

$$\nu_p(N!) = \sum_{j=1}^\infty \left\lfloor \frac{N}{p^j} \right\rfloor = \frac{N - s_p(N)}{p-1}. \tag{1.3}$$

To bound the factorial, we have the explicit Stirling approximation [4]

$$N \log N - N + \log\sqrt{2\pi N} + \frac{1}{12N+1} \leq \log N! \leq N \log N - N + \log\sqrt{2\pi N} + \frac{1}{12N}, \tag{1.4}$$

valid for all natural numbers N.

In addition to the usual asymptotic notation, we use $O_{\leq}(X)$ to denote any quantity whose magnitude is bounded by at most X (note the absence of an additional constant factor).

2. Criteria for upper bounding $t(N)$

We have the trivial upper bound $t(N) \leq (N!)^{1/N}$. This can be improved to $t(N) \leq N/e$ for $N \neq 1, 2, 4$, answering a conjecture of Guy and Selfridge [3]; see [6]. This was derived from the following slightly stronger criterion, which asymptotically gives $\frac{t(N)}{N} \leq \frac{1}{e} - \frac{c_0+o(1)}{\log N}$:

Lemma 2.1 (Upper bound criterion)**.** [6, Lemma 2.1] *Suppose that* $1 \leq t \leq N$ *are such that*

$$\sum_{p > \frac{t}{\lfloor \sqrt{t} \rfloor}} \left\lfloor \frac{N}{p} \right\rfloor \log\left(\frac{p}{t}\left\lceil \frac{t}{p} \right\rceil\right) > \log N! - N \log t \tag{2.1}$$

Then $t(N) < t$.

A surprisingly sharp upper bound comes from linear programming.

Lemma 2.2 (Linear programming bound)**.** *Let* N *be an natural number and* $1 \leq t \leq N/2$. *Suppose for each prime* $p \leq N$, *one has a non-negative real number* w_p *which is weakly non-decreasing in* p *(thus* $w_p \leq w_{p'}$ *when* $p \leq p'$*), and such that*

$$\sum_p w_p \nu_p(j) \geq 1 \tag{2.2}$$

for all $t \leq j \leq N$, *and such that*

$$\sum_p w_p \nu_p(N!) < N. \tag{2.3}$$

Then $t(N) < t$.

Proof. We first observe that the bound (2.2) in fact holds for all $j \geq t$, not just for $t \leq j \leq N$. Indeed, if this were not the case, consider the first $j \geq t$ where (2.2) fails. Take a prime p dividing j and replace it by a prime in the interval $[p/2, p)$ which exists by Bertrand's postulate (or remove p entirely, if $p = 2$); this creates a new j' in $[j/2, j)$ which is still at least t. By the weakly decerasing hypothesis on w_p, we have

$$\sum_p w_p \nu_p(j) \geq \sum_p w_p \nu_p(j')$$

and hence by the minimality of j we have

$$\sum_p w_p \nu_p(j) > 1,$$

a contradiction.

Now suppose for contradiction that $t(N) \geq t$, thus we have a factorization $N! = \prod_{j \geq t} j^{m_j}$ for some natural numbers m_j summing to N. Taking p-valuations, we conclude that

$$\sum_{j \geq t} m_j \nu_p(j) \leq \nu_p(N!)$$

for all $p \leq N$. Multiplying by w_p and summing, we conclude from (2.2) that

$$N = \sum_{j \geq t} m_j \leq \sum_p w_p \nu_p(N!),$$

contradicting (2.3). □

This bound is sharp for all $N \leq 600$, with the exception of $N = 155$, where it gives the upper bound $t(155) \leq 46$. A more precise integer program gives $t(155) = 45$.

Remark 2.3. A variant of the linear programming method also gives good lower bound constructions. Specifically, one can use linear programming to find non-negative real numbers m_j for $t \leq j \leq N$ that maximize the quantity $\sum_{t \leq j \leq N} m_j$ subject to the constraints

$$\sum_{t \leq j \leq N} m_j \nu_p(j) \leq \nu_p(N!).$$

The expression $\prod_{t \leq j \leq N} j^{\lfloor m_j \rfloor}$ will then be a subfactorization of $N!$ into $\sum_{t \leq j \leq N} \lfloor m_j \rfloor$ factors j, each of which is at least t. If $\sum_{t \leq j \leq N} \lfloor m_j \rfloor \geq N$, this demonstrates that $t(N) \geq t$. Numerically, this procedure attains the exact value of $t(N)$ for all $N \leq 600$; for instance for $N = 155$, it shows that $t(155) \geq 45$.

2.1. **Asymptotic analysis of upper bound.** We refine the upper bound in [6] slightly.

Proposition 2.4. *For large N, one has*

$$\frac{t(N)}{N} \leq \frac{1}{e} - \frac{c_0}{\log N} + O\left(\frac{1}{\log^2 N}\right).$$

Proof. We apply Lemma 2.1 with

$$t := \frac{1}{e} - \frac{c_0}{\log N} + \frac{C_0}{\log^2 N}$$

with C_0 a large absolute constant to be chosen later. From the Stirling approximation one sees that

$$\log N! - N \log t \geq ec_0 \frac{N}{\log N} + (C_0 - O(1)) \frac{N}{\log^2 N}$$

so it will suffice to establish the upper bound

$$\sum_{p > \frac{t}{\lfloor \sqrt{t} \rfloor}} \left\lfloor \frac{N}{p} \right\rfloor \log\left(\frac{p}{t}\left\lceil \frac{t}{p} \right\rceil\right) \leq ec_0 \frac{N}{\log N} + O\left(\frac{N}{\log^2 N}\right).$$

For N large enough, we have $\frac{t}{\lfloor \sqrt{t} \rfloor} \leq \frac{N}{\log N}$, so it suffices to show that

$$\sum_{\frac{N}{\log N} \leq p \leq N} \left\lfloor \frac{N}{p} \right\rfloor \log\left(\frac{p}{t}\left\lceil \frac{t}{p} \right\rceil\right) \leq ec_0 \frac{N}{\log N} + O\left(\frac{N}{\log^2 N}\right).$$

The summand is a piecewise monotone function of p, with $O(\log N)$ pieces, and bounded in size by $O(N)$. A routine application of the prime number theorem (with classical error term) and summation by parts then allows one to express the left-hand side as

$$\int_{N/\log N}^{N} \left\lfloor \frac{N}{x} \right\rfloor \log\left(\frac{x}{t}\left\lceil \frac{t}{x}\right\rceil\right) \frac{dx}{\log x} + O\left(\frac{N}{\log^2 N}\right)$$

(in fact the error term can be made much stronger than this). We use the approximation

$$\frac{1}{\log x} = \frac{1}{\log N} + O\left(\frac{\log(N/x)}{\log^2 N}\right).$$

To control the error term, we observe from Taylor expansion that

$$\log\left(\frac{x}{t}\left\lceil \frac{t}{x}\right\rceil\right) \ll \frac{\lceil \frac{t}{x}\rceil - \frac{t}{x}}{t/x} \ll \frac{x}{t} \ll \frac{x}{N} \tag{2.4}$$

and the contribution of the error term is

$$\ll \int_{N/\log N}^{N} \frac{N}{x}\frac{x}{N}\frac{\log(N/x)}{\log^2 N}\, dx \ll \frac{N}{\log^2 N}$$

which is acceptable. As for the main term, we can rescale it to

$$\frac{et}{\log N}\int_{N/et\log N}^{N/et} \left\lfloor \frac{N/et}{x} \right\rfloor \log\left(ex\left\lceil \frac{1}{ex}\right\rceil\right) dx.$$

Since $N/et = 1 + O(1/\log N)$, we see that the integrand here is within $O(1/\log N)$ of $\lfloor \frac{1}{x} \rfloor \log\left(ex\lceil \frac{1}{ex}\rceil\right)$ unless $\frac{1}{x}$ is within $O(1/\log N)$ of an integer, which one can calculate to occur on a set of measure zero. A variant of (2.4) shows that both integrandd are bounded by $O(1)$ for all $x \in [0, N/et]$, so by the triangle inequality the above expression can be rewritten as

$$\frac{N}{\log N}\int_0^1 \left\lfloor \frac{1}{x} \right\rfloor \log\left(ex\left\lceil \frac{1}{ex}\right\rceil\right) dx + O\left(\frac{N}{\log^2 N}\right),$$

and the claim follows from (1.2). □

3. A GENERAL FACTORIZATION ALGORITHM

In this section we present and then analyze an algorithm that, when given parameters $1 \leq t \leq N$, will attempt to construct a factorization $N! = \prod \mathcal{B}$ of $N!$ by a finite multiset $\mathcal{B}$ of N elements that are all at least t. The algorithm will not always succeed, but when it does, it will certify that $t(N) \geq t$.

3.1. **Notational preliminaries.** We begin with some key definitions.

Let $\mathcal{B} = \{b_1, \dots, b_M\}$ be a finite multiset of natural numbers (thus each natural number may appear in $\mathcal{B}$ multiple times); the ordering of elements in the multiset will not be of relevance to us. The *cardinality* $|\mathcal{B}| = M$ of the multiset is the number of elements counting multiplicity; for example,

$$|\{2, 2, 3\}| = 3.$$

The *product* $\prod \mathcal{B}$ of the finite multiset is defined by $\prod \mathcal{B} := \prod_{b \in \mathcal{B}} b$, where we count for multiplicity; for example

$$\prod\{2,2,3\} = 12.$$

The tuple $\mathcal{B}$ is a *factorization* of a natural number M if $\mathcal{B} = M$, and a *subfactorization* if $\mathcal{B}|M$. For example, $\{2,2,3\}$ is a factorization of 12 and a subfactorization of 24.

By the fundamental theorem of arithmetic (or (1.1)), we see that a finite multiset $\mathcal{B}$ is a factorization of M if and only if

$$\nu_p(M/\prod \mathcal{B}) = 0$$

for all primes p, and a subfactorization if and only if

$$\nu_p(M/\prod \mathcal{B}) \geq 0$$

for all primes p. We refer to $\nu_p(M/\prod \mathcal{B})$ as the *p-surplus* of $\mathcal{B}$ (as an attempted factorization of M) at prime p, and $-\nu_p(M/\prod \mathcal{B}) = \nu_p(\prod \mathcal{B}/M)$ as the *p-deficit*, and say that the factorization is *p-balanced* if $\nu_p(M/\prod \mathcal{B}) = 0$. Thus a subfactorization (resp. factorization) occurs when one has non-negative surpluses (resp. balance) at all primes p.

Example 3.1. Suppose one wishes to factorize $5! = 2^3 \times 3 \times 5$. The attempted factorization $\mathcal{B} := \{3,4,5,5\}$ has a 2-surplus of $\nu_2(5!/\prod \mathcal{B}) = 1$, is in balance at 3, and has a 5-deficit of $\nu_2(\prod \mathcal{B}/5!) = 1$, so it is not a factorization or subfactorization of $5!$. However, if one replaces one of the copies of 5 in $\mathcal{B}$ with a 2, this erases both the 2-surplus and the 5-deficit, and creates a factorization $\{2,3,4,5\}$ of $5!$.

A finite multiset $\mathcal{B}$ is said to be *t-admissible* for some $t > 0$ if $b \geq t$ for all $b \in \mathcal{B}$. Then $t(N)$ is largest quantity such that there exists a $t(N)$-admissible factorization of $N!$ of cardinality N.

Call a natural number 3-*smooth* if it is of the form $2^n 3^m$ for some natural numbers n, m. Given a positive real number x, we use $\lceil x \rceil^{\langle 2,3 \rangle}$ to denote the smallest 3-smooth number greater than or equal to x. For instance, $\lceil 5 \rceil^{\langle 2,3 \rangle} = 6$ and $\lceil 10 \rceil^{\langle 2,3 \rangle} = 12$.

3.2. **Description of algorithm.** We now describe an algorithm that, for given $1 \leq t \leq N$, either successfully demonstrates that $t(N) \geq t$, or halts with an error.

(0) Select natural numbers A, K such that $K^2(1+\sigma) < t$, where
$$\sigma := \frac{3N/t}{A}. \tag{3.1}$$
There is some freedom to select parameters here, but roughly speaking one would like to have $1 \lll A \lll K \lll \sqrt{t}$.

(1) Let I denote the elements of the interval[1] $(t, t(1+\sigma)]$ that are coprime to 6. Let $\mathcal{B}^{(1)}$ be the elements of I, each occurring with multiplicity A. This multiset is t-admissible,

[1] Numerically, it would be slightly better to use the closed interval $[t, t(1+\sigma)]$ instead of the half-open interval $(t, t(1+\sigma)]$, but we will consistently aim to use half-open intervals here to be compatible with standard notation for the prime counting function $\pi(x)$.

and $\prod \mathcal{B}^{(1)}$ is not divisible by tiny primes $2,3$. (It will have approximately the right number of primes for $3 < p \leq t/K$, though it may have quite different prime factorization at primes $p > t/K$.)

(2) Remove any element from $\mathcal{B}^{(1)}$ that contains a prime factor p with $p > t/K$, and call this new multiset $\mathcal{B}^{(2)}$. It remains t-admissible with no tiny prime factors, though it tends to acquire a p-surplus in the range $3 < p \leq K$.

(3) For each $p > t/K$, add in $\nu_p(N!)$ copies of the number $p\lceil t/p \rceil$ to $\mathcal{B}^{(2)}$, and call this new multiset $\mathcal{B}^{(3)}$. Now $\mathcal{B}^{(3)}$ is t-admissible and in balance at all primes $p > t/K$, but will typically be in a slight deficit at primes $3 < p \leq t/K$, particularly in the range $3 < p \leq K$. (It will now also contain a few tiny prime factors, but will generally still have a large surplus at those primes.)

(4) For each prime $3 < p \leq t/K$ at which there is a surplus $\nu_p(N!/\prod \mathcal{B}) > 0$, replace $\nu_p(N!/\prod \mathcal{B})$ copies of p in the prime factorizations of elements of $\mathcal{B}^{(3)}$ with $\lceil p \rceil^{\langle 2,3 \rangle}$ instead, and call this new multiset $\mathcal{B}^{(4)}$. Thus $\mathcal{B}^{(4)}$ has no surplus at primes $3 < p \leq t/K$ (and is still t-admissible and in balance for $p > t/K$).

(5) For the primes $3 < p \leq t/K$ at which there is a deficit $\nu_p(\prod \mathcal{B}/N!) > 0$, multiply all these primes together, and use the greedy algorithm to group them into factors $x_1, \dots, x_M$ in the range $(\sqrt{t/K}, t/K]$, together with possibly one exceptional factor x_* in the range $(1, t/K]$. For each of these factors x_i or x_*, add the quantity $x_i \lceil t/x_i \rceil^{\langle 2,3 \rangle}$ or $x_* \lceil t/x_* \rceil^{\langle 2,3 \rangle}$ to $\mathcal{B}^{(4)}$, and call this new multiset $\mathcal{B}^{(5)}$.

(6) By construction, $\mathcal{B}^{(5)}$ is t-admissible and will be in balance at all primes $p > 3$, and is thus $N!/\prod \mathcal{B}^{(5)}$ is of the form $2^n 3^m$ for some integers n, m. If at least one of n, m is negative, then `HALT` the algorithm with an error. Otherwise, select a 3-smooth number $2^{n_1} 3^{m_1}$ greater than equal to t with $n_1/m_1 \leq n/m$ (which one can interpret as $n_1 m \leq n m_1$ in case some of the denominators here vanish), and similarly select a 3-smooth number $2^{n_2} 3^{m_2}$ greater than or equal to t with $n_2/m_2 \geq n/m$. (It is reasonable to select the smallest such 3-smooth numbers in both cases, although this is not absolutely necessary for the algorithm to be successful.) By construction, we can express (n, m) as a positive linear combination $\alpha_1(n_1, m_1) + \alpha_2(n_2, m_2)$ of (n_1, m_1) and (n_2, m_2). Add $\lfloor \alpha_1 \rfloor$ copies of $2^{n_1} 3^{m_1}$ and $\lfloor \alpha_2 \rfloor$ copies of $2^{n_2} 3^{m_2}$ to $\mathcal{B}^{(5)}$, and call this tuple $\mathcal{B}^{(6)}$. (This will largely eliminate the surplus at 2 and 3.)

(7) If the multiset $\mathcal{B}^{(6)}$ has cardinality less than N, `HALT` the algorithm with an error. Otherwise, delete elements from $\mathcal{B}^{(6)}$ to bring the cardinality to N, and arbitrarily distribute any surplus primes to one of the remaining elements, and call the resulting multiset $\mathcal{B}^{(7)}$. By construction, $\mathcal{B}^{(7)}$ is a t-admissible factorization of $N!$ into N numbers, demonstrating that $t(N) \geq t$.

It will be convenient to divide the set of primes into the following ranges:

- *Tiny primes* $p = 2, 3$.
- *Small primes* $3 < p \leq K$.
- *Borderline small primes* $K < p \leq K(1 + \sigma)$.
- *Medium primes* $K(1 + \sigma) < p \leq t/K$.
- *Large primes* $p > t/K$.

The expected p-surpluses or p-deficits at various stages of this process are summarized in Table 1.

	Tiny p	Small p	Borderline p	Medium p	Large p
$\mathcal{B}^{(1)}$	Max. surplus	Near balance	Near balance	Near balance	???
$\mathcal{B}^{(2)}$	Max. surplus	Med. surplus	Med. surplus?	Near balance	Max. surplus
$\mathcal{B}^{(3)}$	Lg. surplus	Sm. surplus?	Med. surplus?	Near balance	Balance
$\mathcal{B}^{(4)}$	Lg. surplus	Balance?	Balance?	Balance/sm. deficit	Balance
$\mathcal{B}^{(5)}$	Lg. surplus	Balance	Balance	Balance	Balance
$\mathcal{B}^{(6)}$	Sm. surplus	Balance	Balance	Balance	Balance
$\mathcal{B}^{(7)}$	Balance	Balance	Balance	Balance	Balance

TABLE 1. Evolution of the surpluses and deficits of the multisets $\mathcal{B}^{(i)}$, $i = 1, \dots, 7$; we describe the size of these surpluses and deficits informally as "small", "medium", "large", or "maximal". For entries with a question mark, we allow the possibility of a tiny deficit. For the entry marked ???, all behavior from large surpluses to large deficits are possible.

3.3. **Analysis of Step 7.** We now analyze the above algorithm, starting from the final Step 7 and working backwards to Step 1, to establish sufficient conditions for the algorithm to successfully demonstrate that $t(N) \geq t$.

It will be convenient to introduce the following notation. For $a_+, a_- \in [0, +\infty]$, we define the asymmetric norm $|x|_{a_+,a_-}$ of a real number x by the formula

$$|x|_{a_+,a_-} := \begin{cases} a_+|x| & x \geq 0 \\ a_-|x| & x \leq 0. \end{cases}$$

If a_+, a_- are finite, this function is Lipschitz with constant $\max(a_+, a_-)$. One can think of a_+ as the "cost" of making x positive, and a_- as the "cost" of making x negative.

We now begin the analysis of Step 9. This procedure will terminate successfully as long as the length $|\mathcal{B}^{(6)}|$ of the tuple is at least N. To ensure this, we introduce the *t-excess* of a multiset $\mathcal{B}$ by the formula

$$\mathcal{E}_t(\mathcal{B}) := \prod_{b \in \mathcal{B}} \log \frac{b}{t} = \log \prod \mathcal{B} - |\mathcal{B}| \log t.$$

Thus, to ensure the success of this step, it suffices to establish the inequality

$$\mathcal{E}_t(\mathcal{B}^{(6)}) \leq \log \prod \mathcal{B}^{(6)} - N \log t.$$

From (1.1) we have

$$\log \prod \mathcal{B}^{(6)} = \log N! - \sum_p \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(6)}} \right) \log p,$$

so we can rewrite the previous condition (using the fact that $\mathcal{B}^{(6)}$ is a subfactorization of $N!$) as

$$\mathcal{E}_t(\mathcal{B}^{(6)}) + \sum_p \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(6)}} \right) \right|_{\log p, \infty} \leq \log N! - N \log t.$$

If we assume that $t = N/e^{1+\delta}$ for some $\delta > 0$, we can use the Stirling approximation (1.4) to reduce to the sufficient condition

$$\mathcal{E}_t(\mathcal{B}^{(6)}) + \sum_p \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(6)}} \right) \right|_{\log p, \infty} \leq \delta N + \log \sqrt{2\pi N}. \tag{3.2}$$

3.4. **Analysis of Step 6.** Now we analyze Step 6. For any $L \geq 1$, let κ_L be the least quantity such that

$$x \leq \lceil x \rceil^{\langle 2,3 \rangle} \leq \exp(\kappa_L) x \tag{3.3}$$

holds for all $x \geq L$. Just from considering the powers of two, we have the trivial upper bound

$$\kappa_L \leq \log 2. \tag{3.4}$$

We shall obtain better estimates on this quantity in Section A. For now we use this quantity to help achieve efficient subfactorizations of 3-smooth numbers, as follows.

Lemma 3.2. *Let $L \geq 1$. Let $t > 3L$ and let $2^n 3^m$ be a 3-smooth number with $n, m > 0$ obeying the conditions*

$$\frac{\log(3L) + \kappa_L}{\log t - \log(3L)} \leq \frac{n \log 2}{m \log 3} \leq \frac{\log t - \log(2L)}{\log(2L) + \kappa_L}. \tag{3.5}$$

Then one can find a t-admissible subfactorization $\mathcal{B}$ of $2^n 3^m$ such that

$$\mathcal{E}_t(\mathcal{B}) \leq \kappa_L \frac{n \log 2 + m \log 3}{\log t} \tag{3.6}$$

and

$$|\nu_2(2^n 3^m / \mathcal{B})|_{\log 2, \infty} + |\nu_3(2^n 3^m / \mathcal{B})|_{\log 3, \infty} \leq 2(\log t + \kappa_L). \tag{3.7}$$

In practice, $\log t$ will be significantly larger than $\log(2L)$ or $\log(3L)$, and so the hypothesis (3.5) will be quite mild, as long as n and m are both reasonably large.

Proof. Let $2^{n_0}, 3^{m_0}$ be the largest powers of 2 and 3 less than or equal to t/L respectively, thus

$$L \leq \frac{t}{2^{n_0}} \leq 2L \tag{3.8}$$

and

$$L \leq \frac{t}{3^{m_0}} \leq 3L. \tag{3.9}$$

From (3.3), the 3-smooth numbers $\lceil t/2^{n_0} \rceil^{\langle 2,3 \rangle} = 2^{n_1} 3^{m_1}$, $\lceil t/3^{m_0} \rceil^{\langle 2,3 \rangle} = 2^{n_2} 3^{m_2}$ obey the estimates

$$\frac{t}{2^{n_0}} \leq 2^{n_1} 3^{m_1} \leq e^{\kappa_L} \frac{t}{2^{n_0}} \tag{3.10}$$

and

$$\frac{t}{3^{m_0}} \leq 2^{n_2} 3^{m_2} \leq e^{\kappa_L} \frac{t}{3^{m_0}}, \tag{3.11}$$

or equivalently

$$t \leq 2^{n_0 + n_1} 3^{m_1}, 2^{n_2} 3^{m_0 + m_2} \leq e^{\kappa_L} t. \tag{3.12}$$

We can use (3.8), (3.10) to bound

$$\begin{aligned}\frac{n_0+n_1}{m_1} &\geq \frac{n_0}{\log(e^{\kappa_L}\frac{t}{2^{n_0}})/\log 3}\\ &\geq \frac{(\log t - \log(2L))/\log 2}{(\log(2L)+\kappa_L)/\log 3}\end{aligned}$$

(with the convention that this bound is vacuously true for $m_1 = 0$). Similarly, from (3.9), (3.11) we have

$$\begin{aligned}\frac{n_2}{m_0+m_2} &\leq \frac{\log(e^{\kappa}\frac{t}{3^{m_0}})/\log 2}{m_0}\\ &\leq \frac{(\log(3L)+\kappa)/\log 2}{(\log t - \log(3L))/\log 3}\end{aligned}$$

and hence by (3.5)

$$\frac{n_2}{m_0+m_2} \leq \frac{n}{m} \leq \frac{n_0+n_1}{m_1}. \tag{3.13}$$

Thus we can write (n,m) as a non-negative linear combination

$$(n,m) = \alpha_1(n_0+n_1, m_1) + \alpha_2(n_2, m_0+m_2)$$

for some real $\alpha_1, \alpha_2 \geq 0$. We now take our subfactorization $\mathcal{B}$ to consist of $\lfloor \alpha_1 \rfloor$ copies of the 3-smooth number $2^{n_0+n_1}3^{m_1}$ and $\lfloor \alpha_2 \rfloor$ copies of the 3-smooth number $2^{n_2}3^{m_0+m_2}$. By (3.12), each term $2^{n'}3^{m'}$ here is admissible and contributes a t-excess of at most κ_L, which is in turn bounded by $\kappa_L \frac{n'\log 2 + m'\log 3}{\log t}$. Adding these bounds together, we obtain (3.6).

The expression $2^n3^m/\prod \mathcal{B}$ contains at most $n_0+n_1+n_2$ factors of 2 and at most $m_0+m_2+m_1$ factors of 3, hence

$$\nu_2(2^n3^m/\prod \mathcal{B})\log 2 + \nu_3(2^n3^m/\prod \mathcal{B})\log 3 \leq \log 2^{n_0+n_1}3^{m_1} + \log 2^{n_2}3^{m_0+m_2},$$

and the bound (3.7) follows from (3.12). □

We now use this lemma to analyze Step 6 as follows.

Proposition 3.3. *Let $L \geq 1$. Let $3L < t = N/e^{1+\delta}$ for some $\delta > 0$, and let $1 \leq K \leq t$ and $A \geq 1$. Suppose that the algorithm in Section 3.2 with the indicated parameters reaches the end of Step 5 with a multiset $\mathcal{B}^{(5)}$ obeying the following hypotheses:*

(i) *(Small excess and surplus at non-tiny primes)*

$$\mathcal{E}_t(\mathcal{B}^{(5)}) + \sum_{p>3}\left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(5)}}\right)\right|_{\log p,\infty} \leq \delta N + \log\sqrt{2\pi} - \frac{3}{2}\log N - (\kappa_L \log\sqrt{12})\frac{N}{\log t}. \tag{3.14}$$

(ii) *(Large surpluses at tiny primes) The surpluses $\nu_2(N!/\prod \mathcal{B}^{(5)})$, $\nu_3(N!/\prod \mathcal{B}^{(5)})$ lie in the sector $\Gamma_{t,L} \subset \mathbb{R}^2$, defined to be the set of pairs (n,m) with $n,m > 0$ and*

$$\frac{\log(3L)+\kappa_L}{\log t - \log(3L)} \leq \frac{n\log 2}{m\log 3} \leq \frac{\log t - \log(2L)}{\log(2L)+\kappa_L}.$$

Then $t(N) \geq t$.

Proof. Write $n := \nu_2(N!/\prod \mathcal{B}^{(5)})$ and $m := \nu_3(N!/\prod \mathcal{B}^{(5)})$. From (1.3) we have $n \leq N$ and $m \leq N/2$, hence

$$n \log 2 + m \log 3 \leq N \log \sqrt{12}.$$

Applying Lemma 3.2, we can find a subfactorization $\mathcal{B}'$ of $2^n 3^m$ with an excess of at most $(\kappa_L \log \sqrt{12})\frac{N}{\log t}$, and with

$$\left|\nu_2\left(\frac{2^n 3^m}{\prod \mathcal{B}'}\right)\right|_{\log 2,\infty} + \left|\nu_3\left(\frac{2^n 3^m}{\prod \mathcal{B}'}\right)\right|_{\log 3,\infty} \leq 2(\log t + \kappa_L) \leq 2 \log N$$

where we have used (3.4) and the fact that $\log t \leq \log N - 1$. Then $\mathcal{B}^{(6)} = \mathcal{B}^{(5)} \cup \mathcal{B}'$ is another t-admissible multiset, and from (3.14), we obtain the previous sufficient condition (3.2). □

3.5. **Analysis of Step 5.**

Proposition 3.4. *Let $1 \leq K \leq t \leq N$, $A \geq 1$, and $L \geq 1$ be parameters such that $9L < t = N/e^{1+\delta}$ for some $\delta > 0$. Suppose that the algorithm in Section 3.2 with the indicated parameters reaches the end of Step 4 to produce a multiset $\mathcal{B}^{(4)}$ obeying the following hypotheses.*

(i) *(Small excess and surplus at small/medium primes)*

$$\begin{aligned}\mathcal{E}_t(\mathcal{B}^{(4)}) + \sum_{3<p\leq t/K} \left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)\right|_{\kappa_K \min(\frac{\log p}{\log \sqrt{t/K}},1),\infty} \\ \leq \delta N - \frac{3}{2}\log N - \kappa_L (\log \sqrt{12})\frac{N}{\log t}.\end{aligned} \tag{3.15}$$

(ii) *(Large surpluses at tiny primes) Whenever n_{**}, m_{**} are natural numbers obeying the bounds*

$$n_{**} \log 2 + m_{**} \log 3 \leq \sum_{3<p\leq t/K} \left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)\right|_{\frac{\log \sqrt{tK}+\kappa_K}{\log \sqrt{t/K}} \log p,\infty} + \log t + \kappa_L,$$

then one has

$$\left(\nu_2\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right) - n_{**}, \nu_3\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right) - m_{**}\right) \in \Gamma_{t,L}.$$

Then $t(N) \geq t$.

Proof. By (3.15), $\mathcal{B}^{(4)}$ is a subfactorization of $N!$, and by construction it is in balance at all primes $p > t/K$. Consider all the p-surplus primes in the small, borderline small, and medium range $3 < p \leq t/K$, thus each such prime is considered with multiplicity $\nu_p(N!/\prod \mathcal{B}^{(4)})$. Using the greedy algorithm, one can factor the product of all these primes into M factors $c_1, \dots, c_M$ in the interval $(\sqrt{t/K}, t/K]$, times at most one exceptional factor c_* in $(1, \sqrt{t/K}]$, for some M. If we let M' denote the number of factors in $c_1, \dots, c_M$ that are not divisible by a prime larger than $\sqrt{t/K}$, we have the bound

$$\left(\sqrt{t/K}\right)^{M'} \leq \prod_{3<p\leq\sqrt{t/K}} \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)$$

and hence on taking logarithms

$$M' \leq \sum_{3<p\leq\sqrt{t/K}} \left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)\right|_{\frac{\log p}{\log\sqrt{t/K}},\infty}.$$

Restoring the factors divisible by primes $p > \sqrt{t/K}$, we conclude that

$$M \leq \sum_{3<p\leq t/K} \left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)\right|_{\min(\frac{\log p}{\log\sqrt{t/K}},1),\infty}. \tag{3.16}$$

For each of the M factors c_i, we introduce the 3-smooth number $\lceil t/c_i\rceil^{\langle 2,3\rangle} = 2^{n_i}3^{m_i}$, which by (3.3) lies in the interval $[t/c_i, e^{\kappa_K}t/c_i]$; similarly, for the exceptional factor c_* we introduce a 3-smooth number $\lceil t/c_*\rceil^{\langle 2,3\rangle} = 2^{n_*}3^{m_*}$ in the interval $[t/c_*, e^{\kappa_K}t/c_*]$. If we then adjoin the 3-smooth numbers $\lceil t/c_i\rceil^{\langle 2,3\rangle}c_i = 2^{n_i}3^{m_i}c_i$ for $i = 1, \dots, M$ as well as $\lceil t/c_*\rceil^{\langle 2,3\rangle}c_* = 2^{n_*}3^{m_*}c_*$ to the t-admissible multiset $\mathcal{B}^{(4)}$ to create a new t-admissible multiset $\mathcal{B}^{(5)}$. The quantity $\log\lceil t/c_*\rceil^{\langle 2,3\rangle} = n_i\log 2 + m_i\log 3$ is bounded by $\log\sqrt{tK} + \kappa_K$, and the quantity $\log\lceil t/c_*\rceil^{\langle 2,3\rangle} = n_*\log 2 + m_*\log 3$ is similarly bounded by $\log t + \kappa$, hence if we denote $n_{**} := n_1 + \dots + n_M + n_*$ and $m_{**} := m_1 + \dots + m_M + m_*$, we have

$$n_{**}\log 2 + m_{**}\log 3 \leq \frac{\log\sqrt{tK} + \kappa_K}{\log\sqrt{t/K}} \sum_{3<p\leq t/K} \left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)\right|_{\log p,\infty} + \log t + \kappa_K.$$

Each of the new factors in $\mathcal{B}^{(5)}$ contributes an excess of at most κ_K, so the total excess of $\mathcal{B}^{(5)}$ is at most

$$\mathcal{E}_t(\mathcal{B}^{(4)}) + \kappa_K M + \kappa_K$$

which by (3.16) is bounded by

$$\mathcal{E}_t(\mathcal{B}^{(4)}) + \sum_{3<p\leq t/K} \left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)\right|_{\kappa_K\min(\frac{\log p}{\log\sqrt{t/K}},1),\infty} + \kappa_K.$$

We conclude that $\mathcal{B}^{(5)}$ obeys the hypotheses of Proposition 3.3 (using (3.4) to bound κ_K by $\log\sqrt{2\pi}$), and the claim follows. □

3.6. **Analysis of Step 4.**

Proposition 3.5. *Let $L \geq 1$. Let $9L < t = N/e^{1+\delta}$ for some $\delta > 0$, and suppose that the algorithm reaches the end of Step 3 to produce a multiset $\mathcal{B}^{(3)}$ obeying the following hypotheses:*

(i) *(Small excess and surplus at small/medium primes) One has*

$$\begin{aligned}\mathcal{E}_t(\mathcal{B}^{(3)}) + \sum_{3<p\leq t/K} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right)\right|_{\kappa_K \min(\frac{\log p}{\log \sqrt{t/K}},1),\kappa_p} \\ \leq \delta N - \frac{3}{2}\log N - \kappa_L(\log\sqrt{12})\frac{N}{\log t}.\end{aligned} \tag{3.17}$$

(ii) *(Large surpluses at tiny primes) Whenever n_{**}, m_{**} are natural numbers obeying the bounds*

$$\begin{aligned}n_{**}\log 2 + m_{**}\log 3 \leq \sum_{3<p\leq t/K} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right)\right|_{\frac{\log\sqrt{tK}+\kappa_K}{\log\sqrt{t/K}}\log p, \log p+\kappa_p} \\ + \log t + \kappa_K,\end{aligned} \tag{3.18}$$

then one has

$$\left(\nu_2\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right) - n_{**}, \nu_3\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right) - m_{**}\right) \in \Gamma_{t,L}. \tag{3.19}$$

Then $t(N) \geq t$.

Proof. Suppose there is a non-tiny prime $p > 3$ with a positive p-deficit $|\nu_p(N!/\prod \mathcal{B}^{(3)})|_{0,1} > 0$. Since $\mathcal{B}^{(3)}$ is in balance at all large primes, we have $3 < p \leq t/K$. We locate an element of $\mathcal{B}^{(3)}$ that contains p as a factor, and replaces it with $\lceil p \rceil^{\langle 2,3\rangle} = 2^{n_p}3^{m_p}$, which increases that factor by at most $\exp(\kappa_p)$ thanks to (3.3). This procedure reduces the p-deficit by one, adds at most κ_p to the t-excess, and decrements $\nu_2(N!/\prod \mathcal{B}^{(3)})$, $\nu_3(N!/\prod \mathcal{B}^{(3)})$ by n_p, m_p respectively. Since $n_p \log 2 + m_p \log 3 \leq \log p + \kappa_p$, if we apply this procedure to clear all deficits at non-tiny primes, the resulting multiset $\mathcal{B}^{(4)}$ has a t-excess of

$$\mathcal{E}_t(\mathcal{B}^{(4)}) \leq \mathcal{E}_t(\mathcal{B}^{(3)}) + \sum_{p>3} |\nu_p(N!/\prod \mathcal{B})|_{0,\kappa_p}$$

and we have

$$\nu_2(N!/\prod \mathcal{B}^{(4)}) = \nu_2(N!/\prod \mathcal{B}^{(3)}) - n', \quad \nu_3(N!/\prod \mathcal{B}^{(4)}) = \nu_3(N!/\prod \mathcal{B}^{(3)}) - m'$$

with

$$n'\log 2 + m'\log 3 \leq \sum_{p>3} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right)\right|_{0,\log p+\kappa_p}.$$

The hypotheses of Proposition 3.4 are now satisfied, and we are done. □

To simplify the criteria here, we introduce the quantities

$$X_1 := \sum_{3<p\leq K} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right)\right|_{\frac{\log p}{\log\sqrt{t/K}},0} + \sum_{K<p\leq t/K} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right)\right| + \tag{3.20}$$

$$X_2 := \sum_{3<p\leq K} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right)\right|_{0,1}. \tag{3.21}$$

Since $\kappa_K \min(\frac{\log p}{\log\sqrt{t/K}}, 1), \kappa_p$ are both bounded by κ_K for $p \geq K$, and bounded by $\kappa_K \frac{\log p}{\log\sqrt{t/K}}$, κ_5 respectively for $3 < p \leq K$, we can replace (3.17) with

$$\mathcal{E}_t(\mathcal{B}^{(3)}) + \kappa_K X_1 + \kappa_5 X_2 \leq \delta N - \frac{3}{2}\log N - \kappa_L(\log\sqrt{12})\frac{N}{\log t}. \tag{3.22}$$

Similarly, since $\frac{\log\sqrt{tK}+\kappa_K}{\log\sqrt{t/K}}\log p, \log p+\kappa_p$ are both bounded by $\log(tK)+2\kappa_K$ for $K < p \leq t/K$, $\frac{\log\sqrt{tK}+\kappa_K}{\log\sqrt{t/K}}\log p$ is bounded by $(\log(tK)+2\kappa_K)\frac{\log p}{\log\sqrt{t/K}}$ for $3 < p \leq K$, and $\log p+\kappa_p$ is bounded by $\log K + \kappa_5$ for $3 < p \leq K$, we can replace (3.18) with

$$n_{**}\log 2 + m_{**}\log 3 \leq (\log(tK) + 2\kappa_K)(X_1 + 1) + (\log K + \kappa_5)X_2. \tag{3.23}$$

3.7. **Analysis of Steps 1,2,3.** To apply Proposition 3.5, we now compute the various statistics of $\mathcal{B}^{(3)}$ produced by Steps 1-3.

We begin with the analysis of $\mathcal{B}^{(1)}$, constructed in Step 1 of the algorithm. To count elements coprime to 6, we use the following lemma:

Lemma 3.6. *For any interval $(a, b]$ with $0 \leq a \leq b$, the number of natural numbers in the interval that are coprime to 6 is $\frac{b-a}{3} + O_{\leq}(4/3)$.*

Proof. By the triangle inequality, it suffices to show that the number of natural numbers coprime to 6 in $[0, a]$, minus $a/3$, is $O_{\leq}(2/3)$. The claim is easily verified for $0 \leq a \leq 6$, and the quantity in question is 6-periodic in a, giving the claim. ☐

The excess of $\mathcal{B}^{(1)}$ can be computed as

$$\mathcal{E}_t(\mathcal{B}^{(1)}) = A\sum_{n\in I}\log\frac{n}{t}.$$

By the fundamental theorem of calculus, this is

$$A\int_0^{3t/A} |I \cap (t, t+h]|\,\frac{dh}{t+h}.$$

Bounding $\frac{1}{t+h}$ by $\frac{1}{t}$ and applying Lemma 3.6, we conclude that

$$\mathcal{E}_t(\mathcal{B}^{(1)}) \leq A\int_0^{3N/A}\left(\frac{h}{3}+\frac{4}{3}\right)\frac{dh}{t} = \frac{3N^2}{2tA} + 4. \tag{3.24}$$

Next, we compute p-valuations $\nu_p(\mathcal{B}^{(1)})$. By construction, this quantity vanishes at tiny primes $p = 2, 3$. For $p > 3$, we can use Lemma 3.6 again to conclude

$$\begin{aligned}
\nu_p(\mathcal{B}^{(1)}) &= A \sum_{1 \le j \le \frac{\log N}{\log p}} |I \cap p^j \mathbb{Z}| \\
&= A \sum_{1 \le j \le \frac{\log N}{\log p}} \left(\frac{N}{p^j A} + O_{\le}(4/3) \right) \\
&= \frac{N}{p-1} + O_{\le}\left(\frac{1}{p-1} \right) + O_{\le}\left(\frac{4A}{3} \left\lceil \frac{\log N}{\log p} \right\rceil \right) \\
&= \frac{N}{p-1} + O_{\le}\left(\frac{4A + 0.75}{3} \left\lceil \frac{\log N}{\log p} \right\rceil \right)
\end{aligned}$$

since $\frac{1}{p-1} \le \frac{0.75}{3}$. Meanwhile, from (1.3) one has

$$\nu_p(N!) = \frac{N}{p-1} + O_{\le}\left(\left\lceil \frac{\log N}{\log p} \right\rceil \right)$$

and thus

$$\nu_p(N!/\mathcal{B}^{(1)}) = O_{\le}\left(\frac{4A + 3.75}{3} \left\lceil \frac{\log N}{\log p} \right\rceil \right). \tag{3.25}$$

Now we pass to $\mathcal{B}^{(2)}$ by performing Step 2 of the algorithm. Removing elements from a t-admissible multiset cannot increase the t-excess, so from (3.24) we have

$$\mathcal{E}_t(\mathcal{B}^{(2)}) \le \frac{3N^2}{2tA} + 4. \tag{3.26}$$

The elements removed are of the form pm with $m \le K(1+\nu)$ coprime to 6, and p in the interval $(\frac{t}{\min(m,K)}, \frac{t(1+\sigma)}{m}]$. We conclude that

$$\nu_p(N!/\mathcal{B}^{(2)}) = \nu_p(N!/\mathcal{B}^{(1)}) \tag{3.27}$$

for medium primes $K(1+\sigma) < p \le t/K$. For small and borderline small primes $3 < p \le K(1+\sigma)$ one has

$$\nu_p(N!/\mathcal{B}^{(2)}) = \nu_p(N!/\mathcal{B}^{(1)}) + A \sum_{\substack{m \le K(1+\sigma) \\ (m,6)=1}} \nu_p(m) \left(\pi\left(\frac{t(1+\sigma)}{m} \right) - \pi\left(\frac{t}{\min(m,K)} \right) \right). \tag{3.28}$$

Finally, for tiny primes $p = 2, 3$ we have the maximal surplus:

$$\nu_p(N!/\mathcal{B}^{(2)}) = \nu_p(N!).$$

We now pass to $\mathcal{B}^{(3)}$ by performing Step 3 of the algorithm. In other words, we add $\nu_p(N!)$ copies of $p\lceil t/p \rceil$ for each largeprime $p > t/K$. The t-excess is now given by

$$\mathcal{E}_t(\mathcal{B}^{(3)}) = \mathcal{E}_t(\mathcal{B}^{(2)}) + \sum_{p > t/K} \nu_p(N!) \log \frac{\lceil t/p \rceil}{t/p}$$

and hence by (3.26)

$$\mathcal{E}_t(\mathcal{B}^{(3)}) = \frac{3N^2}{2tA} + 4 + \sum_{p>t/K} \nu_p(N!) \log \frac{\lceil t/p \rceil}{t/p}. \tag{3.29}$$

By construction one has balance (3.30) at large primes $p > t/K$,

$$\nu_p(N!/\mathcal{B}^{(3)}) = 0 \tag{3.30}$$

and no modification at borderline small or medium primes $K < p \leq t/K$,

$$\nu_p(N!/\mathcal{B}^{(3)}) = \nu_p(N!/\mathcal{B}^{(2)}) \tag{3.31}$$

but now the p-surplus or p-deficit at small primes $3 < p \leq K$ is modified:

$$\nu_p(N!/\mathcal{B}^{(3)}) = \nu_p(N!/\mathcal{B}^{(2)}) - \sum_{p'>t/K} \nu_{p'}(N!) \nu_p(\lceil t/p' \rceil). \tag{3.32}$$

Similarly, at tiny primes $p = 2, 3$ we have

$$\nu_p(N!/\mathcal{B}^{(3)}) = \nu_p(N!) - \sum_{p'>t/K} \nu_{p'}(N!) \nu_p(\lceil t/p' \rceil). \tag{3.33}$$

At medium primes, $K(1+\sigma) < p \leq t/K$, we see from (3.31), (3.25) that

$$\left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(3)}} \right) \right| \leq \frac{4A + 3.75}{3} \left\lceil \frac{\log N}{\log p} \right\rceil .$$

For borderline primes $K \leq p < K(1+\sigma)$, we have from (3.31), (3.28) that

$$\left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(3)}} \right) \right| \leq \frac{4A + 3.75}{3} \left\lceil \frac{\log N}{\log p} \right\rceil + A \sum_{\substack{m \leq K(1+\sigma) \\ (m,6)=1}} \nu_p(m) \left(\pi \left(\frac{t(1+\sigma)}{m} \right) - \pi \left(\frac{t}{\min(m,K)} \right) \right).$$

The only m which contributes here is $m = p$, thus we may simplify to

$$\begin{aligned} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(3)}} \right) \right| &\leq \frac{4A + 3.75}{3} \left\lceil \frac{\log N}{\log p} \right\rceil + A \left(\pi \left(\frac{t}{p} + \frac{\sigma t}{p} \right) - \pi \left(\frac{t}{K} \right) \right) \\ &\leq \frac{4A + 3.75}{3} \left\lceil \frac{\log N}{\log p} \right\rceil + A \left(\pi \left(\frac{t(1+\sigma)}{K} \right) - \pi \left(\frac{t}{K} \right) \right). \end{aligned}$$

For the small primes $3 < p \leq K$, we see from (3.25), (3.28), (3.32) that we have the upper bound

$$\nu_p \left(\frac{N!}{\prod \mathcal{B}^{(3)}} \right) \leq \frac{4A + 3.75}{3} \left\lceil \frac{\log N}{\log p} \right\rceil + Y_p + Z_p$$

and the lower bound

$$\nu_p \left(\frac{N!}{\prod \mathcal{B}^{(3)}} \right) \leq -\frac{4A + 3.75}{3} \left\lceil \frac{\log N}{\log p} \right\rceil + Y_p$$

where Y_p is the quantity

$$\begin{aligned} Y_p := A &\sum_{\substack{m \leq K \\ (m,6)=1}} \nu_p(m) \left(\pi \left(\frac{t(1+\sigma)}{m} \right) - \pi \left(\frac{t}{\min(m,K)} \right) \right) \\ &- \sum_{p'>t/K} \lceil \lfloor \frac{N}{p'} \rceil \rfloor \nu_p(\lceil t/p' \rceil) \end{aligned}$$

and Z_p is the (non-negative) error term

$$Z_p := A \sum_{\substack{K<m\leq K(1+\sigma)\\(m,6)=1}} \nu_p(m)\left(\pi\left(\frac{t(1+\sigma)}{m}\right)-\pi\left(\frac{t}{\min(m,K)}\right)\right).$$

Here we have used (1.3) to write $\nu_{p'}(N!)$ as $\lceil N/p' \rceil$. An important phenomenon for us will be that Y_p is usually positive (so that $\mathcal{B}^{(3)}$ typically enjoys a (modest) surplus at small primes rather than a deficit). From the triangle inequality we now have

$$\begin{aligned} X_1 \leq{} & \frac{4A+3.75}{3}\sum_{3<p\leq t/K}\left\lceil\frac{\log N}{\log p}\right\rceil \\ &+ A\left(\pi(K(1+\sigma))-\pi(K)\right)\left(\pi\left(\frac{t(1+\sigma)}{K}\right)-\pi\left(\frac{t}{K}\right)\right) \\ &+ \sum_{3<p\leq K}\frac{\log p}{\log\sqrt{t/K}}(|Y_p|_{1,0}+Z_p) \end{aligned} \tag{3.34}$$

and

$$X_2 \leq \frac{4A+3.75}{3}\sum_{3<p\leq K}\left\lceil\frac{\log N}{\log p}\right\rceil + \sum_{3<p\leq K}|Y_p|_{0,1}. \tag{3.35}$$

4. Asymptotic evaluation of $t(N)$

In this section we establish the lower bound

$$\frac{t}{N} \geq \frac{1}{e} - \frac{c_0}{\log N} + \frac{1}{\log^{1+c_1} N} \tag{4.1}$$

for some small absolute constant $0 < c_1 < 1$, if N is sufficiently large. With this choice of parameters, one has

$$\delta = \frac{ec_0}{\log N} + \frac{1}{\log^{1+c_1} N} + O\left(\frac{1}{\log^2 N}\right).$$

Let N be sufficiently large. We introduce parameters

$$A := \lfloor \log^2 N \rfloor$$

and

$$K := \lfloor \log^3 N \rfloor$$

and

$$L := N^{0.1}.$$

We apply the algorithm from Section 3.2, using the first option for Step 3, and invoke Proposition 3.5. With these parameters, we see from Lemma A.3 that the right-hand side of (3.17) is at least

$$ec_0\frac{N}{\log N} + \frac{N}{2\log^{1+c_1} N}$$

if c_1 is small enough and N is large enough. Thus, in view of (3.22), (3.23), it suffices to establish the bound

$$\mathcal{E}_t(\mathcal{B}^{(3)}) + \kappa_K X_1 + \kappa_5 X_2 \leq ec_0 \frac{N}{\log N} + O\left(\frac{N(\log\log N)^3}{\log^2 N} \right) \tag{4.2}$$

as well as the condition (3.19) whenever (3.23) holds.

By repeating the proof of Proposition 2.4, we see that

$$\sum_{p>t/K} \nu_p(N!) \log \frac{\lceil t/p \rceil}{t/p} = ec_0 \frac{N}{\log N} + O\left(\frac{N}{\log^2 N} \right)$$

and hence by (3.29)

$$\mathcal{E}_t(\mathcal{B}^{(2)}) = ec_0 \frac{N}{\log N} + O\left(\frac{N}{\log^2 N} \right).$$

Next, from (3.34), (3.35) and the prime number theorem, we can calculate that

$$X_1 \ll \sum_{3<p\leq K} \frac{\log p}{\log N} (|Y_p|_{1,0} + Z_p) + \frac{N}{\log^2 N}$$

and

$$X_2 \ll \sum_{3<p\leq K} |Y_p|_{0,1} + \log^6 N$$

so to verify (4.2) will suffice by Mertens' theorem to show that

$$|Y_p|_{1,0}, Z_p \ll \frac{N(\log\log N)^2}{p \log N} \tag{4.3}$$

and

$$|Y_p|_{0,1} \ll \frac{N(\log\log N)^2}{p \log^2 N}, \tag{4.4}$$

as this will imply that

$$X_1, X_2 \ll \frac{N(\log\log N)^3}{\log^2 N}. \tag{4.5}$$

Note the need to obtain stronger control on $|Y_p|_{0,1}$ than on $|Y_p|_{1,0}$.

The error term Z_p is easily disposed of:

$$\begin{aligned} Z_p &\ll A \sum_{K<m\leq K(1+\sigma)} \nu_p(m) \left(\pi\left(\frac{t(1+\sigma)}{K} \right) - \pi\left(\frac{t}{K} \right) \right) \\ &\ll \frac{A\sigma t}{K \log N} \sum_{K<m\leq K(1+\sigma)} \sum_{j \ll \log K} 1_{p^j | m} \\ &\ll \frac{N}{K \log N} \sum_{j\ll \log K} \left(\frac{K\sigma}{p^j} + 1\right) \\ &\ll \frac{N}{p \log^2 N}. \end{aligned}$$

As for Y_1, we see from Corollary C.2 that

$$A \sum_{m \leq K(1+\sigma); (m,6)=1} \nu_p(m) \left(\pi\left(\frac{t(1+\sigma)}{m}\right) - \pi\left(\frac{t}{m}\right) \right)$$

$$= (1+O(\sigma)) \sum_{m \leq K(1+\sigma); (m,6)=1} \nu_p(m) \frac{t\sigma}{\log(t/m)}; \qquad = \left(1 + O\left(\frac{\log\log N}{\log N}\right)\right) \frac{N}{\log N} \sum_{m \leq K(1+\sigma); (m,6)=1} \nu_p(m) \frac{3}{m}$$

since

$$\begin{aligned} \sum_{m \leq K(1+\sigma)} \frac{\nu_p(m)}{m} &= \sum_{j=1}^{\infty} \sum_{m \leq K(1+\sigma): p^j | m} \frac{1}{m} \\ &\ll \sum_{j=1}^{\infty} \frac{\log K}{p^j} \\ &\ll \frac{\log\log N}{p} \end{aligned} \tag{4.6}$$

we conclude the desired upper bound

$$|Y_1|_{1,0} \ll \frac{N \log\log N}{p \log N}.$$

To control the negative part, we apply Corollary C.2 again to calculate

$$\begin{aligned} \sum_{p' > t/K} \nu_{p'}(N!) \nu_p(\lceil t/p' \rceil) &\leq \sum_{p' > t/K} \frac{N}{p'} \nu_p(\lceil t/p' \rceil) \\ &= \sum_{m \leq K} \nu_p(m) \sum_{t/m \leq p' < t/(m-1)} \frac{N}{p'} \\ &= \left(1 + O\left(\frac{\log\log N}{\log N}\right)\right) \sum_{m \leq K} \nu_p(m) \int_{t/m}^{t/(m-1)} \frac{N}{x} \frac{dx}{\log N} \\ &= \left(1 + O\left(\frac{\log\log N}{\log N}\right)\right) \frac{N}{\log N} \sum_{m \leq K} \nu_p(m) \log \frac{m}{m-1} \\ &= \frac{N}{\log N} \sum_{m \leq K} \nu_p(m) \log \frac{m}{m-1} + O\left(\frac{N (\log\log N)^2}{p \log^2 N}\right) \end{aligned}$$

where to justify the last line we used (4.6). Putting this together, we find that

$$|Y_1|_{0,1} \ll \frac{N}{\log N} \Big| \sum_{m \leq K(1+\sigma); (m,6)=1} \nu_p(m) \frac{3}{m} - \sum_{m \leq K} \nu_p(m) \log \frac{m}{m-1} \Big|_{0,1} + \frac{N (\log\log N)^2}{p \log^2 N}. \tag{4.7}$$

We now have a crucial inequality:

Lemma 4.1 (Key inequality)**.** *For $p \geq 5$ and $K > 0$, we have*

$$0 \leq \sum_{m \leq K; (m,6)=1} \nu_p(m) \frac{3}{m} - \sum_{m \leq K} \nu_p(m) \log \frac{m}{m-1} \leq \frac{2}{p-1}.$$

We defer the proof of this lemma to Appendix B. Because of this lemma, the first term on the right-hand side of (4.7) vanishes, and we recover the desired bounds(4.3), (4.4), and hence (4.5), (4.2), (3.22), and (3.17).

Now we turn to the verification of (3.19) assuming (3.23). In view of (4.5), the latter condition implies that

$$n_{**}, m_{**} \ll \frac{N(\log\log N)^3}{\log N}.$$

Meanwhile, for a tiny prime $p = 2, 3$, we see from (3.33), (1.3), the prime number theorem, and (4.6) that

$$\begin{aligned}
\nu_p(N!/\mathcal{B}^{(3)}) &= \frac{N}{p-1} + O(\log N) - O\left(\sum_{p'>t/K} \frac{N}{p'}\nu_p\left(\lceil t/p' \rceil\right)\right) \\
&= \frac{N}{p-1} + O(\log N) - O\left(\sum_{m\le K} \nu_p(m) \sum_{t/m \le p' < t/(m-1)} \frac{N}{p'}\right) \\
&= \frac{N}{p-1} + O(\log N) - O\left(\frac{N}{\log N}\sum_{m\le K}\frac{\nu_p(m)}{m}\right) \\
&= \frac{N}{p-1} + O\left(\frac{N\log\log N}{\log N}\right).
\end{aligned}$$

We conclude that

$$\begin{aligned}
\nu_2(N!/\mathcal{B}^{(3)}) - n_{**} &= N + O\left(\frac{N(\log\log N)^2}{\log N}\right) \\
\nu_3(N!/\mathcal{B}^{(3)}) - m_{**} &= \frac{N}{2} + O\left(\frac{N(\log\log N)^2}{\log N}\right).
\end{aligned}$$

By choice of L, this implies (3.19) for N large enough. The proof of (4.1) is now complete.

5. Guy–Selfridge conjecture

We now establish the Guy–Selfridge conjecture $t(N) \ge \delta$ in the range

$$N \ge ???.$$

We will apply Proposition 3.5 with the choice of parameters

$$\begin{aligned}
t &:= N/3 \\
A &:= 90 \\
K &:= 342 \\
L &:= 342.
\end{aligned}$$

Clearly $\delta = \log\frac{3}{e} = 0.09861\dots$, and

$$\sigma = \frac{9}{A} = 0.1.$$

From Lemma A.1, we have

$$\kappa_K = \kappa_L \le \log\frac{9}{8} = 0.11778\dots \tag{5.1}$$

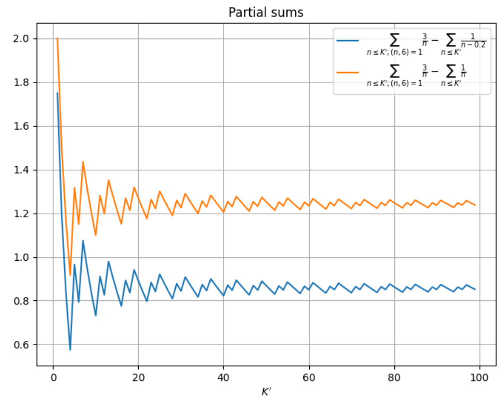

FIGURE 1. A plot of $f(x)$. The integral $c_1 = \int_{1/K}^1 f(x)\, dx \approx 0.9200$ is slightly larger than $ec_0 \approx 0.8244$.

Thus the right-hand side of (3.22) is at least

$$N \log \frac{3}{e} - \frac{3}{2} \log N - (\log \frac{9}{8})(\log \sqrt{12}) \frac{N}{\log(N/3)}.$$

Now we compute the t-excess. From (3.29) and (1.3) (noting that $N/3K \geq \sqrt{N}$) we have

$$\mathcal{E}_t(\mathcal{B}^{(3)}) \leq \frac{9N}{2A} + 4 + \sum_{p > N/3K} \left\lfloor \frac{N}{p} \right\rfloor \log \frac{\lceil N/3p \rceil}{N/3p}$$

and thus

$$\mathcal{E}_t(\mathcal{B}^{(3)}) \leq \frac{9N}{2A} + 4 + \frac{1}{\log(N/3K)} \sum_{N/3K < p \leq N} f(p/N) \log p$$

where $f : (1/3K, 1] \to \mathbb{R}$ is the piecewise smooth function

$$f(x) := \left\lfloor \frac{1}{x} \right\rfloor \log \frac{\lceil 1/3x \rceil}{1/3x}.$$

Applying Lemma C.3 and a change of variables, we thus have

$$\mathcal{E}_t(\mathcal{B}^{(3)}) \leq \frac{9N}{2A} + 4 + \frac{N}{\log(N/3K)} \int_{1/3K}^1 (1 - \frac{2}{\sqrt{Nx}}) f(x)\, dx + \frac{E(N)}{\log(N/3K)} (f(1/3K+) + f(1) + \|f\|_{\mathrm{TV}}).$$

We discard the $\frac{2}{\sqrt{Nx}}$ term as it gives a negative contribution. Direct numerical calculation (cf. Figure 1) reveals that

$$\int_{1/3K}^1 f(x)\, dx \leq 0.9201$$
$$f(1/3K+) + f(1) + \|f\|_{\mathrm{TV}} \leq 2044$$

and thus

$$\mathcal{E}_t(\mathcal{B}^{(3)}) \leq \frac{9N}{2A} + 4 + \frac{N}{\log(N/3K)} (0.9201 + 2044 \frac{E(N)}{N}).$$

Now we consider the expression

$$\sum_{3<p\leq t/K}\left\lceil \frac{\log N}{\log p}\right\rceil$$

appearing in (3.34). The quantity $\left\lceil \frac{\log N}{\log p}\right\rceil$ equals 1 for $p > \sqrt{t}$ and is at most $\frac{\log N}{\log 5}+1$ for $3 < p \leq \sqrt{t}$, so we have

$$\sum_{3<p\leq t/K}\left\lceil \frac{\log N}{\log p}\right\rceil \leq \pi(t/K) + \frac{\log N}{\log 5}\pi(\sqrt{t}).$$

Appendix A. Powers of 2 and 3

We now obtain good bounds on the quantity κ_L. Clearly κ_L is a non-increasing function of L with $\kappa_1 = \log 2$. The following lemma gives improved control on κ_L for large L:

Lemma A.1. *If n_1, n_2, m_1, m_2 are natural numbers such that $n_1+n_2, m_1+m_2 \geq 1$ and*

$$1 \leq \frac{3^{m_1}}{2^{n_1}}, \frac{2^{n_2}}{3^{m_2}}$$

then

$$\kappa_{\min(2^{n_1+n_2},3^{m_1+m_2})/6} \leq \log\max\left(\frac{3^{m_1}}{2^{n_1}}, \frac{2^{n_2}}{3^{m_2}}\right).$$

Proof. If $\min(2^{n_1+n_2}, 3^{m_1+m_2})/6 \leq t \leq 2^{n_2-1}3^{m_1-1}$, then we have

$$t \leq 2^{n_2-1}3^{m_1-1} \leq \max\left(\frac{3^{m_1}}{2^{n_1}}, \frac{2^{n_2}}{3^{m_2}}\right)t, \tag{A.1}$$

so we are done in this case. Now suppose that $t > 2^{n_2-1}3^{m_1-1}$. If we write $\lceil t\rceil^{\langle 2,3\rangle} = 2^n3^m$ be the smallest 3-smooth number that is at least t, then we must have $n \geq n_2$ or $m \geq m_1$ (or both). Thus at least one of $\frac{2^{n_1}}{3^{m_1}}2^n3^m$ and $\frac{3^{m_2}}{3^{n_2}}2^n3^m$ is an integer, and is thus at most t by construction. This gives (A.1), and the claim follows. □

Some efficient choices of parameters for this lemma are given in Table 2. For instance, $\kappa_{4.5} \leq 0.28768\dots$ and $\kappa_{40.5} \leq 0.16989\dots$.

Remark A.2. It should be unsurprising that the continued fraction convergents $1/1, 2/1, 3/2, 8/5, 19/12, \dots$ to

$$\frac{\log 3}{\log 2} = 1.5849\dots = [1; 1, 1, 2, 2, 3, 1, \dots]$$

are often excellent choices for n_1/m_1 or n_2/m_2, although other approximants such as $5/3$ or $11/7$ are also usable.

Asymptotically, we have logarithmic-type decay:

Lemma A.3 (Baker bound)**.** *We have*

$$\kappa_L \ll \log^{-c} L$$

for all $L \geq 2$ and some absolute constant $c > 0$.

n_1	m_1	n_2	m_2	$\min(2^{n_1+n_2}, 3^{m_1+m_2})/6$	$\log\max(3^{m_1}/2^{n_1}, 2^{n_2}/3^{m_2})$
1	1	**1**	**0**	$1/2 = 0.5$	$\log 2 = 0.69314\dots$
1	**1**	2	1	$2^2/3 = 1.33\dots$	$\log(3/2) = 0.40546\dots$
3	2	**2**	**1**	$3^2/2 = 4.5$	$\log(2^2/3) = 0.28768\dots$
3	2	**5**	**3**	$3^4/2 = 40.5$	$\log(2^5/3^3) = 0.16989\dots$
3	**2**	8	5	$2^{10}/3 = 341.33\dots$	$\log(3^2/2^3) = 0.11778\dots$
11	**7**	8	5	$2^{18}/3 = 87381.33\dots$	$\log(3^7/2^{11}) = 0.06566\dots$
19	12	**8**	**5**	$3^{17}/2 \approx 6.4\times 10^7$	$\log(2^8/3^5) = 0.05211\dots$
19	12	**27**	**17**	$3^{29}/2 \approx 3.4\times 10^{13}$	$\log(2^{27}/3^{17}) = 0.03856\dots$
19	12	**46**	**29**	$3^{41}/2 \approx 1.8\times 10^{19}$	$\log(2^{46}/3^{29}) = 0.02501\dots$

TABLE 2. Efficient parameter choices for Lemma A.1. The parameters used to attain the minimum or maximum are indicated in **boldface**. Note how the number of rows in each group matches the terms $1, 1, 2, 2, 3, \dots$ in the continued fraction expansion.

Proof. From the classical theory of continued fractions, we can find rational approximants

$$\frac{p_{2j}}{q_{2j}} \leq \frac{\log 3}{\log 2} \leq \frac{p_{2j+1}}{q_{2j+1}} \tag{A.2}$$

to the irrational number $\log 3/\log 2$, where the convergents p_j/q_j obey the recursions

$$p_j = b_j p_{j-1} + p_{j-2}; \quad q_j = b_j q_{j-1} + q_{j-2}$$

with $p_{-1} = 1, q = -1 = 0, p_0 = b_0, q_0 = 1$, and

$$[b_0; b_1, b_2, \dots] = [1; 1, 1, 2, 2, 3, 1 \dots]$$

is the continued fraction expansion of $\frac{\log 3}{\log 2}$. Furthermore, $p_{2j+1}q_{2j} - p_{2j}q_{2j+1} = 1$, and hence

$$\frac{\log 3}{\log 2} - \frac{p_{2j}}{q_{2j}} = \frac{1}{q_{2j}q_{2j+1}}. \tag{A.3}$$

By Baker's theorem, $\frac{\log 3}{\log 2}$ is a Diophantine number, giving a bound of the form

$$q_{2j+1} \ll q_{2j}^{O(1)} \tag{A.4}$$

and a similar argument (using $p_{2j+2}q_{2j+1} - p_{2j+1}q_{2j+2} = -1$) gives

$$q_{2j+2} \ll q_{2j+1}^{O(1)}. \tag{A.5}$$

We can rewrite (A.2) as

$$1 \leq \frac{3^{q_{2j}}}{2^{p_{2j}}}, \frac{2^{p_{2j+1}}}{3^{q_{2j+1}}}$$

and routine Taylor expansion using (A.3) gives the upper bounds

$$\frac{3^{q_{2j}}}{2^{p_{2j}}}, \frac{2^{p_{2j+1}}}{3^{q_{2j+1}}} \leq \exp\left(O\left(\frac{1}{q_{2j}}\right)\right).$$

From Lemma A.1 we obtain

$$\kappa_{\min(2^{p_{2j}+p_{2j+1}}, 3^{q_{2j}+q_{2j+1}})/6} \ll \frac{1}{q_{2j}}.$$

The claim then follows from (A.4), (A.5) after optimizing in j.

□

It seems reasonable to conjecture that c can be taken to be arbitrarily close to 1, but this is essentially equivalent to the open problem of determining that irrationality measure of $\log 3/\log 2$ is equal to 2.

Appendix B. Key inequality

We now prove Lemma 4.1. Writing $\nu_p(m) = \sum_{j\geq 1} 1_{p^j|m}$, it suffices to show that

$$0 \leq \sum_{m\leq K;(m,6)=1,p^j|m} \frac{3}{m} - \sum_{m\leq K,p^j|m} \log\frac{m}{m-1} \leq \frac{2}{p^j}$$

for all j. Making the change of variables $m = p^j n$, it suffices to show that

$$0 \leq \sum_{n\leq K'} \frac{3}{n} 1_{(n,6)=1} - p^j \log\frac{p^j n}{p^j n - 1} \leq 2$$

for any $K' > 0$. Using the bound

$$\log\frac{p^j n}{p^j n - 1} = \int_{p^j n-1}^{p^j n} \frac{dx}{x} \in [\frac{1}{p^j n}, \frac{1}{p^j n - 1}]$$

and $p^j \geq 5$, we have

$$\frac{1}{n} \leq p^j \log\frac{p^j n}{p^j n - 1} \leq \frac{1}{n - 0.2}$$

and so it suffices to show that

$$0 \leq \sum_{n\leq K'} \frac{3}{n} 1_{(n,6)=1} - \frac{1}{n-0.2} \leq \sum_{n\leq K'} \frac{3}{n} 1_{(n,6)=1} - \frac{1}{n} \geq 2. \tag{B.1}$$

Since

$$\sum_{n=1}^{\infty} \frac{1}{n-0.2} - \frac{1}{n} = \psi(0.8) - \psi(1) = 0.353473,$$

where ψ here denotes the digamma function rather than the von Mangoldt summatory function, it will suffice to show that

$$0.4 \leq \sum_{n\leq K'} \frac{3}{n} 1_{(n,6)=1} - \frac{1}{n} \geq 2. \tag{B.2}$$

This can be numerically verified for $K' \leq 100$, with substantial room to spare for K' large; see Figure 2. On a block $6a-1 \leq n \leq 6a+4$, the sum is positive:

$$\begin{aligned}\sum_{6a-1\leq n\leq 6a+4} \frac{3}{n} 1_{(n,6)=1} - \frac{1}{n} &= \left(\frac{1}{6a-1} - \frac{1}{6a}\right) + \left(\frac{1}{6a-1} - \frac{1}{6a+2}\right) \\ &\quad + \left(\frac{1}{6a+1} - \frac{1}{6a+3}\right) + \left(\frac{1}{6a+1} - \frac{1}{6a+4}\right) \\ &> 0.\end{aligned}$$

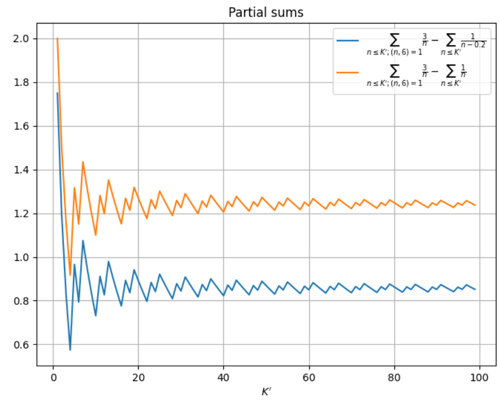

FIGURE 2. A plot of (B.1).

Similarly, on a block $6a-4 \leq n \leq 6a+1$, the sum is negative:

$$\begin{aligned}\sum_{6a-4 \leq n \leq 6a+1} \frac{3}{n} 1_{(n,6)=1} - \frac{1}{n} &= \left(\frac{1}{6a+1} - \frac{1}{6a}\right) + \left(\frac{1}{6a+1} - \frac{1}{6a-2}\right) \\ &\quad + \left(\frac{1}{6a-1} - \frac{1}{6a-3}\right) + \left(\frac{1}{6a-1} - \frac{1}{6a-4}\right) \\ &\quad < 0.\end{aligned}$$

Thus the sum in (B.2) is increasing for $K' = 4$ (6) and decreasing for $K' = 1$ (6), and the inequality for $K' > 100$ is then easily verified from the $K' \leq 100$ data and the triangle inequality

From this and the triangle inequality one can easily establish (B.1) in the remaining ranges $K' \geq 98$.

APPENDIX C. ESTIMATING SUMS OVER PRIMES

Lemma C.1 (Integration by parts)**.** *Let $(y,x]$ be a half-open interval in $(0,+\infty)$. Suppose that one has a function $a\colon \mathbb{N} \to \mathbb{R}$ and a continuous function $f : (y,x] \to \mathbb{R}$ such that*

$$\sum_{y<n\leq z} a_n = \int_z^y f(t)\, dt + C + O_{\leq}(A)$$

for all $y \leq z \leq x$, and some $C \in \mathbb{R}$, $A > 0$. Then, for any function $b : (y,x] \to \mathbb{R}$ of bounded total variation, one has

$$\sum_{y<n\leq x} b(n) a_n = \int_x^y b(t) f(t)\, dt + O_{\leq}(A(|b(y^+)| + |b(x)| + \|b\|_{\mathrm{TV}})), \tag{C.1}$$

where $b(y^+) := \lim_{t\to y^+} b(t)$ denotes the right limit of b at y.

Proof. If, for every natural number $y < n \leq x$, one modifies b to be equal to the constant $b(n)$ in a small neighborhood of n, then one does not affect the left-hand side of (C.1) or increase the total variation of b, while only modifying the integral in (C.1) by an arbitrarily small amount. Hence, by the usual limiting argument, we may assume without loss of generality that b is locally constant at each such n. If we define the function $g : (y, x] \to \mathbb{R}$ by

$$g(z) := \sum_{y<n\leq z} a_n - \int_z^y f(u)\, du - C$$

then g has jump discontinuities at the natural numbers, but is otherwise continuously differentiable, and is also bounded uniformly in magnitude by A. We can then compute the Riemann–Stieltjes integral

$$\int_{(y,x]} b\, dg = \sum_{y<n\leq x} b(n)a_n - \int_y^x f(t)b(t)\, dt.$$

Since the discontinuities of g and b do not coincide, we may integrate by parts to obtain

$$\int_{(y,x]} b\, dg = b(x)g(x) - b(y^+)g(y^+) - \int_{(y,x]} g db.$$

The left-hand side is $O_\leq(A(|b(y^+)| + |b(x)| + \|b\|_{\mathrm{TV}}))$, and the claim follows. □

By combining the above lemma with the prime numebr theorem with classical error term, we obtain

Corollary C.2. *Under the above hypotheses with* $1 \leq y \leq x$*, we have*

$$\sum_{y<p\leq x} b(p) = \int_y^x b(t)\frac{dt}{\log t} + O\left((|b(y^+)| + |b(x)| + \|b\|_{\mathrm{TV}})x \exp(-c\sqrt{\log x})\right)$$

for some absolute constant $c > 0$*.*

We have the following explicit version of the above estimate, where it is convenient to apply the Chebyshev weighting of assigning each prime p a weight of $\log p$.

Lemma C.3 (Buthe effective prime number theorem)**.** *Under the above hypotheses with* $1423 \leq y \leq x$*, one has*

$$\sum_{y<p\leq x} b(p)\log p = \int_y^x b(t)(1 - \frac{2}{\sqrt{t}})\, dt + O\left((|b(y^+)| + |b(x)| + \|b\|_{\mathrm{TV}})E(x)\right)$$

where

$$E(x) := 0.95\sqrt{x} + \frac{\sqrt{x}}{8\pi}\log x(\log x - 3)1_{x\geq 10^{19}} + 1.11742\times 10^{-8}x1_{x\geq e^{45}}.$$

Proof. Observe that E is monotone non-decreasing. Thus by Lemma C.1, it will suffice to show that

$$\sum_{p\leq x}\log p = x - \sqrt{x} + O_\leq(E(x)) = \int_0^x (1 - \frac{2}{\sqrt{t}})\, dt + O_\leq(E(x))$$

where

$$E(x) := 0.95\sqrt{x} + \frac{\sqrt{x}}{8\pi}\log x(\log x - 3)1_{x\geq 10^{19}} + 1.11742\times 10^{-8}x1_{x\geq e^{45}}.$$

For $1423 \leq x \leq 10^{19}$, this follows from [2, Theorem 2]. In the range For $10^{19} \leq x \leq 10^{21} \approx e^{48.35}$, we use the bound

$$\psi(x) = x + O_{\leq}(\frac{\sqrt{x}}{8\pi} \log x(\log x - 3))$$

which was established for $5000 \leq x \leq 10^{21}$ in [1, (7.3)], where $\psi(x) := \sum_{n \leq x} \Lambda(n)$ is the usual von Mangoldt summatory function. To use this, we apply [1, (6.10), (6.11)] to conclude that

$$\sum_{p \leq x} \log p = \psi(x) - \psi(\sqrt{x}) + O_{\leq}(1.03883(x^{1/3} + x^{1/5} + 2(\log x)x^{1/13})).$$

From [5, Theorems 10,12] we have

$$\psi(\sqrt{x}) = \sqrt{x} + O_{\leq}(0.18\sqrt{x}).$$

Since

$$0.18\sqrt{x} + 1.03883(x^{1/3} + x^{1/5} + 2(\log x)x^{1/13}) \leq 0.95\sqrt{x}$$

for $x \geq 10^{19}$, the claim follows.

Finally, in the range $x \geq 10^{21}$, we see from [1, Theorem 1, Table 1] that one has the bound

$$\psi(x) = x + O_{\leq}(1.11742 \times 10^{-8} x)$$

for $x \geq e^{45} \approx 10^{19.54}$, and the claim follows by repeating the previous arguments. □

We remark that assuming the Riemann hypothesis, the final term in the definition of $E(x)$ may be deleted, since [1, (7.3)] then holds for all $x \geq 5000$.

REFERENCES

[1] J. Büthe, *Estimating $\pi(x)$ and related functions under partial RH assumptions*, Math. Comp., 85(301), 2483–2498, Jan. 2016.

[2] J. Büthe, *An analytic method for bounding $\psi(x)$* . Math. Comp., **87** (312), 1991–2009.

[3] R. K. Guy, J. L. Selfridge, *Factoring factorial n*, Amer. Math. Monthly **105** (1998) 766–767.

[4] H. Robbins, *A Remark on Stirling's Formula*, Amer. Math. Monthly **62** (1955) 26–29.

[5] J. Rosser, L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. 6 (1962), 64–94.

[6] T. Tao, *Decomposing factorials into bounded factors*, preprint, 2025. `https://arxiv.org/abs/2503.20170`