# New bounds for the sum of the first $n$ prime numbers

Christian Axler

June 23, 2016

**Abstract**

In this paper we establish a general asymptotic formula for the sum of the first $n$ prime numbers, which leads to a generalization of the most accurate asymptotic formula given by Massias and Robin. Further we prove a series of results concerning Mandl's inequality on the sum of the first $n$ prime numbers. We use these results to find new explicit estimates for the sum of the first $n$ prime numbers, which improve the currently best known estimates.

## 1 Introduction

In this paper, we study the sum of the first $n$ prime numbers. At the beginning of the 20th century, Landau [12] showed that

$$\sum_{k \leq n} p_k \sim \frac{n^2}{2} \log n \qquad (n \to \infty).$$

Using a result of Robin [16], Massias and Robin [15, p. 217] found the more accurate asymptotic formula

$$\sum_{k \leq n} p_k = \frac{n^2}{2} \left( \log n + \sum_{k=0}^{m} \frac{B_{k+1}(\log \log n)}{\log^k n} \right) + O\left( \frac{n^2 (\log \log n)^{m+1}}{\log^{m+1} n} \right),$$

where $m \in \mathbb{N}$ and the polynomials are given by the formulas $B_0(x) = 1$ and

$$B'_{k+1}(x) = B'_k(x) - (k-1)B_k(x). \tag{1.1}$$

Further, they gave

$$\sum_{k \leq n} p_k = \frac{n^2}{2} \left( \log n + \log \log n - \frac{3}{2} + \frac{\log \log n - 5/2}{\log n} \right) + O\left( \frac{n^2 (\log \log n)^2}{\log^2 n} \right). \tag{1.2}$$

Unfortunately, the formula (1.1) does not yields any of the polynomials $B_1, \ldots, B_{m+1}$. So, the asymptotic formula given in (1.2) is currently the most accurate for the sum of the first $n$ primes. Our first goal in this paper is to derive the following general asymptotic formula for the sum of the first $n$ primes.

**Theorem 1.1** (See Corollary 2.8)**.** *There exist unique monic polynomials $T_s \in \mathbb{Q}[x]$, where $1 \leq s \leq m$ and $\deg(T_s) = s$, such that*

$$\sum_{k \leq n} p_k = \frac{n^2}{2} \left( \log n + \log \log n - \frac{3}{2} + \sum_{s=1}^{m} \frac{(-1)^{s+1} T_s(\log \log n)}{s \log^s n} \right) + O(nc_m(n)).$$

*The polynomials $T_s$ can be computed explicitly. In particular, $T_1(x) = x - 5/2$ and $T_2(x) = x^2 - 7x + 29/2$.*

The motivation for this paper is an inequality conjectured by Mandl concerning an upper bound for the sum of the first $n$ prime numbers, namely that

$$\frac{np_n}{2} \geq \sum_{k \leq n} p_k \tag{1.1}$$

for every $n \geq 9$, where $p_n$ is the $n$th prime number. This inequality originally appeared in a paper of Rosser and Schoenfeld [19, p. 1] from 1975 without proof. In his thesis, Dusart [7] used the identity

$$\sum_{k \leq n} p_k = np_n - \int_2^{p_n} \pi(x)\, dx,$$

where $\pi(x)$ denotes the number of primes $\leq x$, and explicit estimates for $\pi(x)$ to prove that (1.1) holds for every $n \geq 9$. One of the goals of this paper is to study the sequence $(B_n)_{n \in \mathbb{N}}$ with

$$B_n = \frac{np_n}{2} - \sum_{k \leq n} p_k$$

more detailed (see also [22]). For this purpose, we derive the following asymptotic formula for $B_n$.

**Theorem 1.2** (see Corollary 3.2)**.** *There exist uniquely determined polynomials $V_2, \ldots, V_m \in \mathbb{Q}[x]$ with $\deg(V_s) = s - 1$ and positive leading coefficient, so that*

$$B_n = \frac{n^2}{4} + \frac{n^2}{4 \log n} + \frac{n^2}{2} \sum_{s=2}^{m} \frac{(-1)^{s+1} V_s(\log \log n)}{s \log^s n} + O\left(\frac{n^2 (\log \log n)^{m+1}}{\log^{m+1} n}\right).$$

*The polynomials $V_s$ can be computed explicitly. In particular, $V_2(x) = x - 7/2$.*

Since it is still difficult to compute $B_n$ for large $n$, we are interested in explicit estimates for $B_n$. By (1.1), we get that $B_n > 0$ for every $n \geq 9$. Hassani [11, Corollary 1.5] found that the inequality

$$B_n > \frac{n^2}{12}$$

holds for every $n \geq 10$. Up to now, the sharpest lower bound for $B_n$ is due to Sun [23]. In 2012, he proved that

$$B_n > \frac{n^2}{4} \tag{1.2}$$

for every $n \geq 417$. According to Theorem 1.2, we improve the inequality (1.2) by showing the following

**Theorem 1.3** (see Corollary 4.8)**.** *If $n \geq 348247$, then*

$$B_n > \frac{n^2}{4} + \frac{n^2}{4 \log n} - \frac{n^2(\log \log n - 2.1)}{4 \log^2 n}.$$

Theorem 1.2 yields that for each $\varepsilon > 0$ there exists an $n_0(\varepsilon) \in \mathbb{N}$, so that

$$B_n < \frac{n^2}{4} + \frac{n^2}{4 \log n} - \frac{n^2(\log \log n - (7/2 + \varepsilon))}{4 \log^2 n} \tag{1.3}$$

for every $n \geq n_0(\varepsilon)$. In view of (1.3), we give the following explicit estimate for $B_n$, which improves the only known upper bound for $B_n$ given by Hassani [11, Corollary 1.5], namely that for every $n \geq 2$,

$$B_n < \frac{9n^2}{4}. \tag{1.4}$$

**Theorem 1.4** (See Corollary 4.10)**.** *If $n \geq 26220$, then*

$$B_n < \frac{n^2}{4} + \frac{n^2}{4 \log n} - \frac{n^2(\log \log n - 5.22)}{4 \log^2 n}.$$

Since it is still difficult to compute the sum of the first $n$ primes for large $n$, we are interested in explicit estimates for this sum. In 1989, Massias, Nicolas and Robin [14, Lemma 4] showed that the upper bounds

$$\sum_{k \leq n} p_k \leq \frac{n^2}{2} \left(\log n + \log \log n - \frac{3}{2} + \frac{1.866 \log \log n}{\log n}\right)$$

and

$$\sum_{k \leq n} p_k \leq \frac{n^2}{2}(\log n + \log\log n - 1.4416)$$

hold for every $n \geq 3688$. Massias and Robin [15, Théorème D(vi)] improved these inequalities by showing that

$$\sum_{k \leq n} p_k \leq \frac{n^2}{2}\left(\log n + \log\log n - \frac{3}{2} + \frac{1.805 \log\log n}{\log n}\right)$$

for every $n \geq 18$ and that

$$\sum_{k \leq n} p_k \leq \frac{n^2}{2}(\log n + \log\log n - 1.463)$$

for every $n \geq 779$. Further, Massias and Robin [15, Théorème A(vi)] proved under the assumption that the Riemann hypothesis is true that

$$p_n \leq n\left(\log n + \log\log n - 1 + \frac{\log\log n - 1.8}{\log n}\right) \tag{1.3}$$

for every $n \geq 27076$ und deduced [15, Théorème D(viii)] under the assumption of (1.3), that the inequality

$$\sum_{k \leq n} p_k \leq \frac{n^2}{2}\left(\log n + \log\log n - \frac{3}{2} + \frac{\log\log n - 2.29}{\log n}\right) \tag{1.4}$$

holds for every $n \geq 10134$. In 1998, Dusart [7, Théoròme 1.7] proved that (1.3) holds unconditionally for every $n \geq 27076$, which implies that the inequality (1.4) holds for every $n \geq 10134$. In this paper, we find the following improvement of (1.4).

**Theorem 1.5** (See Corollary 5.2)**.** *For every $n \geq 355147$, we have*

$$\sum_{k \leq n} p_k < \frac{n^2}{2}\left(\log n + \log\log n - \frac{3}{2} + \frac{\log\log n - 5/2}{\log n} - \frac{(\log\log n)^2 - 7\log\log n + 12.373}{2\log^2 n}\right).$$

According to lower bounds for the sum of the first $n$ primes, Massias and Robin [15, Théorème C(vii)] proved that the inequality

$$\sum_{k \leq n} p_k \geq \frac{n^2}{2}(\log n + \log\log n - 1.5034)$$

holds for every $n \geq 127042$ and that

$$\sum_{k \leq n} p_k \geq \frac{n^2}{2}\left(\log n + \log\log n - \frac{3}{2} - \frac{3.568}{\log n}\right)$$

for every $n \geq 2$. Further, they found that

$$\sum_{k \leq n} p_k \geq \frac{n^2}{2}\left(\log n + \log\log n - \frac{3}{2}\right) \tag{1.5}$$

for every $305494 \leq n \leq e^{530}$. Under the assumption that the Riemann hypothesis is true, they the inequality (1.5) holds for every $n \geq 305494$. In 1998, Dusart [7, Lemme 1.7] proved that the inequality (1.5) holds unconditionally for every $n \geq 305494$. We improve this inequality by showing the following

**Theorem 1.6** (See Corollary 5.5)**.** *For every $n \geq 2$, we have*

$$\sum_{k \leq n} p_k > \frac{n^2}{2}\left(\log n + \log\log n - \frac{3}{2} + \frac{\log\log n - 5/2}{\log n} - \frac{(\log\log n)^2 - 7\log\log n + 17.067}{2\log^2 n}\right).$$

In the last section we derive new explicit estimates for the step function

$$S(x) = \sum_{p \leq x} p,$$

which improve the currently best known estimates for this step function.

## 2 An asymptotic formula for $\sum_{k \leq n} p_k$

To prove a general asymptotic formula for the sum of the first $n$ primes, we introduce the following definition.

**Definition.** Let $s, i, j, r \in \mathbb{N}_0$ with $j \geq r$. We define the integers $b_{s,i,j,r} \in \mathbb{Z}$ as follows:

- If $j = r = 0$, then
$$b_{s,i,0,0} = 1. \tag{2.1}$$
- If $j \geq 1$, then
$$b_{s,i,j,j} = b_{s,i,j-1,j-1} \cdot (-i + j - 1). \tag{2.2}$$
and
$$b_{s,i,j,0} = b_{s,i,j-1,0} \cdot (s + j - 1). \tag{2.3}$$
- If $j > r \geq 1$, then
$$b_{s,i,j,r} = b_{s,i,j-1,r} \cdot (s + j - 1) + b_{s,i,j-1,r-1} \cdot (-i + r - 1). \tag{2.4}$$

**Proposition 2.1.** *If $r > i$, then*
$$b_{s,i,j,r} = 0.$$

*Proof.* Let $r > i$. From (2.2), it follows that
$$b_{s,i,i+1,i+1} = 0 \tag{2.5}$$
and hence
$$b_{s,i,k,k} = 0 \tag{2.6}$$
for every $k \geq i + 1$. We use (2.4) repeatedly and (2.5) to get
$$b_{s,i,k,i+1} = b_{s,i,i+1,i+1} \cdot (s + k - 1) \cdot \ldots (s + (i + 2) - 1) = 0$$
for every $k \geq i + 2$. By using (2.5), it follows that
$$b_{s,i,k,i+1} = 0 \tag{2.7}$$
for every $k \geq i + 1$. Next, we prove by induction that
$$b_{s,i,k,i+n} = 0 \tag{2.8}$$
for every $n \in \mathbb{N}$ and every $k \geq i + n$. If $k = i + n$, then $b_{s,i,k,i+n} = 0$ by (2.6). So, it suffices to prove (2.8) for every $n \in \mathbb{N}$ and every $k \geq i + n + 1$. If $n = 1$, the claim follows from (2.7). Now we write $k = i + n + t$ with an arbitrary $t \in \mathbb{N}$. By (2.4) and the induction hypothesis, we obtain
$$b_{s,i,t+i+n,i+n} = b_{s,i,i+n,i+n} \cdot (s + (t + n + i) - 1) \cdot \ldots \cdot (s + (i + n + 1) - 1).$$
Since $b_{s,i,i+n,i+n} = 0$ by (2.6), we get $b_{s,i,k,i+n} = 0$ and the result follows. □

Let $m \in \mathbb{N}$. To prove a more accurate asymptoic formula for the sum of the first $n$ primes, we first note a result of Cipolla [5] from 1902 concerning an asymptotic formula for the $n$th prime number.

**Lemma 2.2** (Cipolla, [5])**.** *There exist uniquely determined computable rational numbers $a_{is} \in \mathbb{Q}$ with $a_{ss} = 1$ for every $1 \leq s \leq m$, such that*
$$p_n = n \left( \log n + \log\log n - 1 + \sum_{s=1}^{m} \frac{(-1)^{s+1}}{s \log^s n} \sum_{i=0}^{s} a_{is} (\log\log n)^i \right) + O\left( \frac{n(\log\log n)^{m+1}}{\log^{m+1} n} \right).$$

**Corollary 2.3.** *There exist uniquely determined monic polynomials $R_1, \ldots, R_m \in \mathbb{Q}[x]$ with $\deg(R_k) = k$, so that*

$$p_n = n\left(\log n + \log\log n - 1 + \sum_{k=1}^{m} \frac{(-1)^{k+1} R_k(\log\log n)}{k \log^k n}\right) + O\left(\frac{n(\log\log n)^{m+1}}{\log^{m+1} n}\right).$$

*The polynomials $R_k$ can be computed explicitly. In particular, $R_1(x) = x - 2$ and $R_2(x) = x^2 - 6x + 11$.*

The next two lemmata include the *logarithmic integral* $\mathrm{li}(x)$ defined for every real $x \geq 2$ by

$$\mathrm{li}(x) = \int_0^x \frac{dt}{\log t} = \lim_{\varepsilon \to 0} \left\{ \int_0^{1-\varepsilon} \frac{dt}{\log t} + \int_{1+\varepsilon}^x \frac{dt}{\log t} \right\} \approx \int_2^x \frac{dt}{\log t} + 1.04516\ldots.$$

**Lemma 2.4.** *For every $x > 1$ and every $n \in \mathbb{N}$, we have*

$$\mathrm{li}(x) = \sum_{j=1}^{n} \frac{(j-1)!}{\log^j x} + O\left(\frac{x}{\log^{n+1} x}\right).$$

*Proof.* Integration by parts. □

**Lemma 2.5.** *Let $x, a \in \mathbb{R}$ be such that $x \geq a \geq 2$. Then*

$$\int_a^x \frac{t\,dt}{\log t} = \mathrm{li}(x^2) - \mathrm{li}(a^2).$$

*Proof.* See Dusart [7, Lemme 1.6]. □

In the following lemma, we compare the integral of a function with its partial sum.

**Lemma 2.6.** *Let $k, k_0 \in \mathbb{N}$ be such that $k \geq k_0$ and let $f$ be a continuous function on $[k_0, \infty)$ which is non-negativ and increasing on $[k, \infty)$. Then*

$$\sum_{j=k_0}^{n} f(j) = \int_{k_0}^{n} f(x)\,dx + O(f(n)).$$

*Proof.* We estimate the integral by upper and lower sums. □

Now, we set

$$c_m(n) = \frac{n(\log\log n)^{m+1}}{\log^{m+1} n}$$

and

$$g(n) = \log n + \log\log n - \frac{3}{2}.$$

Further, let

$$h_m(n) = \sum_{j=1}^{m} \frac{(j-1)!}{2^j \log^j n}.$$

Then, the following theorem yields a general asymptotic formula for the sum of the first $n$ primes.

**Theorem 2.7.** *For every $n \in \mathbb{N}$,*

$$\sum_{k \leq n} p_k = \frac{n^2}{2}\left(g(n) - h_m(n) + \sum_{s=1}^{m} \frac{(-1)^{s+1}}{s \log^s n} \sum_{i=0}^{s} a_{is} \sum_{j=0}^{m-s} \sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r}(\log\log n)^{i-r}}{2^j \log^j n}\right) + O(nc_m(n)).$$

*Proof.* We set

$$\tau(x) = x\left(\log x + \log\log x - 1 + \sum_{s=1}^{m} \frac{(-1)^{s+1}}{s\log^s x}\sum_{i=0}^{s} a_{is}(\log\log x)^i\right).$$

By Lemma 2.2, we obtain

$$p_k = \tau(k) + O(c_m(k)).$$

Using $\tau(n) \sim n\log n$ as $n \to \infty$ and Lemma 2.6, we get

$$\sum_{k\le n} p_k = \sum_{k=3}^{n} \tau(k) + O(nc_m(n)) = \int_3^n \tau(x)\,dx + O(nc_m(n)). \tag{2.9}$$

First, we integrate the first three terms of $\tau(x)$. We have

$$\int_3^n (x\log x - x)\,dx = \frac{n^2\log n}{2} - \frac{3n^2}{4} + O(1).$$

Next, using integration by parts, Lemma 2.4 and Lemma 2.5, we get

$$\int_3^n x\log\log x\,dx = \frac{n^2\log\log n}{2} - \frac{n^2}{2}h_m(n) + O\left(\frac{n^2}{\log^{m+1} n}\right).$$

Hence, by (2.9),

$$\sum_{k\le n} p_k = \frac{n^2}{2}(g(n) - h_m(n)) + \sum_{s=1}^{m}\frac{(-1)^{s+1}}{s}\sum_{i=0}^{s} a_{is}\int_3^n \frac{x(\log\log x)^i}{\log^s x}\,dx + O(nc_m(n)). \tag{2.10}$$

Now let $1 \le s \le m$ and $0 \le i \le s$. We prove by induction that for every $t \in \mathbb{N}_0$,

$$\int_3^n \frac{x(\log\log x)^i}{\log^s x}\,dx = \sum_{j=0}^{t}\sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r}n^2(\log\log n)^{i-r}}{2^{j+1}\log^{s+j} n} + \int_3^n \sum_{r=0}^{\min\{i,t+1\}} \frac{b_{s,i,t+1,r}x(\log\log x)^{i-r}}{2^{t+1}\log^{s+t+1} x}\,dx + O(1). \tag{2.11}$$

Integration by parts gives

$$\int_3^n \frac{x(\log\log x)^i}{\log^s x}\,dx = \frac{n^2(\log\log n)^i}{2\log^s n} - \frac{i}{2}\int_3^n \frac{x(\log\log x)^{i-1}}{\log^{s+1} x}\,dx + \frac{s}{2}\int_3^n \frac{x(\log\log x)^i}{\log^{s+1} x}\,dx + O(1),$$

so (2.11) holds for $t = 0$. By induction hypothesis, we get

$$\int_3^n \frac{x(\log\log x)^i}{\log^s x}\,dx = \sum_{j=0}^{t-1}\sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r}n^2(\log\log n)^{i-r}}{2^{j+1}\log^{s+j} n} + \sum_{r=0}^{\min\{i,t\}}\int_3^n \frac{b_{s,i,t,r}x(\log\log x)^{i-r}}{2^t\log^{s+t} x}\,dx + O(1).$$

Using integration by parts of the integral on the right hand side, we obtain

$$\int_3^n \frac{x(\log\log x)^i}{\log^s x}\,dx = \sum_{j=0}^{t}\sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r}n^2(\log\log n)^{i-r}}{2^{j+1}\log^{s+j} n} + \int_3^n \sum_{r=0}^{\min\{i,t\}} \frac{b_{s,i,t,r}((s+t)(\log\log x)^{i-r} - (i-r)(\log\log x)^{i-r-1})x}{2^{t+1}\log^{s+t+1} x}\,dx + O(1). \tag{2.12}$$

Since

$$\sum_{r=0}^{\min\{i,t\}} b_{s,i,t,r}((s+t)(\log\log x)^{i-r} - (i-r)(\log\log x)^{i-r-1})$$

$$= \sum_{r=1}^{\min\{i,t\}} (b_{s,i,t,r}(s+t) - b_{s,i,t,r-1}(i-(r-1)))(\log\log x)^{i-r} + b_{s,i,t,0}(s+t)(\log\log x)^{i}$$
$$- b_{s,i,t,\min\{i,t\}}(i-\min\{i,t\})(\log\log x)^{i-(\min\{i,t\}+1)},$$

we can use (2.3) and (2.4) to get

$$\sum_{r=0}^{\min\{i,t\}} b_{s,i,t,r}((s+t)(\log\log x)^{i-r} - (i-r)(\log\log x)^{i-r-1})$$
$$= \sum_{r=0}^{\min\{i,t\}} b_{s,i,t+1,r}(\log\log x)^{i-r} - b_{s,i,t,\min\{i,t\}}(i-\min\{i,t\})(\log\log x)^{i-(\min\{i,t\}+1)}. \quad (2.13)$$

It is easy to see that

$$-b_{s,i,t,\min\{i,t\}}(i-\min\{i,t\}) = b_{s,i,t+1,\min\{i,t\}+1}.$$

Hence, by (2.13), we obtain

$$\sum_{r=0}^{\min\{i,t\}} b_{s,i,t,r}((s+t)(\log\log x)^{i-r} - (i-r)(\log\log x)^{i-r-1}) = \sum_{r=0}^{\min\{i,t\}+1} b_{s,i,t+1,r}(\log\log x)^{i-r}.$$

Since $b_{s,i,t+1,i+1} = 0$ for $t \geq i$ by Proposition 2.1, it follows that

$$\sum_{r=0}^{\min\{i,t\}} b_{s,i,t,r}((s+t)(\log\log x)^{i-r} - (i-r)(\log\log x)^{i-r-1}) = \sum_{r=0}^{\min\{i,t+1\}} b_{s,i,t+1,r}(\log\log x)^{i-r}.$$

Using (2.12), we obtain (2.11). Now we choose $t = m-s$ in (2.11) and we get

$$\int_3^n \frac{x(\log\log x)^i}{\log^s x}\,dx = \sum_{j=0}^{m-s}\sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r}n^2(\log\log n)^{i-r}}{2^{j+1}\log^{s+j} n} + O\left(\frac{n^2(\log\log n)^i}{\log^{m+1} n}\right).$$

We substitute this in (2.10) to obtain

$$\sum_{k\leq n} p_k = \frac{n^2}{2}(g(n)-h_m(n)) + \sum_{s=1}^{m}\frac{(-1)^{s+1}}{s}\sum_{i=0}^{s} a_{is}\sum_{j=0}^{m-s}\sum_{r=0}^{\min\{i,j\}}\frac{b_{s,i,j,r}n^2(\log\log n)^{i-r}}{2^{j+1}\log^{s+j} n}$$
$$+ O\left(\sum_{s=1}^{m}\frac{(-1)^{s+1}}{s}\sum_{i=0}^{s} a_{is}\frac{n^2(\log\log n)^i}{\log^{m+1} n}\right) + O(nc_m(n))$$

and our theorem is proved. □

The following corollary generalizes the asymptotic formula (1.2).

**Corollary 2.8.** *There exist uniquely determined monic polynomials $T_s \in \mathbb{Q}[x]$, where $1 \leq s \leq m$ and $\deg(T_s) = s$, such that*

$$\sum_{k\leq n} p_k = \frac{n^2}{2}\left(\log n + \log\log n - \frac{3}{2} + \sum_{s=1}^{m}\frac{(-1)^{s+1}T_s(\log\log n)}{s\log^s n}\right) + O(nc_m(n)).$$

*The polynomials $T_s$ can be computed explicitly. In particular, $T_1(x) = x - 5/2$ and $T_2(x) = x^2 - 7x + 29/2$.*

*Proof.* Since $a_{ss} = 1$ and $b_{s,s,0,0} = 1$, the first claim follows from Theorem 2.7. Let $m = 2$. By Cipolla [5], we have $a_{01} = -2, a_{11} = 1, a_{02} = 11, a_{12} = -6$ and $a_{22} = 1$. Further, we use the formulas (2.1)–(2.4) to compute the integers $b_{s,i,j,r}$. Then, using Theorem 2.7, we obtain the polynomials $T_1$ and $T_2$. □

*Remark.* The first part of Corollary 2.8 was already proved by Sinha [20, Theorem 2.3] in 2010. Due to a calculation error, he gave the polynomials $T_1(x) = x - 3$ and $T_2(x) = x^2 - 7x + 27/2$.

## 3 An asymptotic formula for $B_n$

Let $m \in \mathbb{N}$. We obtain the following asymptotic formula for the sum of the first $n$ prime numbers.

**Theorem 3.1.** *We have*

$$B_n = \frac{n^2}{4} + \frac{n^2}{2} h_m(n) + \frac{n^2}{2} \sum_{s=1}^{m} \frac{(-1)^s}{s \log^s n} \sum_{i=0}^{s} a_{is} \sum_{j=1}^{m-s} \sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r} (\log\log n)^{i-r}}{2^j \log^j n} + O(nc_m(n)).$$

*Proof.* First we multiply the asymptotic formula in Lemma 2.2 with $n/2$ and then we subtract the asymptotic formula in Theorem 2.7 to get

$$B_n = \frac{n^2}{4} + \frac{n^2}{2} h_m(n) + \frac{n^2}{2} \sum_{s=1}^{m} \frac{(-1)^{s+1}}{s \log^s n} \sum_{i=0}^{s} a_{is} \left( (\log\log n)^i - \sum_{j=0}^{m-s} \sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r} (\log\log n)^{i-r}}{2^j \log^j n} \right) + O(nc_m(n)).$$

Since

$$(\log\log n)^i - \sum_{j=0}^{m-s} \sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r} (\log\log n)^{i-r}}{2^j \log^j n} = -\sum_{j=1}^{m-s} \sum_{r=0}^{\min\{i,j\}} \frac{b_{s,i,j,r} (\log\log n)^{i-r}}{2^j \log^j n},$$

our theorem is proved. □

Using Corollary 2.3 and 2.8, we obtain the following corollary.

**Corollary 3.2.** *There exist uniquely determined polynomials $V_2, \ldots, V_m \in \mathbb{Q}[x]$ with $\deg(V_s) = s-1$ and positive leading coefficient, so that*

$$B_n = \frac{n^2}{4} + \frac{n^2}{4 \log n} + \frac{n^2}{2} \sum_{s=2}^{m} \frac{(-1)^{s+1} V_s(\log\log n)}{s \log^s n} + O(nc_m(n)).$$

*The polynomials $V_s$ can be computed explicitly. In particular, $V_2(x) = x - 7/2$.*

*Remark.* Based on two calculation errors, Sinha [20, Lemma 3.1] gave the polynomial $V_2(x) = x - 49/2$.

## 4 Explicit estimates for $B_n$

In this section we give some explicit estimates for $B_n$.

### 4.1 Some auxiliaries lemmata

In order to find new explicit estimates for $B_n$, we first note some useful lemmata concerning upper and lower bound for the severel prime functions. We start with upper bounds for the prime counting function $\pi(x)$ given in [4, Corollary 3.4].

**Lemma 4.1.** *For every $x \geq 21.95$, we have*

$$\pi(x) < \frac{x}{\log x - 1 - \frac{1}{\log x} - \frac{3.35}{\log^2 x} - \frac{12.65}{\log^3 x} - \frac{89.6}{\log^4 x}}. \tag{4.1}$$

*If $x \geq 14.36$, then*

$$\pi(x) < \frac{x}{\log x - 1 - \frac{1}{\log x} - \frac{3.35}{\log^2 x} - \frac{15.43}{\log^3 x}} \tag{4.2}$$

*and for every $x \geq 9.25$, we have*

$$\pi(x) < \frac{x}{\log x - 1 - \frac{1}{\log x} - \frac{3.83}{\log^2 x}}. \tag{4.3}$$

*If $x \geq 5.43$, then*

$$\pi(x) < \frac{x}{\log x - 1 - \frac{1.17}{\log x}}. \tag{4.4}$$

In [4, Theorem 1.4 and Corollary 3.5] the following lower bounds for $\pi(x)$ are found.

**Lemma 4.2.** *If $x \geq x_0$, then*

$$\pi(x) > \frac{x}{\log x - 1 - \frac{1}{\log x} - \frac{a_1}{\log^2 x} - \frac{a_2}{\log^3 x} - \frac{a_3}{\log^4 x} - \frac{a_4}{\log^5 x} - \frac{a_5}{\log^6 x}},$$

*where*

| | | | | | |
|---|---|---|---|---|---|
| $a_1$ | 2.65 | 2.65 | 2.65 | 2.65 | 0 |
| $a_2$ | 13.35 | 13.35 | 13.35 | 4.6 | 0 |
| $a_3$ | 70.3 | 70.3 | 5 | 0 | 0 |
| $a_4$ | 455.6275 | 69 | 0 | 0 | 0 |
| $a_5$ | 3404.4225 | 0 | 0 | 0 | 0 |
| $x_0$ | 1332479531 | 909050897 | 374123969 | 38168363 | 468049 |

In 1962, Rosser and Schoenfeld [18, Corollary 1] already found the following lower bound for $\pi(x)$.

**Lemma 4.3** (Rosser and Schoenfeld, [18])**.** *For every $x \geq 17$, we have*

$$\pi(x) > \frac{x}{\log x}.$$

In 2010, Dusart [9, Proposition 6.6] gave the following upper bound for the $n$th prime number.

**Lemma 4.4** (Dusart, [9])**.** *For every $n \geq 688383$, we have*

$$p_n \leq n \left( \log n + \log\log n - 1 + \frac{\log\log n - 2}{\log n} \right).$$

The next lemma gives an inequality between $\log n$ and $\log p_n$.

**Lemma 4.5.** *For every $n \geq 255$, we have*

$$\log n \geq 0.75 \log p_n.$$

*Proof.* First, we note an upper bound for $p_n$ found by Rosser and Schoenfeld [18, p. 69], namely that

$$p_n < n(\log n + \log\log n)$$

for every $n \geq 6$. Further, we have $x \geq 8\log^3 x$ for every $x \geq 4913$. Hence,

$$n^4 \geq n^3(2\log n)^3 \geq n^3(\log n + \log\log n)^3 \geq p_n^3$$

for every $n \geq 4913$. We use a computer to check the required inequality holds for the remaining cases. ☐

## 4.2 New lower bounds for $B_n$

The goal in this subsection is to improve the lower bound (1.2). We start with the following lemma.

**Lemma 4.6.** *For every $n \in \mathbb{N}$, we have*

$$927p_n^2 \log p_n + 5620p_n^2 > 756np_n \log^2 p_n + 117n^2 \log^3 p_n.$$

*Proof.* First, we consider the case $n \geq 6077$. Then, we have $p_n \geq 60184$ and by [7],

$$p_n > n(\log p_n - 1.1). \tag{4.5}$$

Hence, we obtain

$$927p_n^2 \log p_n > 756np_n \log^2 p_n + 171np_n \log^2 p_n - 1019.7np_n \log p_n. \tag{4.6}$$

Again, by using (4.5), we get that

$$171np_n\log^2 p_n > 54n^2\log^3 p_n + 117n^2\log^3 p_n - 188.1n^2\log^2 p_n$$

and in combination with (4.6), we have

$$927p_n^2\log p_n > 756np_n\log^2 p_n + 54n^2\log^3 p_n + 117n^2\log^3 p_n - 188.1n^2\log^2 p_n - 1019.7np_n\log p_n. \quad (4.7)$$

Since $\log p_n \geq 11 > 188.1/54$, we get, by (4.7),

$$927p_n^2\log p_n > 756np_n\log^2 p_n + 117n^2\log^3 p_n - 1019.7np_n\log p_n. \quad (4.8)$$

We have $\log p_n \geq 1.1/0.75$. Hence, we get $p_n > n(\log p_n - 1.1) \geq 0.25n\log p_n$ and therefore $5620p_n^2 > 1019.7np_n\log p_n$. In combination with (4.8), the claim follows for every $n \geq 6077$. For every $1 \leq n \leq 6076$, we check the required inequality with a computer. □

Now we set

$$\rho(n) = \frac{2.1\log^2 n}{4\log^2 p_n} + \frac{\log^2 p_n\log^2 n + 16.3\log^2 n - \log^3 p_n\log n + \log^3 p_n\log\log n}{4\log^3 p_n}$$

to obtain the following theorem.

**Theorem 4.7.** *For every $n \geq 842857$, we have*

$$B_n > \frac{n^2}{4} + \frac{n^2}{4\log n} - \frac{n^2\log\log n}{4\log^2 n} + \frac{\rho(n)n^2}{\log^2 n}.$$

*Proof.* By [3, Theorem 3], we have

$$np_n - \sum_{k\leq n} p_k \geq \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \Theta(n) \quad (4.9)$$

for every $n \geq 52703656$, where

$$\Theta(n) = \frac{43.6p_n^2}{8\log^4 p_n} + \frac{90.9p_n^2}{4\log^5 p_n} + \frac{927.5p_n^2}{8\log^6 p_n} + \frac{5620.5p_n^2}{8\log^7 p_n} + \frac{39537.75p_n^2}{8\log^8 p_n}.$$

Let $n \geq 52703656$. By (4.9) and the definition of $B_n$ it suffices to prove that

$$\frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \Theta(n) > \frac{np_n}{2} + \frac{n^2}{4} + \frac{n^2}{4\log n} - \frac{n^2\log\log n}{4\log^2 n} + \frac{\rho(n)n^2}{\log^2 n}. \quad (4.10)$$

For convenience, we write $p = p_n, y = \log n$ und $z = \log p$. By Lemma 4.6 and the definition of $\rho(n)$, we have

$$\begin{aligned}&2n^2z^7y^2 + 4.2n^2z^6y^2 + 32.6n^2z^5y^2 + 927.5p^2z^2y^2 + 5620.5p^2zy^2 + 39537.75p^2y^2\\&\quad > 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 116.964n^2z^4y^2 + 755.592npz^3y^2\end{aligned}$$

which is equvalent to

$$\begin{aligned}&2n^2z^7y^2 + 4.2n^2z^6y^2 + 53.5n^2z^4y^2(z-1.1) + 927.5p^2z^2y^2 + 5620.5p^2zy^2 + 39537.75p^2y^2\\&\quad > 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 20.9n^2z^5y^2 + 58.114n^2z^4y^2 + 755.592npz^3y^2.\end{aligned}$$

Using (4.5), we get

$$\begin{aligned}&2n^2z^7y^2 + 4.2n^2z^6y^2 + 53.5npz^4y^2 + 927.5p^2z^2y^2 + 5620.5p^2zy^2 + 39537.75p^2y^2\\&\quad > 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 20.9n^2z^5y^2 + 58.114n^2z^4y^2 + 755.592npz^3y^2.\end{aligned}$$

This inequality is equvalent to

$$\begin{aligned}2n^2z^7y^2 &+ 4.2n^2z^6y^2 + 181.8npz^3y^2(z-1.1) + 927.5p^2z^2y^2 + 5620.5p^2zy^2 + 39537.75p^2y^2\\ &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 20.9n^2z^5y^2 + 58.114n^2z^4y^2 + 128.3npz^4y^2\\ &\quad + 555.612npz^3y^2.\end{aligned}$$

Again, we use (4.5) to obtain that the inequality

$$\begin{aligned}2n^2z^7y^2 &+ 4.2n^2z^6y^2 + 181.8p^2z^3y^2 + 927.5p^2z^2y^2 + 5620.5p^2zy^2 + 39537.75p^2y^2\\ &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 20.9n^2z^5y^2 + 58.114n^2z^4y^2 + 128.3npz^4y^2\\ &\quad + 555.612npz^3y^2 \qquad (4.11)\end{aligned}$$

holds. We set

$$\Psi_1(x) = x - 1 - \frac{1.17}{x}$$

to see that (4.11) is equivalent to the inequality

$$\begin{aligned}2n^2z^7y^2 &+ 10.2n^2z^5y^2\Psi_1(z) + 181.8p^2z^3y^2 + 927.5p^2z^2y^2 + 5620.5p^2zy^2 + 39537.75p^2y^2\\ &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 6n^2z^6y^2 + 10.7n^2z^5y^2 + 46.18n^2z^4y^2\\ &\quad + 128.3npz^4y^2 + 555.612npz^3y^2.\end{aligned}$$

Now we use (4.4) to get

$$\begin{aligned}2n^2z^7y^2 &+ 10.2npz^5y^2 + 181.8p^2z^3y^2 + 927.5p^2z^2y^2 + 5620.5p^2zy^2 + 39537.75p^2y^2\\ &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 6n^2z^6y^2 + 10.7n^2z^5y^2 + 46.18n^2z^4y^2\\ &\quad + 128.3npz^4y^2 + 555.612npz^3y^2,\end{aligned}$$

which is equivalent to

$$\begin{aligned}2n^2z^7y^2 &+ 43.6npz^4y^2\Psi_1(z) + 181.8p^2z^3y^2 + 927.5p^2z^2y^2 + 5620.5p^2zy^2 + 39537.75p^2y^2\\ &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 6n^2z^6y^2 + 10.7n^2z^5y^2 + 46.18n^2z^4y^2 + 33.4npz^5y^2\\ &\quad + 84.7npz^4y^2 + 504.6npz^3y^2.\end{aligned}$$

By (4.4) and the definition of $\Theta(n)$, we obtain

$$\begin{aligned}2n^2z^7y^2 + 8\Theta(n)z^8y^2 &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 6n^2z^6y^2 + 10.7n^2z^5y^2 + 46.18n^2z^4y^2\\ &\quad + 33.4npz^5y^2 + 84.7npz^4y^2 + 504.6npz^3y^2. \qquad (4.12)\end{aligned}$$

Now we set

$$\Psi_2(x) = x - 1 - \frac{1}{x} - \frac{3.83}{x^2}$$

to obtain that (4.12) is equivalent to the inequality

$$\begin{aligned}4n^2z^6y^2\Psi_2(z) + 8\Theta(n)z^8y^2 &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 2n^2z^7y^2 + 2n^2z^6y^2 + 6.7n^2z^5y^2\\ &\quad + 30.86n^2z^4y^2 + 33.4npz^5y^2 + 84.7npz^4y^2 + 504.6npz^3y^2\end{aligned}$$

and by (4.3) we get

$$\begin{aligned}4npz^6y^2 + 8\Theta(n)z^8y^2 &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 2n^2z^7y^2 + 2n^2z^6y^2 + 6.7n^2z^5y^2\\ &\quad + 30.86n^2z^4y^2 + 33.4npz^5y^2 + 84.7npz^4y^2 + 504.6npz^3y^2\end{aligned}$$

which is equivalent to

$$\begin{aligned}14npz^5y^2\Psi_2(z) + 8\Theta(n)z^8y^2 &> 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 2n^2z^7y^2 + 2n^2z^6y^2 + 6.7n^2z^5y^2\\ &\quad + 30.86n^2z^4y^2 + 10npz^6y^2 + 19.4npz^5y^2 + 70.7npz^4y^2 + 450.98npz^3y^2.\end{aligned}$$

Again by (4.3), we get that the inequality

$$\begin{aligned}14p^2z^5y^2 + 8\Theta(n)z^8y^2 > 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 2n^2z^7y^2 + 2n^2z^6y^2 + 6.7n^2z^5y^2 \\ + 30.86n^2z^4y^2 + 10npz^6y^2 + 19.4npz^5y^2 + 70.7npz^4y^2 + 450.98npz^3y^2.\end{aligned}$$

holds. By putting

$$\Psi_3(x) = x - 1 - \frac{1}{x} - \frac{3.35}{x^2} - \frac{15.43}{x^3},$$

the last inequality is equivalent to

$$\begin{aligned}2n^2z^7y^2\Psi_3(z) + 14p^2z^5y^2 + 8\Theta(n)z^8y^2 > 2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 10npz^6y^2 \\ + 19.4npz^5y^2 + 70.7npz^4y^2 + 450.98npz^3y^2\end{aligned}$$

and by (4.2) we get that

$$\begin{aligned}2npz^7y^2 + 14p^2z^5y^2 + 8\Theta(n)z^8y^2 > 2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 10npz^6y^2 \\ + 19.4npz^5y^2 + 70.7npz^4y^2 + 450.98npz^3y^2.\end{aligned}$$

This inequality is equivalent to

$$\begin{aligned}6npz^6y^2\Psi_3(z) + 14p^2z^5y^2 + 8\Theta(n)z^8y^2 > 2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 4npz^7y^2 \\ + 4npz^6y^2 + 13.4npz^5y^2 + 50.6npz^4y^2 + 358.4npz^3y^2\end{aligned}$$

and, again by (4.2), it follows that

$$\begin{aligned}6p^2z^6y^2 + 14p^2z^5y^2 + 8\Theta(n)z^8y^2 > 2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8 + 4npz^7y^2 \\ + 4npz^6y^2 + 13.4npz^5y^2 + 50.6npz^4y^2 + 358.4npz^3y^2. \qquad (4.13)\end{aligned}$$

Finally we set

$$\Psi_4(x) = x - 1 - \frac{1}{x} - \frac{3.35}{x^2} - \frac{12.65}{x^3} - \frac{89.6}{x^4},$$

then, by (4.13), we have

$$\begin{aligned}4npz^7y^2\Psi_4(z) + 6p^2z^6y^2 + 14p^2z^5y^2 + 8\Theta(n)z^8y^2 \\ > 4npz^8y^2 + 2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8\end{aligned}$$

Now we apply the inequality (4.1) to get

$$\begin{aligned}4p^2z^7y^2 + 6p^2z^6y^2 + 14p^2z^5y^2 + 8\Theta(n)z^8y^2 \\ > 4npz^8y^2 + 2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\rho(n)n^2z^8.\end{aligned}$$

By dividing the last inequality by $8\log^8 p \log^2 n$, we obtain the inequality (4.10) and the claim follows for every $n \geq 52703656$. For every $842857 \leq n \leq 52703655$, we check the required inequality with a computer. □

By proving the next corollary, we improve the inequality (1.2).

**Corollary 4.8.** *If $n \geq 348247$, then*

$$B_n > \frac{n^2}{4} + \frac{n^2}{4\log n} - \frac{n^2(\log\log n - 2.1)}{4\log^2 n}.$$

*Proof.* For convenience, we write again $y = \log n$ and $z = \log p_n$. First, we consider the case $n \geq 842857$. By Theorem 4.7 it suffices to show that $\rho(n) \geq 2.1/4$. Since $x^2 - 5.2x + 16.3 \cdot 0.75^2 > 0$ for every $x \in \mathbb{R}$, we get

$$(\log^2 y - (1 + 4.2)\log y + 16.3 \cdot 0.75^2)z^2 + 2.1z\log^2 y - 2.1z(\log y - 1) \geq 0.$$

Using Lemma 4.5, the last inequality implies that

$$16.3y^2 + (\log^2 y - (1 + 4.2)\log y)z^2 + 2.1z\log^2 y - 2.1z(\log y - 1) \geq 0. \tag{4.14}$$

From Lemma 4.4, it follows that

$$z \leq y + \log y + \log\left(1 + \frac{\log y - 1}{y} + \frac{\log y - 2}{y^2}\right).$$

Now we apply the inequality $\log(1 + t) \leq t$, which holds for every $t > -1$, to get

$$z \leq y + \log y + \frac{\log y - 1}{y} + \frac{\log y - 2}{y^2}. \tag{4.15}$$

Using the result of Rosser [17, Theorem 1] that $p_n > n\log n$ for every $n \in \mathbb{N}$, we obtain

$$-z + \log y \leq -y, \tag{4.16}$$

and therefore

$$-2.1z^2\log y + 2.1z\log^2 y \leq -2.1zy\log y.$$

Hence, by using (4.14),

$$16.3y^2 + z^2(\log y - 1 - 2.1)\log y - 2.1zy\log y - 2.1z(\log y - 1) \geq 0. \tag{4.17}$$

Let $f(x) = \log x - 15(\log\log x - 1)/2$. Then, $f'(x) \geq 0$ if and only if $x \geq e^{15/2}$. Further, we have $f(e^{15/2}) \geq 4.26$. So, $f(x) \geq 0$ for every $x > 1$. Therefore,

$$\frac{3.1(\log\log n - 2)}{\log n} + \frac{2.1(\log\log n - 1)}{\log n} + \frac{2.1(\log\log n - 2)}{\log^2 n} \leq (3.1 + 2.1 + 2.1)\cdot\frac{2}{15} < 1.$$

We multiply this inequality by $z^2$ and combine this to (4.17) to obtain

$$\begin{aligned} &z^2 + 16.3y^2 + z^2(\log y - 1 - 2.1)\log y \\ &\quad\geq \frac{(3.1z^2)(\log y - 2)}{y} + 2.1z(y\log y + \log y - 1) + \frac{2.1z^2}{y}\left(\log y - 1 + \frac{\log y - 2}{y}\right). \end{aligned}$$

Since $z^2 > z$ and $\log y - 2 > 0$ we get that

$$\begin{aligned} &z^2 + 16.3 + z^2(\log y - 1 - 2.1)\log y \\ &\quad\geq \frac{(z^2 + 2.1z)(\log y - 2)}{y} + 2.1z(y\log y + \log y - 1) + \frac{2.1z^2}{y}\left(\log y - 1 + \frac{\log y - 2}{y}\right), \end{aligned}$$

which is equivalent to

$$\begin{aligned} &z^2 + 16.3y^2 + 2.1zy^2 + z^2(\log y - 1 - 2.1)\log y \\ &\quad\geq 2.1zy\left(y + \log y + \frac{\log y - 1}{y} + \frac{\log y - 2}{y^2}\right) + \frac{z^2(\log y - 2)}{y} + \frac{2.1z^2}{y}\left(\log y - 1 + \frac{\log y - 2}{y}\right). \end{aligned}$$

Now we use (4.15) to obtain

$$\begin{aligned} &z^2 + 16.3y^2 + 2.1zy^2 + z^2(\log y - 1 - 2.1)\log y \\ &\quad\geq 2.1z^2y + \frac{z^2(\log y - 2)}{y} + \frac{2.1z^2}{y}\left(\log y - 1 + \frac{\log y - 2}{y}\right) \end{aligned}$$

and this inequality is equvalent to

$$\begin{aligned} &z^2 + 16.3y^2 + 2.1zy^2 + z^2(\log y - 1)\log y \\ &\quad\geq 2.1z^2\left(y + \log y + \frac{\log y - 1}{y} + \frac{\log y - 2}{y^2}\right) + \frac{z^2(\log y - 2)}{y}. \end{aligned}$$

Again, by using (4.15),

$$z^2 + 16.3y^2 + 2.1zy^2 + z^2(\log y - 1)\log y \geq 2.1z^3 + \frac{z^2(\log y - 2)}{y}.$$

Now, by (4.16),

$$16.3y^2 + 2.1zy^2 + z^2(z - y)\log y \geq z^2(\log y - 1) + 2.1z^3 + \frac{z^2(\log y - 2)}{y},$$

which is equivalent to the inequality

$$z^2y^2 + 16.3y^2 + 2.1zy^2 \geq z^2y\left(y + \log y + \frac{\log y - 1}{y} + \frac{\log y - 2}{y^2}\right) - z^3\log y + 2.1z^3.$$

Now we use the inequality (4.15) to get

$$z^2y^2 + 16.3y^2 + 2.1zy^2 \geq z^3y - z^3\log y + 2.1z^3.$$

Finally we divide this inequality by $4z^3$ to obtain $\rho(n) \geq 2.1/4$. Hence, the claim follows from Theorem 4.7 for every $n \geq 842857$. We check the remaining cases for $n$ with a computer. □

*Remark.* Based on some computations, Sinha [20, p. 4] conjectured that

$$B_n > \frac{n^2}{4} + \frac{n^2}{2\log n} - \frac{n^2\log\log n}{4\log^2 n} \tag{4.18}$$

for every $n \geq 835$. However, Corollary 3.2 implies that this inequality does not hold for sufficiently large $n$. In fact, the smallest counterexample $n \geq 835$ for (4.18) is given by $n = 835$.

### 4.3 New upper bounds for $B_n$

We set

$$\kappa(n) = \frac{\log p_n \log^2 n + 4.9\log^2 n - \log^2 p_n \log n + \log^2 p_n \log\log n}{4\log^2 p_n} + \frac{r(\log p_n)\log^2 n}{8\log^6 p_n},$$

where $r(x)$ is defined by

$$r(x) = 35.4x^3 + 213.9x^2 + 1478.78x + 30199.015, \tag{4.19}$$

to obtain the following theorem.

**Theorem 4.9.** *For every $n \geq 2$, we have*

$$B_n < \frac{n^2}{4} + \frac{n^2}{4\log n} - \frac{n^2\log\log n}{4\log^2 n} + \frac{\kappa(n)n^2}{\log^2 n}.$$

*Proof.* By [3, Theorem 4], we have

$$np_n - \sum_{k\leq n} p_k \leq \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \Omega(n) \tag{4.20}$$

for every $n \in \mathbb{N}$, where

$$\Omega(n) = \frac{46.4p_n^2}{8\log^4 p_n} + \frac{95.1p_n^2}{4\log^5 p_n} + \frac{962.5p_n^2}{8\log^6 p_n} + \frac{5809.5p_n^2}{8\log^7 p_n} + \frac{59424p_n^2}{8\log^8 p_n}. \tag{4.21}$$

First, let $n \geq 66775031$; i.e. $p_n \geq 1332479531$. By (4.20) and the definition of $B_n$ it suffices to show that

$$\frac{np_n}{2} + \frac{n^2}{4} + \frac{n^2}{4\log n} - \frac{n^2\log\log n}{4\log^2 n} + \frac{\kappa(n)n^2}{\log^2 n} > \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \Omega(n). \tag{4.22}$$

For convenience, we write $p = p_n, y = \log n$ and $z = \log p$. Using the definiton of $\kappa(n)$ and $r(x)$, we get

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8\\&\qquad = 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 1478.78n^2z^3y^2 + 30199.015n^2z^2y^2.\end{aligned}$$

By setting $\Phi_1(x) = x$, we get that the last equality is equivalent to

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 3993.86n^2z^2y^2\\&\qquad = 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 1478.78n^2z^3y^2 + 34192.875n^2zy^2\Phi_1(z).\end{aligned}$$

Since $p < n\Phi_1(z)$ by Lemma 4.3, we obtain the inequality

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 3993.86n^2z^2y^2\\&\qquad > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 1478.78n^2z^3y^2 + 34192.875npzy^2,\end{aligned}$$

which is equivalent to

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 3993.86n^2z^2y^2 + 25231.125npzy^2\\&\qquad > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 1478.78n^2z^3y^2 + 59424npy^2\Phi_1(z).\end{aligned}$$

Again, we use the inequality $p < n\Phi_1(z)$ to get

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 3993.86n^2z^2y^2 + 25231.125npzy^2\\&\qquad > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 1478.78n^2z^3y^2 + 59424p^2y^2.\end{aligned}$$

By setting $\Phi_2(x) = x - 1$, the last inequality can be rewritten as

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 623.87n^2z^3y^2 + 1891.21n^2z^2y^2 + 25231.125npzy^2\\&\qquad > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 2102.65n^2z^2y^2\Phi_2(z) + 59424p^2y^2.\end{aligned}$$

Since $p < n\Phi_2(z)$ by Dusart [7, Théorème 1.10], we get

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 623.87n^2z^3y^2 + 1891.21n^2z^2y^2 + 25231.125npzy^2\\&\qquad > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 2102.65npz^2y^2 + 59425p^2y^2\end{aligned}$$

which is equivalent to

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 623.87n^2z^3y^2 + 1891.21n^2z^2y^2 + 3706.85npz^2y^2 + 19421.625npzy^2\\&\qquad > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 5809.5npzy^2\Phi_2(z) + 59424p^2y^2.\end{aligned}$$

Now we use again the inequality $p < n\Phi_2(z)$ and obtain

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 623.87n^2z^3y^2 + 1891.21n^2z^2y^2 + 3706.85npz^2y^2 + 19421.625npzy^2\\&\qquad > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 213.9n^2z^4y^2 + 5809.5p^2zy^2 + 59424p^2y^2.\end{aligned}$$

We define

$$\Phi_3(x) = x - 1 - \frac{1}{x}$$

to rewrite the last inequality to

$$\begin{aligned}&2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 113.6n^2z^4y^2 + 296.37n^2z^3y^2 + 1563.71n^2z^2y^2 + 3706.85npz^2y^2\\&\qquad + 19421.625npzy^2\\&\qquad > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 327.5n^2z^3y^2\Phi_3(z) + 5809.5p^2zy^2 + 59424p^2y^2.\end{aligned}$$

By Lemma 4.2, we have $p < n\Phi_3(z)$. Hence,

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 113.6n^2z^4y^2 + 296.37n^2z^3y^2 + 1563.71n^2z^2y^2 + 3706.85npz^2y^2\\ + 19421.625npzy^2\\ > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 327.5npz^3y^2 + 5809.5p^2zy^2 + 59424p^2y^2.\end{aligned}$$

This inequality is equivalent to

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 113.6n^2z^4y^2 + 296.37n^2z^3y^2 + 1563.71n^2z^2y^2 + 635npz^3y^2\\ + 2744.35npz^2y^2 + 18459.125npzy^2\\ > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 962.5npz^2y^2\Phi_3(z) + 5809.5p^2zy^2 + 59424p^2y^2.\end{aligned}$$

Again, by using the inequality $p < n\Phi_3(z)$, we obtain

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 113.6n^2z^4y^2 + 296.37n^2z^3y^2 + 1563.71n^2z^2y^2 + 635npz^3y^2\\ + 2744.35npz^2y^2 + 18459.125npzy^2\\ > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 35.4n^2z^5y^2 + 962.5p^2z^2y^2 + 5809.5p^2zy^2 + 59424p^2y^2.\end{aligned}$$

By setting

$$\Phi_4(x) = x - 1 - \frac{1}{x} - \frac{2.65}{x^2},$$

the last inequality can be rewritten as

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 25.1n^2z^5y^2 + 53.1n^2z^4y^2 + 235.87n^2z^3y^2 + 1403.385n^2z^2y^2\\ + 635npz^3y^2 + 2744.35npz^2y^2 + 18459.125npzy^2\\ > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 60.5n^2z^4y^2\Phi_4(z) + 962.5p^2z^2y^2 + 5809.5p^2zy^2 + 59424p^2y^2.\end{aligned}$$

By Lemma 4.2, we have $p < n\Phi_4(z)$. Therefore,

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 25.1n^2z^5y^2 + 53.1n^2z^4y^2 + 235.87n^2z^3y^2 + 1403.385n^2z^2y^2\\ + 635npz^3y^2 + 2744.35npz^2y^2 + 18459.125npzy^2\\ > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 60.5npz^4y^2 + 962.5p^2z^2y^2 + 5809.5p^2zy^2 + 59424p^2y^2,\end{aligned}$$

which is equivalent to

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 25.1n^2z^5y^2 + 53.1n^2z^4y^2 + 235.87n^2z^3y^2 + 1403.385n^2z^2y^2\\ + 129.7npz^4y^2 + 444.8npz^3y^2 + 2554.17npz^2y^2 + 17955.095npzy^2\\ > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 190.2npz^3y^2\Phi_4(z) + 962.5p^2z^2y^2 + 5809.5p^2zy^2 + 59424p^2y^2\end{aligned}$$

Now we apply again the inequality $p < n\Phi_4(z)$ to get that the inequality

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 25.1n^2z^5y^2 + 53.1n^2z^4y^2 + 235.87n^2z^3y^2 + 1403.385n^2z^2y^2\\ + 129.7npz^4y^2 + 444.8npz^3y^2 + 2554.17npz^2y^2 + 17955.095npzy^2\\ > 2n^2z^7y^2 + 9.8n^2z^6y^2 + 190.2p^2z^3y^2 + 962.5p^2z^2y^2 + 5809.5p^2zy^2 + 59424p^2y^2\end{aligned}$$

holds. Let

$$\Phi_5(x) = x - 1 - \frac{1}{x} - \frac{2.65}{x^2} - \frac{13.35}{x^3}$$

to rewrite the last inequality to

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 6n^2z^6y^2 + 9.3n^2z^5y^2 + 37.3n^2z^4y^2 + 194n^2z^3y^2 + 1192.455n^2z^2y^2\\ + 129.7npz^4y^2 + 444.8npz^3y^2 + 2554.17npz^2y^2 + 17955.095npzy^2\\ > 2n^2z^7y^2 + 15.8n^2z^5y^2\Phi_5(z) + 190.2p^2z^3y^2 + 962.5p^2z^2y^2 + 5809.5p^2zy^2 + 59424p^2y^2.\end{aligned}$$

It follows from Lemma 4.2 that the inequality $p < n\Phi_5(z)$ is fullfilled. So,

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 6n^2z^6y^2 + 9.3n^2z^5y^2 + 37.3n^2z^4y^2 + 194n^2z^3y^2 + 1192.455n^2z^2y^2 \\ + 129.7npz^4y^2 + 444.8npz^3y^2 + 2554.17npz^2y^2 + 17955.095npzy^2 \\ > 2n^2z^7y^2 + 15.8npz^5y^2 + 190.2p^2z^3y^2 + 962.5p^2z^2y^2 + 5809.5p^2zy^2 + 59424p^2y^2,\end{aligned}$$

which is equavalent to

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 6n^2z^6y^2 + 9.3n^2z^5y^2 + 37.3n^2z^4y^2 + 194n^2z^3y^2 + 1192.455n^2z^2y^2 \\ + 30.6npz^5y^2 + 83.3npz^4y^2 + 398.4npz^3y^2 + 2431.21npz^2y^2 + 17335.655npzy^2 \\ > 2n^2z^7y^2 + 46.4npz^4y^2\Phi_5(z) + 190.2p^2z^3y^2 + 962.5p^2z^2y^2 + 5809.5p^2zy^2 + 59424p^2y^2.\end{aligned}$$

Now we use the inequality $p < n\Phi_5(z)$ for a second time and the definition of $\Omega(n)$ from (4.21) to get

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 6n^2z^6y^2 + 9.3n^2z^5y^2 + 37.3n^2z^4y^2 + 194n^2z^3y^2 + 1192.455n^2z^2y^2 \\ + 30.6npz^5y^2 + 83.3npz^4y^2 + 398.4npz^3y^2 + 2431.21npz^2y^2 + 17335.655npzy^2 \\ > 2n^2z^7y^2 + 8\Omega(n)z^8y^2.\end{aligned}$$

Next, we set

$$\Phi_6(x) = x - 1 - \frac{1}{x} - \frac{2.65}{x^2} - \frac{13.35}{x^3} - \frac{70.3}{x^4}$$

to obtain that the last inequality is equivalent to

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 2n^2z^7y^2 + 2n^2z^6y^2 + 5.3n^2z^5y^2 + 26.7n^2z^4y^2 + 140.6n^2z^3y^2 \\ + 911.255n^2z^2y^2 + 30.6npz^5y^2 + 83.3npz^4y^2 + 398.4npz^3y^2 + 2431.21npz^2y^2 \\ + 17335.655npzy^2 \\ > 4n^2z^6y^2\Phi_6(z) + 8\Omega(n)z^8y^2.\end{aligned}$$

Since $p < n\Phi_6(z)$ by Lemma 4.2, we get

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 2n^2z^7y^2 + 2n^2z^6y^2 + 5.3n^2z^5y^2 + 26.7n^2z^4y^2 + 140.6n^2z^3y^2 \\ + 911.255n^2z^2y^2 + 30.6npz^5y^2 + 83.3npz^4y^2 + 398.4npz^3y^2 + 2431.21npz^2y^2 \\ + 17335.655npzy^2 \\ > 4npz^6y^2 + 8\Omega(n)z^8y^2,\end{aligned}$$

which can be rewritten to

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 2n^2z^7y^2 + 2n^2z^6y^2 + 5.3n^2z^5y^2 + 26.7n^2z^4y^2 + 140.6n^2z^3y^2 \\ + 911.255n^2z^2y^2 + 10npz^6y^2 + 16.6npz^5y^2 + 69.3npz^4y^2 + 361.3npz^3y^2 \\ + 2244.31npz^2y^2 + 16351.455npzy^2 \\ > 14npz^5y^2\Phi_6(z) + 8\Omega(n)z^8y^2.\end{aligned}$$

By using again the inequality $p < n\Phi_6(z)$, we have

$$\begin{aligned}2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 2n^2z^7y^2 + 2n^2z^6y^2 + 5.3n^2z^5y^2 + 26.7n^2z^4y^2 + 140.6n^2z^3y^2 \\ + 911.255n^2z^2y^2 + 10npz^6y^2 + 16.6npz^5y^2 + 69.3npz^4y^2 + 361.3npz^3y^2 \\ + 2244.31npz^2y^2 + 16351.455npzy^2 \\ > 14p^2z^5y^2 + 8\Omega(n)z^8y^2.\end{aligned}$$

Now, let

$$\Phi_7(x) = x - 1 - \frac{1}{x} - \frac{2.65}{x^2} - \frac{13.35}{x^3} - \frac{70.3}{x^4} - \frac{455.6275}{x^5}.$$

Then the last inequality is equivalent to

$$\begin{aligned}2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 10npz^6y^2 + 16.6npz^5y^2 + 69.3npz^4y^2 \\ + 361.3npz^3y^2 + 2244.31npz^2y^2 + 16351.455npzy^2 \\ > 2n^2z^7y^2\Phi_7(z) + 14p^2z^5y^2 + 8\Omega(n)z^8y^2.\end{aligned}$$

Lemma 4.2 implies that $p < n\Phi_7(z)$. Hence,

$$\begin{aligned}2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 10npz^6y^2 + 16.6npz^5y^2 + 69.3npz^4y^2 \\ + 361.3npz^3y^2 + 2244.31npz^2y^2 + 16351.455npzy^2 \\ > 2npz^6y^2 + 14p^2z^5y^2 + 8\Omega(n)z^8y^2,\end{aligned}$$

which we rewrite to

$$\begin{aligned}2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 4npz^7y^2 + 4npz^6y^2 + 10.6npz^5y^2 + 53.4npz^4y^2 \\ + 281.2npz^3y^2 + 1822.51npz^2y^2 + 13617.69npzy^2 \\ > 6npz^6y^2\Phi_7(z) + 14p^2z^5y^2 + 8\Omega(n)z^8y^2\end{aligned}$$

Again, by using the inequality $p < n\Phi_7(z)$, we obtain that the inequality

$$\begin{aligned}2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 + 4npz^7y^2 + 4npz^6y^2 + 10.6npz^5y^2 + 53.4npz^4y^2 \\ + 281.2npz^3y^2 + 1822.51npz^2y^2 + 13617.69npzy^2 \\ > 6p^2z^6y^2 + 14p^2z^5y^2 + 8\Omega(n)z^8y^2\end{aligned}$$

holds. Finally, we set

$$\Phi_8(x) = x - 1 - \frac{1}{x} - \frac{2.65}{x^2} - \frac{13.35}{x^3} - \frac{70.3}{x^4} - \frac{455.6275}{x^5} - \frac{3404.4225}{x^6}.$$

Then the last inequality can be rewritten to

$$\begin{aligned}4npz^8y^2 + 2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 \\ > 4npz^7y^2\Phi_8(z) + 6p^2z^6y^2 + 14p^2z^5y^2 + 8\Omega(n)z^8y^2.\end{aligned}$$

Now, we get $p < n\Phi_8(z)$ by Lemma 4.2. Therefore,

$$\begin{aligned}4npz^8y^2 + 2n^2z^8y^2 + 2n^2z^8y - 2n^2z^8\log y + 8\kappa(n)n^2z^8 \\ > 4p^2z^7y^2 + 6p^2z^6y^2 + 14p^2z^5y^2 + 8\Omega(n)z^8y^2.\end{aligned}$$

We divide both sides of this inequality by $8z^8y^2$ to obtain the inequality (4.22) for every $n \geq 66775031$. For every $2 \leq n \leq 66775030$ we check the required inequality with a computer. □

According to Corollary 3.2, Theorem 4.9 implies the following result.

**Corollary 4.10.** *If $n \geq 26220$, then*

$$B_n < \frac{n^2}{4} + \frac{n^2}{4\log n} - \frac{n^2(\log\log n - 5.22)}{4\log^2 n}.$$

*Proof.* First, we consider the case $n \geq e^{19.63}$. For convenience, we set again $y = \log n$ and $z = \log p_n$. Further, we set $a_1 = 0.08$ and

$$h_1(x) = 16a_1 - \frac{4\log x}{x} - \frac{4}{x} - \frac{2\log x}{x^2}.$$

Then, $h_1(x) \geq 0.004$ for every $x \geq 11$. Next, we set

$$h(x) = 8a_1x^2 - 4x\log x - \log^2 x.$$

Since $h'(x) = xh_1(x) \geq 0$ für alle $x \geq 11$ and $h(19.17) \geq 0.008$, we get that $h(x) \geq 0$ for every $x \geq 19.17$. Now, let $x \geq 19.17$. We consider the function

$$f(x) = 4a_1(x + \log x) + (x + 4a_1 - \log x) \log \left(1 + \frac{\log x - 1}{x}\right) - \log^2 x.$$

We have

$$f'(x) = 4a_1 + \frac{4a_1}{x} + \left(1 - \frac{1}{x}\right) \log \left(1 + \frac{\log x - 1}{x}\right) + \frac{2 - \log x}{x + \log x - 1} + \frac{8a_1 - (4a_1 + 2) \log x}{x(x + \log x - 1)} + \frac{\log^2 x}{x(x + \log x - 1)} - \frac{2 \log x}{x}.$$

Since $x \geq e^{4a_1+2}$, we get

$$f'(x) \geq 4a_1 + \frac{4a_1}{x} + \left(1 - \frac{1}{x}\right) \log \left(1 + \frac{\log x - 1}{x}\right) + \frac{2 - \log x}{x + \log x - 1} + \frac{8a_1}{x(x + \log x - 1)} - \frac{2 \log x}{x}.$$

Now we use the inequality $\log(1 + t) \geq t - t^2/2$, which holds for every $t \geq 0$, to obtain that

$$f'(x) \geq 4a_1 + \frac{4a_1}{x} + \frac{\log x - 1}{x} - \frac{\log^2 x - 2 \log x + 1}{2x^2} - \frac{\log x - 1}{x^2} + \frac{\log^2 x - 2 \log x + 1}{2x^3} + \frac{2 - \log x}{x + \log x - 1} + \frac{8a_1}{x(x + \log x - 1)} - \frac{2 \log x}{x}.$$

It follows that

$$f'(x) \geq 4a_1 - \frac{2 \log x}{x} - \frac{\log^2 x}{2x^2} = \frac{h(x)}{2x^2} \geq 0$$

for every $x \geq \max\{e^{4a_1+2}, 19.17\} = 19.17$. In addition, we have $f(19.63) \geq 0.00011$ and it follows that $f(x) \geq 0$ for every $x \geq 19.63$. Since $n \geq e^{19.63}$, we have $y \geq 19.63$. Therefore $f(y) \geq 0$; i.e.

$$\left(y + \log y + \left(1 + \frac{\log y - 1}{y}\right)\right) (4a_1 + y - \log y) \geq y^2. \tag{4.23}$$

In [8], Dusart found that the inequality $p_k \geq k(\log k + \log \log k - 1)$ holds for every $k \geq 2$. So,

$$z \geq y + \log y + \log \left(1 + \frac{\log y - 1}{y}\right). \tag{4.24}$$

Now we use (4.23) to obtain that

$$z(4a_1 + y - \log y) \geq y^2,$$

which is equivalent to the inequality

$$8a_1 z^8 \geq 2z^7 y^2 - 2z^8 y + 2z^8 \log y. \tag{4.25}$$

Next, we set $a_2 = 1.225$ and

$$t(x) = 16a_2 x^5 \log x + 8a_2 x^4 \log^2 x - r(x),$$

where $r(x)$ is defined by (4.19). Let $x \geq 4.4$. Then,

$$\frac{t(x)}{x^5 \log x} \geq 16a_2 - \frac{35.4}{x^2 \log x} - \frac{213.9}{x^3 \log x} - \frac{1478.78}{x^4 \log x} - \frac{30199.015}{x^5 \log x} \geq 1.648.$$

Hence, $t(x) \geq 0$ for every $x \geq 4.4$ and it follows that

$$16a_2 z^5 y^2 \log z + 8a_2 z^4 y^2 \log^2 z - r(z) y^2 + (8a_2 - 9.8) z^6 y^2 \geq 0.$$

The function $w(t) = \log(t)/t$ is decreasing for every $t \geq e$. So, $\log(y)/y \geq \log(z)/z$ and we get

$$8a_2 z^6 (y + \log y)^2 - r(z)y^2 - 9.8z^6y^2 \geq 0.$$

By (4.24), we obtain $z \geq y + \log y$. Hence,

$$8a_2 z^8 \geq r(z)y^2 + 9.8z^6y^2.$$

Now we use (4.25) to obtain that

$$10.44z^8 = 8(a_1 + a_2)z^8 \geq 2z^7y^2 - 2z^8y + 2z^8 \log y + r(z)y^2 + 9.8z^6y^2 = 8\kappa(n)z^8.$$

So, $\kappa(n) \leq 5.22/4$ for every $n \geq e^{19.63}$. Using a computer, we check that the inequality $\kappa(n) \leq 5.22/4$ holds for every $132380 \leq n \leq e^{19.63}$ as well. Applying Theorem 4.9, our corollary is proved for every $n \geq 132380$. A direkt computer check shows that the required inequality also holds for every $26220 \leq n \leq 132379$. □

# 5 Explicit estimates for the sum of the first $n$ primes

In this section, we improve the sharpest known estimates for the sum of the first $n$ primes by using the explicit estimates for $B_n$ obtained in Corollary (4.8) and Corollary (4.10)

## 5.1 A new upper bound for the sum of the first $n$ primes

By Corollary 2.8, for each $\varepsilon > 0$ there is a $N_0(\varepsilon) > 1$, such that

$$\sum_{k \leq n} p_k < \frac{n^2}{2} \left( \log n + \log\log n - \frac{3}{2} + \frac{\log\log n - 5/2}{\log n} - \frac{(\log\log n)^2 - 7\log\log n + 29/2 - \varepsilon}{2\log^2 n} \right) \tag{5.1}$$

for every $n > N_0(\varepsilon)$. Let $m \in \mathbb{N}$. According to Corollary 2.3, let $\widetilde{R}_m \in \mathbb{Q}[x]$ with $\deg(\widetilde{R}_m) = m$ and let $N_1 = N_1(m, \widetilde{R}_m) \in \mathbb{N}$ be minimal, such that

$$p_n < n \left( \log n + \log\log n - 1 + \sum_{s=1}^{m-1} \frac{(-1)^{s+1} R_s(\log\log n)}{s \log^s n} + \frac{(-1)^{m+1} \widetilde{R}_m(\log\log n)}{m \log^m n} \right) \tag{5.2}$$

for every $n \geq N_1$. Further, in view of Corollary 3.2, let $\widetilde{V}_m \in \mathbb{Q}[x]$ with $\deg(\widetilde{V}_m) = m - 1$ and let $N_2 = N_2(m, \widetilde{V}_m) \in \mathbb{N}$ be minimal, such that

$$B_n > \frac{n^2}{4} + \frac{n^2}{4\log n} + \frac{n^2}{2} \sum_{s=2}^{m-1} \frac{(-1)^{s+1} V_s(\log\log n)}{s \log^s n} + \frac{n^2}{2} \frac{(-1)^{m+1} \widetilde{V}_m(\log\log n)}{m \log^m n} \tag{5.3}$$

for every $n \geq N_2$. Then, we obtain the following result.

**Proposition 5.1.** *For every $n \geq \max\{N_1, N_2\}$, we have*

$$\sum_{k \leq n} p_k < \frac{n^2}{2} \left( \log n + \log\log n - \frac{3}{2} + \frac{R_1(\log\log n) - 1/2}{\log n} + \sum_{s=2}^{m-1} \frac{(-1)^{s+1}(R_s(\log\log n) - V_s(\log\log n))}{s \log^s n} \right)$$
$$+ \frac{n^2}{2} \frac{(-1)^{m+1}(\widetilde{R}_m(\log\log n) - \widetilde{V}_m(\log\log n))}{m \log^m n}.$$

*Proof.* The claim follows directly from (5.2) and (5.3). □

According to (5.1) we obtain the following corollary, which leads to an improvement of (1.4).

**Corollary 5.2.** *For every* $n \geq 355147$*, we have*

$$\sum_{k \leq n} p_k < \frac{n^2}{2} \left( \log n + \log\log n - \frac{3}{2} + \frac{\log\log n - 5/2}{\log n} - \frac{(\log\log n)^2 - 7\log\log n + 12.373}{2\log^2 n} \right).$$

*Proof.* By Corollary 2.3, we have $R_1(x) = x - 2$. By setting $\widetilde{R}_2(x) = x^2 - 6x + 10.273$, we get $N_1 = 8009824$ by [2, Korollar 2.11]. Now, we set $\widetilde{V}_2(x) = x - 2.1$ to get $N_2 = 348247$ by Corollary 4.8. So, by Proposition 5.1, the claim follows for every $n \geq 8009824$. By using a computer, we check the asserted inequality for every $355147 \leq n \leq 8009823$. ☐

**Corollary 5.3.** *For every* $n \geq 115149$*, we have*

$$\sum_{k \leq n} p_k < \frac{n^2}{2} \left( \log n + \log\log n - \frac{3}{2} + \frac{\log\log n - 5/2}{\log n} \right).$$

*Proof.* Since $x^2 - 7x + 12.373 > 0$ for every $x \in \mathbb{R}$, the claim follows from Corollary 5.2 for every $n \geq 355147$. It remains to check the required inequality for every $115149 \leq n \leq 355146$. ☐

## 5.2 A new lower bound for the sum of the first $n$ primes

By Corollary 2.8, we get that for each $\varepsilon > 0$ there exists an $N_3(\varepsilon) > 1$, such that

$$\sum_{k \leq n} p_k > \frac{n^2}{2} \left( \log n + \log\log n - \frac{3}{2} + \frac{\log\log n - 5/2}{\log n} - \frac{(\log\log n)^2 - 7\log\log n + 29/2 + \varepsilon}{2\log^2 n} \right) \tag{5.4}$$

for every $n \geq N_3(\varepsilon)$. Let $m \in \mathbb{N}$. According to Corollary 2.3, let $\widehat{R}_m \in \mathbb{Q}[x]$ with $\deg(\widehat{R}_m) = m$ and let $N_4 = N_4(m, \widehat{R}_m) \in \mathbb{N}$ be minimal, such that

$$p_n > n \left( \log n + \log\log n - 1 + \sum_{s=1}^{m-1} \frac{(-1)^{s+1} R_s(\log\log n)}{s \log^s n} + \frac{(-1)^{m+1} \widehat{R}_m(\log\log n)}{m \log^m n} \right) \tag{5.5}$$

for every $n \geq N_4$. In addition, in view of Corollary 3.2, let $\widehat{V}_m \in \mathbb{Q}[x]$ with $\deg(\widehat{V}_m) = m - 1$ and let $N_5 = N_5(m, \widehat{V}_m) \in \mathbb{N}$ be minimal, such that

$$B_n < \frac{n^2}{4} + \frac{n^2}{4\log n} + \frac{n^2}{2} \sum_{s=2}^{m-1} \frac{(-1)^{s+1} V_s(\log\log n)}{s \log^s n} + \frac{n^2}{2} \frac{(-1)^{m+1} \widehat{V}_m(\log\log n)}{m \log^m n} \tag{5.6}$$

for every $n \geq N_5$. Then, we obtain the following lower bound for the sum of the first $n$ prime numbers.

**Proposition 5.4.** *If* $n \geq \max\{N_4, N_5\}$*, then*

$$\sum_{k \leq n} p_k > \frac{n^2}{2} \left( \log n + \log\log n - \frac{3}{2} + \frac{R_1(\log\log n) - 1/2}{\log n} + \sum_{s=2}^{m-1} \frac{(-1)^{s+1}(R_s(\log\log n) - V_s(\log\log n))}{s \log^s n} \right)$$
$$+ \frac{n^2}{2} \frac{(-1)^{m+1}(\widehat{R}_m(\log\log n) - \widehat{V}_m(\log\log n))}{m \log^m n}.$$

*Proof.* The claim follows directly from (5.5) and (5.6). ☐

According to (5.4) we get the following corollary, which improves the inequality (1.5).

**Corollary 5.5.** *For every* $n \geq 2$*, we have*

$$\sum_{k \leq n} p_k > \frac{n^2}{2} \left( \log n + \log\log n - \frac{3}{2} + \frac{\log\log n - 5/2}{\log n} - \frac{(\log\log n)^2 - 7\log\log n + 17.067}{2\log^2 n} \right).$$

*Proof.* We have $R_1(x) = x - 2$ by Corollary 2.3. Setting $\widehat{R}_2(x) = x^2 - 6x + 11.847$, we have $N_4 = 2$ by [2, Korollar 2.25]. Further, we set $\widehat{V}_2(x) = x - 5.22$ to obtain $N_5 = 26220$ by Corollary 4.10. Hence, by Proposition 5.4, the corollary is proved for every $n \geq 26220$. For every $2 \leq n \leq 26219$, we check the asserted inequality with a computer. ☐

# 6 On explicit estimates for the step function $S(x) = \sum_{p \leq x} p$

In this section we study the step function

$$S(x) = \sum_{p \leq x} p,$$

which is constant on every interval of the form $[p_n, p_{n+1})$ with $S(p_n) = \sum_{k \leq n} p_k$. A result of Szalay [24, Lemma 1] implies that there exists a constant $a > 0$, such that

$$S(x) = \mathrm{li}(x^2) + O(x^2 e^{-a\sqrt{\log x}}). \tag{6.1}$$

Under the assumption that the Riemann hypothesis is true, Deléglise and Nicolas [6, Lemma 2.5] improved (6.1) by showing

$$|S(x) - \mathrm{li}(x^2)| \leq \frac{5}{24\pi} x^{3/2} \log x$$

for every $x \geq 41$. The goal of this section is to find explicit estimates for $S(x)$ by a continuous function, which yield explicit bounds for the difference $S(x) - \mathrm{li}(x^2)$.

## 6.1 Auxiliaries lemmata

In order to find new explicit estimates for $S(x)$, we first note two useful lemmata concerning upper and lower bound for the prime counting function $\pi(x)$ given in [4]

**Lemma 6.1.** *If $x > 1$, then*

$$\pi(x) < \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6.35x}{\log^4 x} + \frac{24.35x}{\log^5 x} + \frac{121.75x}{\log^6 x} + \frac{730.5x}{\log^7 x} + \frac{6801.4x}{\log^8 x}.$$

*Proof.* See [4, Theorem 1.1] □

**Lemma 6.2.** *If $x \geq 1332450001$, then*

$$\pi(x) > \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{5.65x}{\log^4 x} + \frac{23.65x}{\log^5 x} + \frac{118.25x}{\log^6 x} + \frac{709.5x}{\log^7 x} + \frac{4966.5x}{\log^8 x}.$$

*Proof.* See [4, Theorem 1.2] □

In the next lemma, we give an lower bound for the elements of a sequence involving the sum of the first $n$ prime numbers.

**Lemma 6.3.** *If $n \geq 52703656$, then*

$$np_n - \sum_{k \leq n} p_k \geq \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \Theta(n),$$

*where*

$$\Theta(n) = \frac{43.6p_n^2}{8\log^4 p_n} + \frac{90.9p_n^2}{4\log^5 p_n} + \frac{927.5p_n^2}{8\log^6 p_n} + \frac{5620.5p_n^2}{8\log^7 p_n} + \frac{79075.5p_n^2}{16\log^8 p_n}.$$

*Proof.* See [3, Theorem 3]. □

In the next two lemmata, we give some integration rules.

**Lemma 6.4.** *Seien $x, a \in \mathbb{R}$ mit $x \geq a > 1$. Dann gilt*

$$\int_a^x \frac{t\,dt}{\log^2 t} = 2\,\mathrm{li}(x^2) - 2\,\mathrm{li}(a^2) - \frac{x^2}{\log x} + \frac{a^2}{\log a}.$$

*Proof.* Siehe Dusart [7, Lemme 1.6]. □

**Lemma 6.5.** *Let $m \in \mathbb{N}$ with $m \geq 2$. Let $a_2, \ldots, a_m \in \mathbb{R}$ and let $r, s \in \mathbb{R}$ with $s \geq r > 1$. Then*

$$\sum_{k=2}^{m} a_k \int_r^s \frac{x\,dx}{\log^k x} = t_{m-1,1} \int_r^s \frac{x\,dx}{\log^2 x} - \sum_{k=2}^{m-1} t_{m-1,k} \left( \frac{s^2}{\log^k s} - \frac{r^2}{\log^k r} \right),$$

*where*

$$t_{i,j} = (j-1)! \sum_{l=j}^{i} \frac{2^{l-j} a_{l+1}}{l!}.$$

*Proof.* See [3, Proposition 9]. □

Finally, we note the following lemma.

**Lemma 6.6** (Abel's identity)**.** *For any function $a : \mathbb{N} \to \mathbb{C}$ let $A(x) = \sum_{n \leq x} a(n)$, where $A(x) = 0$ if $x < 1$. Assume $f$ has a continuous derivative on the interval $[y, x]$, where $0 < y < x$. Then we have*

$$\sum_{y<n\leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)\ dt.$$

*Proof.* See Apostol [1, Theorem 4.2]. □

## 6.2 New upper bounds for $S(x)$

Using (6.1) and the asymptotic formula for $\mathrm{li}(x^2)$ given in Lemma 2.4, we obtain that

$$S(x) \sim \frac{x^2}{2\log x} \sum_{k=0}^{n} \frac{k!}{2^k \log^k x} \qquad (x \to \infty) \tag{6.2}$$

for every $n \in \mathbb{N}$. In particular, the asymptotic formula (6.2) implies that for each $\varepsilon > 0$ there exists an $x_0(\varepsilon) > 1$, so that

$$S(x) < \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x} + \frac{x^2}{4\log^3 x} + \frac{(3+\varepsilon)x^2}{8\log^4 x} \tag{6.3}$$

for every $x \geq x_0(\varepsilon)$. According to (6.3), we show the following theorem, which improves the sharpest known upper bound for $S(x)$ found by Massias and Robin [15, Théorème D(v)] in 1996, namely that

$$S(x) \leq \frac{x^2}{2\log x} + \frac{3x^2}{10\log^2 x}$$

for every $x \geq 24281$.

**Theorem 6.7.** *For every $x \geq 355992$, we have*

$$S(x) < \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x} + \frac{x^2}{4\log^3 x} + \frac{7.2x^2}{8\log^4 x} + \frac{6.5x^2}{4\log^5 x} + \frac{46.5x^2}{8\log^6 x} + \frac{223.5x^2}{8\log^7 x} + \frac{14873.45x^2}{8\log^8 x}.$$

*Proof.* First, we consider the case $x \geq 1038495853$. We denote the right-hand side of the required inequality by $f(x)$. Then,

$$f'(t) = \frac{t}{\log t}\left(1 + \frac{1.05}{\log^3 t} - \frac{0.35}{\log^4 t} + \frac{3.5}{\log^5 t} + \frac{21}{\log^6 t} + \frac{1}{\log^7 t}\left(3522.8 - \frac{14873.45}{\log t}\right)\right) > 0 \tag{6.4}$$

for every $t \geq e^{4.23}$. Now, let $n = \pi(x)$. Then,

$$S(x) = \sum_{k\leq n} p_k = \pi(p_n)p_n - \left(np_n - \sum_{k\leq n} p_k\right). \tag{6.5}$$

Since $x \geq 1038495853$, we have $n \geq 52703656$. Now we use (6.5), Lemma 6.1 and 6.3 to get

$$S(x) \leq \pi(p_n)p_n - \frac{p_n^2}{2\log p_n} - \frac{3p_n^2}{4\log^2 p_n} - \frac{7p_n^2}{4\log^3 p_n} - \Theta(n) \leq f(p_n).$$

In addition to (6.4), the claim follows for every $x \geq 1038495853 = p_{52703656}$. A computer check shows that $f(p_i) \geq S(p_i)$ for every $30457 \leq i \leq 52703655$. Hence, $f(x) \geq S(x)$ for every $x \geq 356023 = p_{30457}$. Using a computer, we check that $f(355992) \geq 5171616645$. So $f(x) > S(x)$ for every $355992 \leq x \leq 356023$. $\square$

According to (6.1), we obtain the following corollary.

**Corollary 6.8.** *If $x \geq 355992$, then*

$$S(x) < \mathrm{li}(x^2) + \frac{0.525x^2}{\log^4 x} + \frac{0.875x^2}{\log^5 x} + \frac{3.9375x^2}{\log^6 x} + \frac{22.3125x^2}{\log^7 x} + \frac{1839.49375x^2}{\log^8 x}.$$

*Proof.* First, we consider the function

$$f(x) = \mathrm{li}(x) - \left( \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \frac{24x}{\log^5 x} + \frac{120x}{\log^6 x} + \frac{720x}{\log^7 x} + \frac{5040x}{\log^8 x} \right).$$

Then, $f(4171) \geq 0.00019$ and $f'(x) = 40320/\log^9 x$. Hence, $f(x) > 0$ for every $x \geq 4171$. Therefore,

$$\mathrm{li}(x^2) > \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x} + \frac{x^2}{4\log^3 x} + \frac{3x^2}{8\log^4 x} + \frac{3x^2}{4\log^5 x} + \frac{15x^2}{8\log^6 x} + \frac{45x^2}{8\log^7 x} + \frac{315x^2}{16\log^8 x}$$

for every $x \geq \sqrt{4171}$. So, the claim follows from Theorem 6.7 for every $x \geq 355992$. If $x = 355991$, the required inequality does not hold. $\square$

## 6.3 Lower bounds for $S(x)$

In this subsection, we find some lower bounds for the step function $S(x)$. The asymptoitic formula (6.2) implies that

$$S(x) \geq \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x} \tag{6.6}$$

for every sufficiently large $x$. In view of (6.6), Massias [13, Lemme 6] proved that

$$S(x) \geq \frac{x^2}{2\log x} + \frac{9x^2}{52\log^2 x} + \frac{x^2}{4\log^3 x} \tag{6.7}$$

for every $x \geq 11813$. In 1989, Massias, Nicolas and Robin [14, Lemma 3] improved (6.7) by showing that

$$S(x) \geq \frac{x^2}{2\log x} + \frac{0.477x^2}{2\log^2 x}$$

for every $x \geq 70001$ and that the inequality

$$S(x) \geq \frac{x^2}{2\log x} \exp\left(\frac{0.475}{\log x}\right)$$

holds for every $x \geq 4256233$. The currently best known lower bound for $S(x)$ is due to Massias and Robin [15, Théorème D(ii)]. They proved that the inequality

$$S(x) \geq \frac{x^2}{2\log x} + \frac{0.954x^2}{4\log^2 x} \tag{6.8}$$

holds for every $x \geq 70841$. We obtain the following proposition, which leads to an improvement of (6.8), where we use Lemma 6.6 and some explicit estimates for the prime counting function $\pi(x)$.

**Proposition 6.9.** *If $x \geq 65405363$, then*

$$S(x) > \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x} + \frac{x^2}{4\log^3 x} - \frac{3x^2}{20\log^4 x} - \frac{x^2}{8\log^5 x} - \frac{33x^2}{16\log^6 x} - \frac{267x^2}{16\log^7 x} - \frac{315065x^2}{128\log^8 x}.$$

*Proof.* First, we consider the case $x \geq 1332450001$. We denote the right-hand side of the required inequality by $f(x)$. Further, let $y = 1, f(t) = t$ and

$$a(n) = \begin{cases} 1 & \text{if } n \in \mathbb{P}, \\ 0 & \text{otherwise.} \end{cases}$$

We use Lemma 6.6 to get

$$S(x) = \sum_{1<n\leq x} a(n)f(n) = x\pi(x) - \int_1^x \pi(t)\,dt = x\pi(x) - 143 - \int_{27}^x \pi(t)\,dt.$$

Using the estimates for $\pi(x)$ given in Lemma 6.1 and 6.2, we obtain that

$$\begin{aligned} S(x) > {} & \frac{x^2}{\log x} + \frac{x^2}{\log^2 x} + \frac{2x^2}{\log^3 x} + \frac{5.65x^2}{\log^4 x} + \frac{23.65x^2}{\log^5 x} + \frac{118.25x^2}{\log^6 x} + \frac{709.5x^2}{\log^7 x} + \frac{4966.5x^2}{\log^8 x} - 143 \\ & - \int_{31}^x \left( \frac{t}{\log t} + \frac{t}{\log^2 t} + \frac{2t}{\log^3 t} + \frac{6.35t}{\log^4 t} + \frac{24.35t}{\log^5 t} + \frac{121.75t}{\log^6 t} + \frac{730.5t}{\log^7 t} + \frac{6801.4t}{\log^8 t} \right) dt. \end{aligned}$$

Now, by Lemma 2.5, Lemma 6.4 and Lemma 6.5, we get that

$$\begin{aligned} S(x) > {} & E_1 + \frac{953927x^2}{6300\log x} + \frac{953927x^2}{12600\log^2 x} + \frac{953927x^2}{12600\log^3 x} + \frac{949517x^2}{8400\log^4 x} + \frac{237563x^2}{1050\log^5 x} + \frac{59207x^2}{105\log^6 x} \\ & + \frac{117679x^2}{70\log^7 x} + \frac{4966.5x^2}{\log^8 x} - \frac{950777}{3150}\operatorname{li}(x^2), \end{aligned} \tag{6.9}$$

where

$$\begin{aligned} E_1 = {} & \frac{950777}{3150}\operatorname{li}(27^2) - \frac{947627 \cdot 27^2}{6300\log 27} - \frac{941327 \cdot 27^2}{12600\log^2 27} - \frac{928727 \cdot 27^2}{12600\log^3 27} - \frac{902057 \cdot 27^2}{8400\log^4 27} \\ & - \frac{425461 \cdot 27^2}{2100\log^5 27} - \frac{187163 \cdot 27^2}{420\log^6 27} - \frac{34007 \cdot 27^2}{35\log^7 27} - 143. \end{aligned}$$

By [3, Lemma 19], we have

$$\operatorname{li}(x^2) \leq \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x} + \frac{x^2}{4\log^3 x} + \frac{3x^2}{8\log^4 x} + \frac{3x^2}{4\log^5 x} + \frac{15x^2}{8\log^6 x} + \frac{45x^2}{8\log^7 x} + \frac{1575x^2}{64\log^8 x} \tag{6.10}$$

for every $x \geq 10^9$. Now we apply this inequality to (6.9) and we get that

$$S(x) > E_1 + f(x).$$

Since $E_1 \geq 110.232 > 0$, our proposition is proved for every $x \geq 1332450001$. We have

$$f'(x) \geq \frac{x}{\log x}\left(1 - \frac{1.05}{\log^3 x} - \frac{3.5}{\log^5 x} - \frac{21}{\log^6 x} - \frac{4806.078125}{\log^7 x}\right) > 0$$

for every $x \geq e^4$. So, we check with a computer that $S(p_i) \geq f(p_{i+1})$ for every $3862984 \leq i \leq \pi(1332450001) + 1 = 66773605$. Hence, $S(x) \geq f(x)$ for every $x \geq p_{3862984} = 65405363$. If $i = 3862983$, then $f(p_{i+1}) - S(p_i) = f(65405363) - S(65405357) > 2 \cdot 10^7 > 0$. Since $f$ is continuous and $S(x)$ is constant on $I = [p_{3862983}, p_{3862984})$, there is an $x_0 \in I$ so that $f(x) - S(x) > 0$ for every $x \in [x_0, p_{3862984})$. ☐

In 1988, Massias, Nicolas and Robin [14, Lemma 3(i)] proved that (6.6) holds for every $302791 \leq x \leq e^{90}$. Under the assumption that the Riemann hypothesis is true, Massias and Robin [15, Théorème D(iv)] showed in 1996 that the inequality (6.6) holds for every $x \geq 302971$. Further, they proved that the inequality (6.6) holds unconditionally for every $302971 \leq x \leq e^{98}$ and for every $x \geq e^{63864}$. With the following corollary, we close this gap.

**Corollary 6.10.** *For every $x \geq 302971$, we have*

$$S(x) > \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x}.$$

*Proof.* If $x \geq 685 \geq e^{6.529}$, then

$$\frac{1}{4} - \frac{3}{20\log x} - \frac{1}{8\log^2 x} - \frac{33}{16\log^3 x} - \frac{267}{16\log^4 x} - \frac{315065}{128\log^5 x} > 0.$$

So, the claim follows from Proposition 6.9 for every $x \geq 65405363$. Similarly to the proof of Proposition 6.9, we obtain that the required inequality holds for every $302971 \leq x < 65405363$ as well. □

Using (6.9) and (6.10), we find the following explicit lower bound for $S(x) - \mathrm{li}(x^2)$.

**Corollary 6.11.** *For every $x \geq 65405363$, we have*

$$S(x) > \mathrm{li}(x^2) - \frac{0.525x^2}{\log^4 x} - \frac{0.875x^2}{\log^5 x} - \frac{3.9375x^2}{\log^6 x} - \frac{22.3125x^2}{\log^7 x} - \frac{2486.0546875x^2}{\log^8 x}.$$

*Proof.* For every $x \geq 10^9$, the claim follows from Proposition 6.9 and (6.10). Similarly to the proof of Proposition 6.9, we obtain that the required inequality also holds for every $65405363 \leq x \leq 10^9$. □

# References


[1] T. Apostol, *Introduction to analytic number theory*, Springer, New York–Heidelberg, 1976.

[2] C. Axler, *Über die Primzahl-Zählfunktion, die n-te Primzahl und verallgemeinerte Ramanujan-Primzahlen*, PhD thesis, Düsseldorf, 2013. Available at http://docserv.uni-duesseldorf.de/servlets/DerivateServlet/Derivate-28284/pdfa-1b.pdf.

[3] C. Axler, *On a sequence involving prime numbers*, J. Integer Seq. **18** (2015), no. 7, Article 15.7.6, 13 pp.

[4] C. Axler, *New bounds for the prime counting function*, Integers **16** (2016), Paper No. A22, 15 pp.

[5] M. Cipolla, *La determinazione assintotica dell' $n^{imo}$ numero primo*, Rend. Accad. Sci. Fis-Mat. Napoli (3) **8** (1902), 132–166.

[6] M. Deléglise and J.-L. Nicolas, *Maximal product of primes whose sum is bounded*, preprint, 2012. Available at `http://arxiv.org/abs/1207.0603`.

[7] P. Dusart, *Autour de la fonction qui compte le nombre de nombres premiers*, Dissertation, Université de Limoges, 1998.

[8] P. Dusart, *The k-th prime is grater than $k(\ln k - \ln\ln k - 1)$ for $k \geq 2$*, Math. Comp. **68** (1999), no. 225, 411–415.

[9] P. Dusart, *Estimates of some functions over primes without R.H.*, preprint, 2010. Available at `http://arxiv.org/abs/1002.0442`.

[10] M. Hassani, *A remark on the Mandl's inequality*, preprint, 2006. Available at `http://arxiv.org/abs/math/0606765`.

[11] M. Hassani, *On the ratio of the arithmetic and geometric means of the prime numbers and the number e*, Int. J. Number Theory **9** (2013), no. 6, 1593–1603.

[12] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Vol. 1, Teubner, Leipzig und Berlin, 1909.

[13] J.-P. Massias, *Majoration explicite de l'ordre maximum d'un élément du groupe symétrique*, Annales de la faculté des sciences de Toulouse, Sér. 5, **6**, no. 3-4 (1984), 269–281.

[14] J.-P. Massias, J.-L. Nicolas and G. Robin, *Effective bounds for the maximal order of an element in the symmetric group*, Math. Comp. **53** (1989), no. 188, 665–678.

[15] J.-P. Massias and G. Robin, *Bornes effectives pour certaines fonctions concernant les nombres premiers*, Journal Th. Nombres de Bordeaux, Vol. **8** (1996), 213–238.

[16] G. Robin, *Permanence de relations de récurrence dans certains développements asymptotiques*, Publ. Inst. Math. (Beograd) (N.S.) **43** (57) (1988), 17–25.

[17] J. B. Rosser, *The $n$-th prime is greater than $n \log n$*, Proc. London Math. Soc., ser. 2, **45** (1939), 21–44.

[18] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6**:1 (1962), 64–94.

[19] J. B. Rosser and L.Schoenfeld, *Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$*, Collection of articles dedicated to Derrick Henry Lehmer on the occasion of his seventieth birthday, Math. Comp. **29** (1975), 243–269.

[20] N. K. Sinha, *On the asymptotic expansion of the sum of the first $n$ primes*, preprint, 2010. Available at `http://arxiv.org/abs/1011.1667`.

[21] N. J. A. Sloane, Sequence A007504, The on-line encyclopedia of integer sequences. `https://oeis.org/A007504`. Accessed 27 Apr 2016.

[22] N. J. A. Sloane, Sequence A124478, The on-line encyclopedia of integer sequences. `https://oeis.org/A124478`. Accessed 27 Apr 2016

[23] Z.-W. Sun, *On some inequalities for primes*, preprint, 2012. Available at `http://arxiv.org/abs/1209.3729`.

[24] M. Szalay, *On the maximal order in $S_n$ and $S_n^*$*, Acta Arith. **37** (1980), 321–331.