Distribution of the Largest Strong Goldbach Numbers Generated by Primes

Pingyuan Zhou[1] & Rong Ao[2]

[1] Beiyuan 35-210, Chengdu University of Technology, Chengdu, Sichuan, China

[2] 10 Oak Grove St. Ottawa, Ontario K2G6R4, Canada

Correspondence: Pingyuan Zhou, Beiyuan 35-210, Chengdu University of Technology, Sichuan 610059, China.
E-mail: pingyuanzhou49@163.com

Abstract

Using the first 4000000 primes to find L_n, the largest strong Goldbach number generated by the n-th prime P_n, we generalize a proposition in our previous work (Zhou 2017) and propose that $L_n \approx 2P_n$ and $L_n/2P_n < 1$ for sufficiently large P_n but the limit of $L_n/2P_n$ as $n \to \infty$ is 1, $L_n \approx P_n + n \log n$ and $L_n/(P_n + n \log n) > 1$ for sufficiently large P_n but the limit of $L_n/(P_n + n \log n)$ as $n \to \infty$ is 1. There are five corollaries of the generalized proposition for getting $L_n \to \infty$ as $n \to \infty$, which is equivalent to Goldbach's conjecture. If every step in distribution curve of L_n is called a Goldbach step, a study on the ratio of width to height for Goldbach steps supports the existence of above two limits but a study on distribution of Goldbach steps supports an estimation that $Q(n) \approx (1 + 1/\log \log n)n/\log n$ and the limit of $Q(n)/((1 + 1/\log \log n)n/\log n)$ as $n \to \infty$ is 1, where $Q(n)$ is the number of Goldbach steps, from which we may expect there are infinitely many Goldbach steps to imply Goldbach's conjecture.

Keywords: prime, largest strong Goldbach number, numerical evidence, Goldbach step, Goldbach's conjecture

2010 Mathematics Subject Classification: 11A41, 11B99

1. Introduction

In our previous work (Zhou 2017), we moved beyond traditional definition of Goldbach number (Montgomery, & Vaughan, 1975; Li 1999; Lu 2010) by introducing three new definitions. First, $G_n = p + q$ is defined as a Goldbach number generated by the n-th prime P_n for $n \geq 2$ if p and q are two odd primes not greater than P_n (although G_n is a number, there is a sequence (G_n), for example, (G_5) generated by $P_5 = 11$ is (6, 8, 10, 12, 14, 16, 18, 22) in which every term is a Goldbach number generated by P_5). Second, G_n is defined as a strong Goldbach number generated by P_n and written as S_n if all even numbers from 6 to G_n are Goldbach numbers generated by P_n (although S_n is a number, there is a sequence (S_n), for example, (S_5) generated by $P_5 = 11$ is (6, 8, 10, 12, 14, 16, 18) in which every term is a strong Goldbach number generated by P_5). Third, S_n is defined as the largest strong Goldbach number generated by P_n and written as L_n if $S_n + 2$ is not a Goldbach number generated by P_n (here L_n is always a number, for example, $L_5 = 18$). Proposition 3.2 in our work (Zhou 2017) proposed that $L_n - P_n \approx n \log n$, or equivalently $L_n \approx P_n + n \log n$ for all P_n. In this paper, we generalized the proposition based on the result of numerical calculation using the first 4000000 primes and provided a different approach for analyzing Goldbach's conjecture. We chose 4000000 because it is large enough to show tendency of L_n, yet small enough for our algorithm to generate data reasonably quickly. We examined the ratio of width to height for Goldbach steps and found an alternative approach to lead to the two limits in the generalized proposition. By studying distribution of Goldbach steps, we obtained an estimation such that $Q(n) \approx (1 + 1/\log \log n)n/\log n$ and the limit of $Q(n)/((1 + 1/\log \log n)n/\log n)$ as $n \to \infty$ is 1, where $Q(n)$ is the number of Goldbach steps.

2. Distribution of the Largest Strong Goldbach Numbers Generated by Sufficiently Large Primes

By definition of L_n, we have the following observations for sufficiently large primes.

Observation 2.1. *Distribution of the ratio $L_n/2P_n$ for $10000 \leq n \leq 4000000$.*

Figure 1 shows distribution of the ratio $L_n/2P_n$ for $10000 \leq n \leq 4000000$. From the distribution we see there is the same property such that $L_n/2P_n \approx 1$ and $L_n/2P_n < 1$ for these primes but there is an obvious and stable increasing tendency, so one may estimate the limit of $L_n/2P_n$ as $n \to \infty$ is 1, which reflects the fact that $2P_n$ is the largest Goldbach number generated by P_n (Zhou 2017).
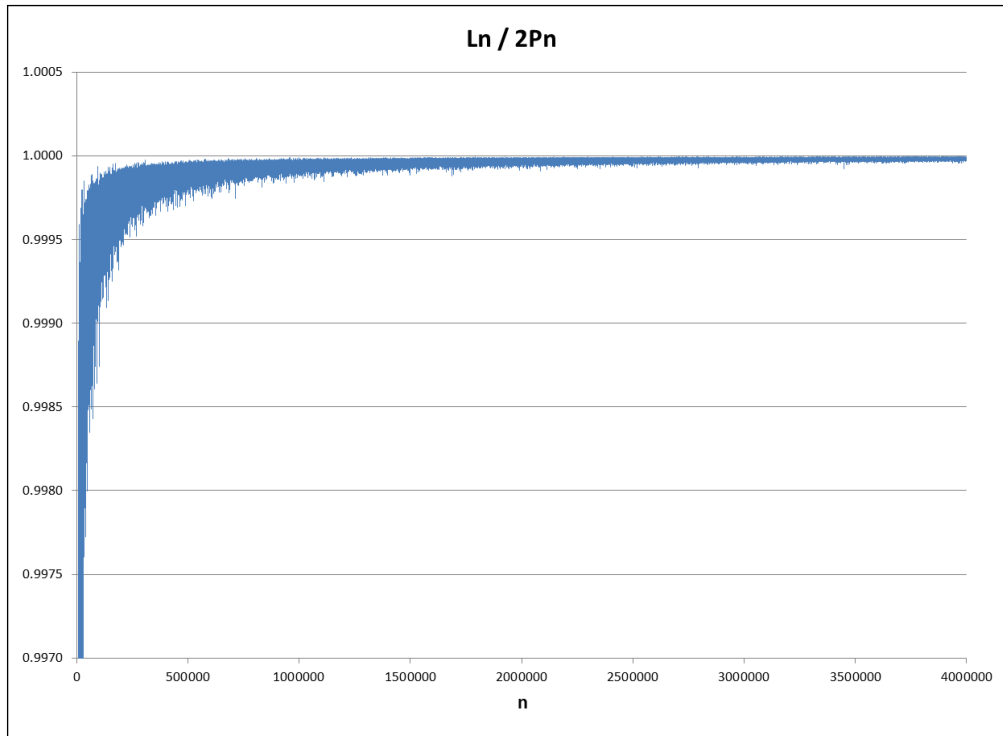
Figure 1. Distribution of the ratio $L_n/2P_n$ for $10000 \le n \le 4000000$

Observation 2.2. *Distribution of the ratio $L_n/(P_n + n \log n)$ for $10000 \le n \le 4000000$.*

Figure 2 shows distribution of the ratio $L_n/(P_n + n \log n)$ for $10000 \le n \le 4000000$. From the distribution we see there is the same property such that $L_n/(P_n + n \log n) \approx 1$ and $L_n/(P_n + n \log n) > 1$ for these primes but there is an obvious and stable decreasing tendency, so one may estimate the limit of $L_n/(P_n + n \log n)$ as $n \to \infty$ is 1.

Based on the above observations, we have the following proposition to lead to Goldbach's conjecture if the proposition is proven.

Proposition 2.3. $L_n \approx 2P_n$ *and $L_n/2P_n < 1$ for sufficiently large P_n but the limit of $L_n/2P_n$ as $n \to \infty$ is 1, $L_n \approx P_n + n \log n$ and $L_n/(P_n + n \log n) > 1$ for sufficiently large P_n but the limit of $L_n/(P_n + n \log n)$ as $n \to \infty$ is 1.*

Remark 2.4. Proposition 2.3 is a generalization of our Proposition 3.2 (Zhou 2017). Any of the following five approaches would lead to $L_n \to \infty$ as $n \to \infty$, which is equivalent to Goldbach's conjecture (Zhou 2017).

Corollary 2.5. *If $L_n \approx 2P_n$ as $n \to \infty$, then $L_n \to \infty$ as $n \to \infty$.*

Proof. Since $2P_n \to \infty$ as $n \to \infty$ and $L_n \approx 2P_n$ as $n \to \infty$, replace $2P_n$ with L_n, so $L_n \to \infty$ as $n \to \infty$. Thus the corollary holds.

Corollary 2.6. *If the limit of $L_n/2P_n$ as $n \to \infty$ is 1, then $L_n \to \infty$ as $n \to \infty$.*

Proof. Since $2P_n \to \infty$ as $n \to \infty$ and $L_n/2P_n \to 1$ as $n \to \infty$, replace $2P_n$ with L_n, so $L_n \to \infty$ as $n \to \infty$. Thus the corollary holds.

Corollary 2.7. *If $L_n \approx P_n + n \log n$ as $n \to \infty$, then $L_n \to \infty$ as $n \to \infty$.*

Proof. It has been proven in our work (Zhou 2017).

Corollary 2.8. *If $L_n/(P_n + n \log n) > 1$ as $n \to \infty$, then $L_n \to \infty$ as $n \to \infty$.*

Proof. Since $P_n + n \log n \to \infty$ as $n \to \infty$ and $L_n > P_n + n \log n$ as $n \to \infty$, we have $L_n \to \infty$ as $n \to \infty$. Thus the corollary holds.

Corollary 2.9. *If the limit of $L_n/(P_n + n \log n)$ as $n \to \infty$ is 1, then $L_n \to \infty$ as $n \to \infty$.*

Proof. Since $P_n + n \log n \to \infty$ as $n \to \infty$ and $L_n/(P_n + n \log n) \to 1$ as $n \to \infty$, replace $P_n + n \log n$ with L_n, so $L_n \to \infty$ as $n \to \infty$. Thus the corollary holds.
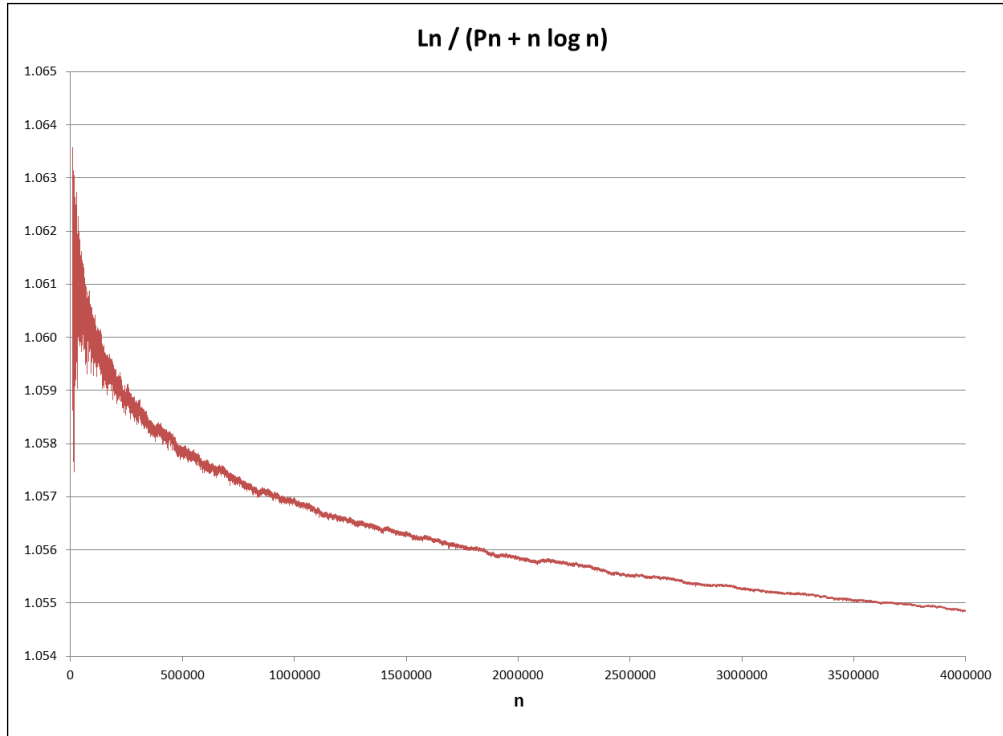
Figure 2. Distribution of the ratio $L_n/(P_n + n \log n)$ for $10000 \le n \le 4000000$

Our algorithm generated the following data for Figure 1 and Figure 2 at n = 10000, 100000, 1000000, 2000000, 3000000, 4000000: $L_n/2P_n$ = 0.99746011, 0.99958221, 0.99997443, 0.99996866, 0.99997855, 0.99998341, and $L_n/(P_n + n \log n)$ = 1.06144108, 1.06011031, 1.05697891, 1.05585627, 1.05526880, 1.05485514. It is obvious that L_n is closer to $2P_n$ than $P_n + n \log n$, but we will use both $2P_n$ and $P_n + n \log n$ to study the ratio of width to height for steps in the distribution curve of L_n.

3. Distribution of the Ratio of Width to Height for Goldbach Steps

It could be derived from the definition of strong Goldbach number that if S_n is a strong Goldbach number generated by P_n, then S_n is also a strong Goldbach number generated by P_{n+k} for $k \ge 1$ (Zhou 2017). Thus we have $L_n \le L_{n+1}$ for any given P_n so that the distribution of the largest strong Goldbach numbers generated by primes is a step-shaped curve growing without bound as Figure 3 shows.

Definition 3.1. Every step in the distribution curve of the largest strong Goldbach numbers L_n generated by primes P_n for $n \ge 2$ is called a *Goldbach step*.

Definition 3.2. For a given Goldbach step, W is called *width* of the Goldbach step if $W = n_2 - n_1$, where n_1 is n-value at the beginning of the Goldbach step and n_2 is n-value at the beginning of next Goldbach step.

Definition 3.3. For a given Goldbach step, H is called *height* of the Goldbach step if $H = L_n/\log L_n$, where L_n is generated by P_n with $n = n_1$ being n-value at the beginning of the Goldbach step.

According to Definition 3.2 and Definition 3.3, the ratio of width to height for a given Goldbach step is $W/H = ((n_2 - n_1) \log L_n)/L_n$. Figure 4 shows the distribution of $W/H = ((n_2 - n_1) \log L_n)/L_n$ for $2 \le n \le 4000000$, Figure 5 shows the distribution of $((n_2 - n_1) \log 2P_n)/2P_n$ for $2 \le n \le 4000000$ and Figure 6 shows the distribution of $((n_2 - n_1) \log (P_n + n \log n))/(P_n + n \log n)$ for $2 \le n \le 4000000$. Comparing Figure 5 and Figure 6 with Figure 4 respectively, we see that the figures are almost identical. By Figure 4, Figure 5 and Figure 6, it is obvious that $W/H \approx ((n_2 - n_1) \log 2P_n)/2P_n$ and $W/H \approx ((n_2 - n_1) \log (P_n + n \log n))/(P_n + n \log n)$ but there is a tendency such that the limit of $((n_2 - n_1) \log L_n)/L_n$ as $n \to \infty$ is 0. Therefore, one may expect that if the limit of $((n_2 - n_1) \log L_n)/L_n$ as $n \to \infty$ is 0 then $n_2 - n_1$ must be smaller than $L_n/\log L_n$ for any given Goldbach step so that there is always a new and higher Goldbach step to follow any given Goldbach step and Goldbach's conjecture is true.
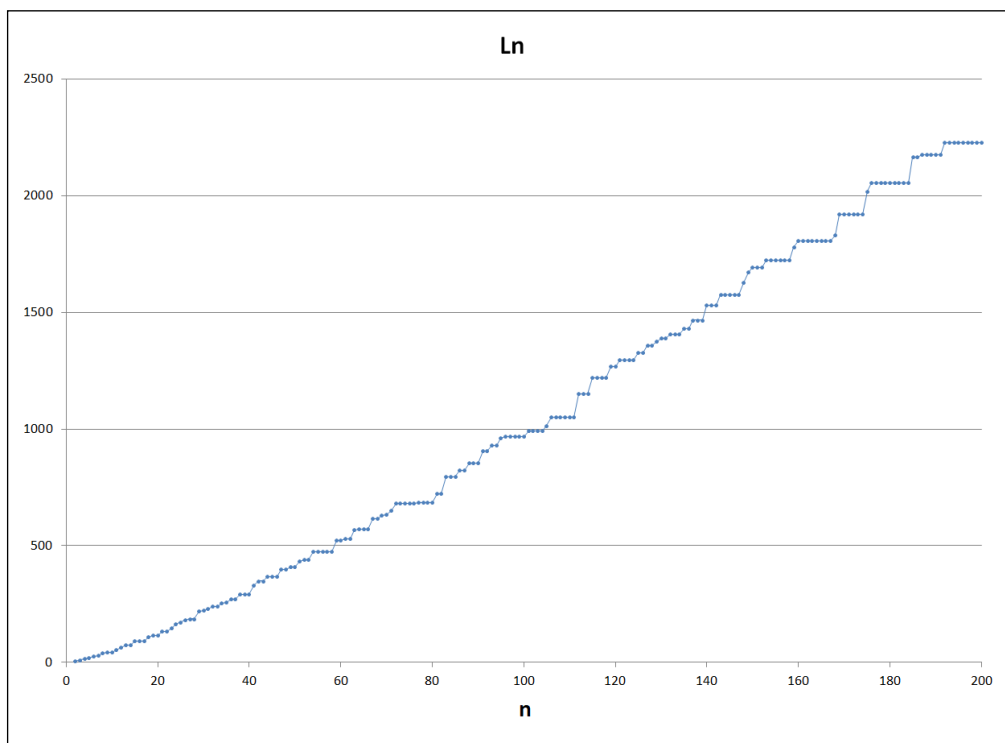
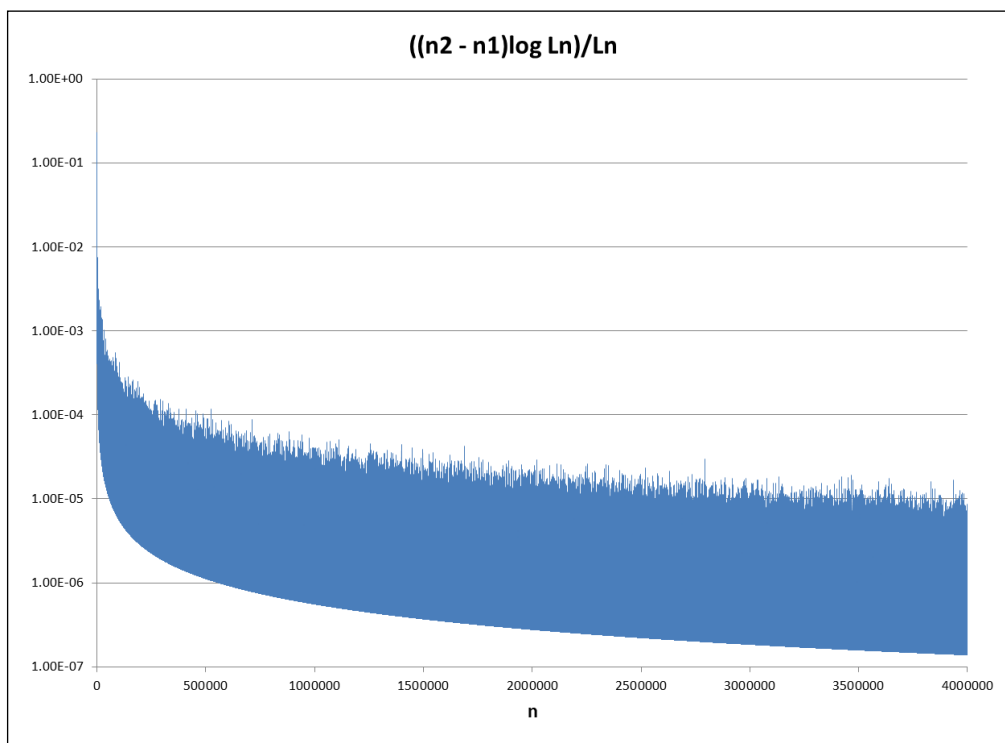Figure 3. Step-shaped distribution of L_n for $2 \leq n \leq 200$

Figure 4. Distribution of $((n_2 - n_1)\log L_n)/L_n$ for $2 \leq n \leq 4000000$

In order to prove the limit of $((n_2 - n_1)\log L_n)/L_n$ as $n \to \infty$ is 0, we obtain a set $\{A, B, C\}$, where $A = ((n_2 - n_1) \log 2P_n)/2P_n$, $B = ((n_2 - n_1)\log L_n)/L_n$ and $C = ((n_2 - n_1)\log (P_n + n \log n))/(P_n + n \log n)$. If it is proven that A is lower bound of the set and the limit of $((n_2 - n_1)\log 2P_n)/2P_n$ as $n \to \infty$ is 0 but C is upper bound of the set and the limit of $((n_2 - n_1)\log (P_n + n \log n))/(P_n + n \log n)$ as $n \to \infty$ is 0, then the limit of $B = ((n_2 - n_1)\log L_n)/L_n$ as $n \to \infty$ is 0.
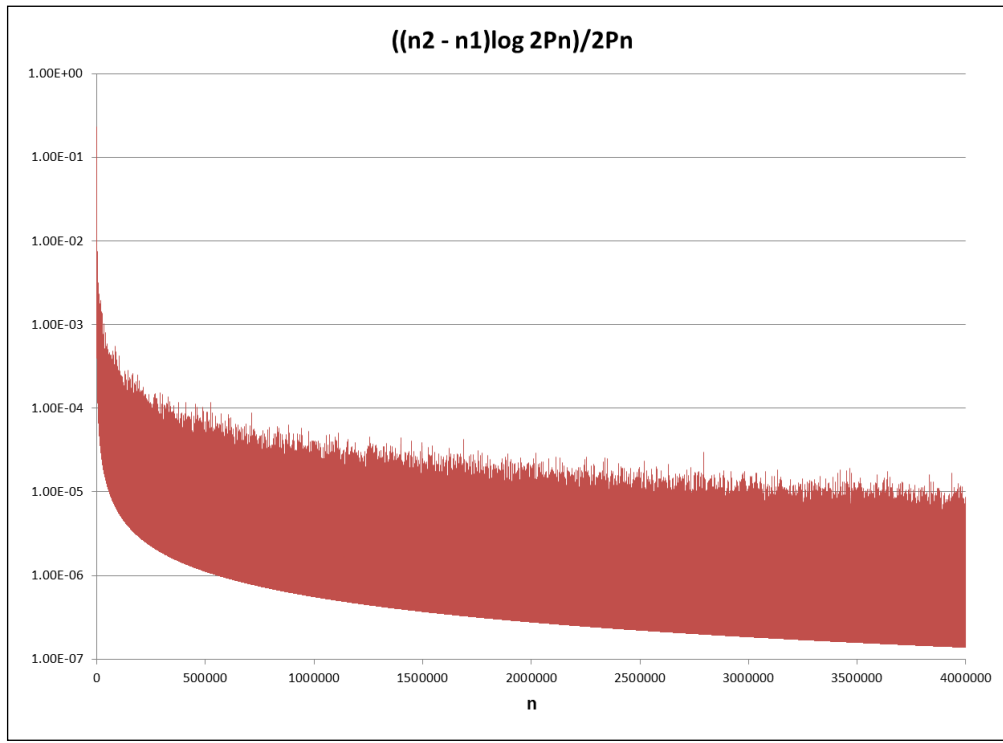
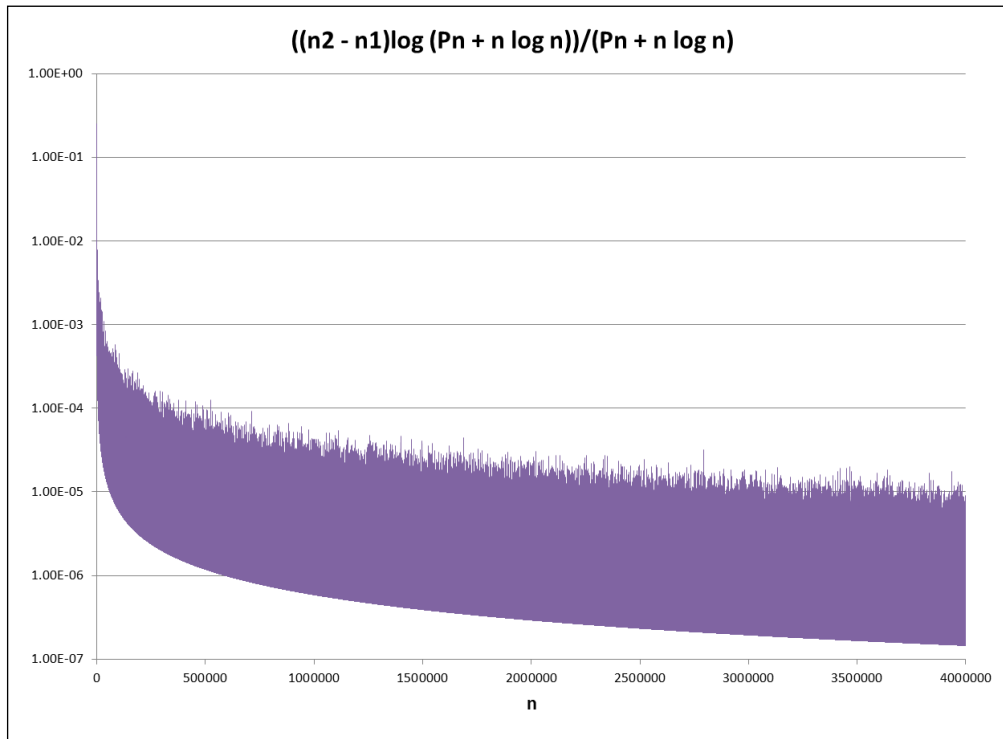Figure 5. Distribution of $((n_2 - n_1)\log 2P_n)/2P_n$ for $2 \leq n \leq 4000000$

Figure 6. Distribution of $((n_2 - n_1)\log(P_n + n \log n))/(P_n + n \log n)$ for $2 \leq n \leq 4000000$

Suppose it has been proven that the limit of $B = ((n_2 - n_1)\log L_n)/L_n$ as $n \to \infty$ is 0. Since the limit of A/B as $n \to \infty$ is 1 but $A/B = (L_n/2P_n)(\log 2P_n)/\log L_n$, if the limit of $(\log 2P_n)/\log L_n$ as $n \to \infty$ is 1 then the limit of $L_n/2P_n$ as $n \to \infty$ is 1. Since the limit of C/B as $n \to \infty$ is 1 but $C/B = (L_n/(P_n + n \log n))(\log (P_n + n \log n))/\log L_n$, if the limit of $(\log (P_n + n \log n))/\log L_n$ as $n \to \infty$ is 1 then the limit of $L_n/(P_n + n \log n)$ as $n \to \infty$ is 1. Thus above results are equivalent to the two limits in Proposition 2.3.
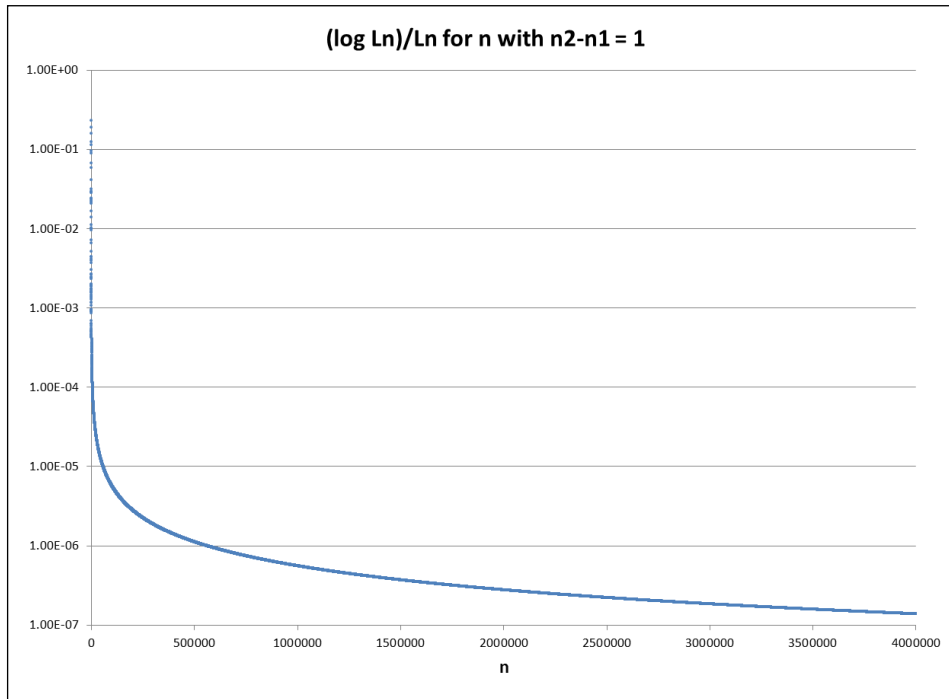

Figure 7. Distribution of $(\log L_n)/L_n$ for n with $n_2 - n_1 = 1$ for $2 \le n \le 4000000$

Further, we study bottom curves in Figure 4, Figure 5 and Figure 6 because these curves look smooth and there is the same tendency such that the three curves will approach 0 as n goes to infinity. First, we should consider the bottom curve of $A = ((n_2 - n_1)\log 2P_n)/2P_n$ in Figure 5 and we are sure that the curve is formed by minimum values of A. Let $n_2 - n_1 = 1$. Then we obtain $A_{min} = (\log 2P_m)/2P_m$ if m denotes the n-value at which $(\log 2P_n)/2P_n$ is a minimum value of A. It is obvious that $2P_m/\log 2P_m$ is the number of primes among all integers from 1 to $2P_m$, therefore, $1/A_{min} = 2P_m/\log 2P_m$ is just the number of primes not greater than $2P_m$. Similarly, we have $1/C_{min} = (P_m + m \log m)/\log (P_m + m \log m)$ if m denotes the n-value at which $(\log (P_n + n \log n))/(P_n + n \log n)$ is a minimum value of C in Figure 6 but we have $1/B_{min} = L_m/\log L_m$ if m denotes the n-value at which $(\log L_n)/L_n$ is a minimum value of B in Figure 4. Finally, Figure 7 shows the distribution of $(\log L_n)/L_n$ for n with $n_2 - n_1 = 1$ for $2 \le n \le 4000000$, that is, the distribution of $B_{min} = (\log L_m)/L_m$ for $2 \le n \le 4000000$. Comparing Figure 7 with Figure 4, we see the curve in Figure 7 is just the bottom curve in Figure 4.

4. Distribution of Goldbach Steps

Let $Q(n)$ denote the number of Goldbach steps formed by the first n primes P_n (the number is just the number of n_1 among the first n primes P_n). Then we have the following observations.

Observation 4.1. *Distribution of $Q(n)$ for $2 \le n \le 4000000$.*

Figure 8 shows the distribution of $Q(n)$ for $2 \le n \le 4000000$, which is a curve with continuous growths. In the figure, another curve shows the distribution of $(1 + 1/\log \log n)n/\log n$, in which $n/\log n$ denotes the number of primes among the first n positive integers.

Observation 4.2. *Distribution of the ratio $Q(n)/((1 + 1/\log \log n)n/\log n)$ for $2 \le n \le 4000000$.*

Figure 9 shows the distribution of $Q(n)/((1 + 1/\log \log n)n/\log n)$ for $2 \le n \le 4000000$. From the distribution we see $Q(n)/((1 + 1/\log \log n)n/\log n) \approx 1$ for these n-values and there is a decreasing tendency, so one may estimate the limit of $Q(n)/((1 + 1/\log \log n)n/\log n)$ as $n \to \infty$ is 1.

Based on the above two observations, we have the following proposition.

Proposition 4.3. *$Q(n) \approx (1 + 1/\log \log n)n/\log n$ and the limit of $Q(n)/((1 + 1/\log \log n)n/\log n)$ as $n \to \infty$ is 1.*

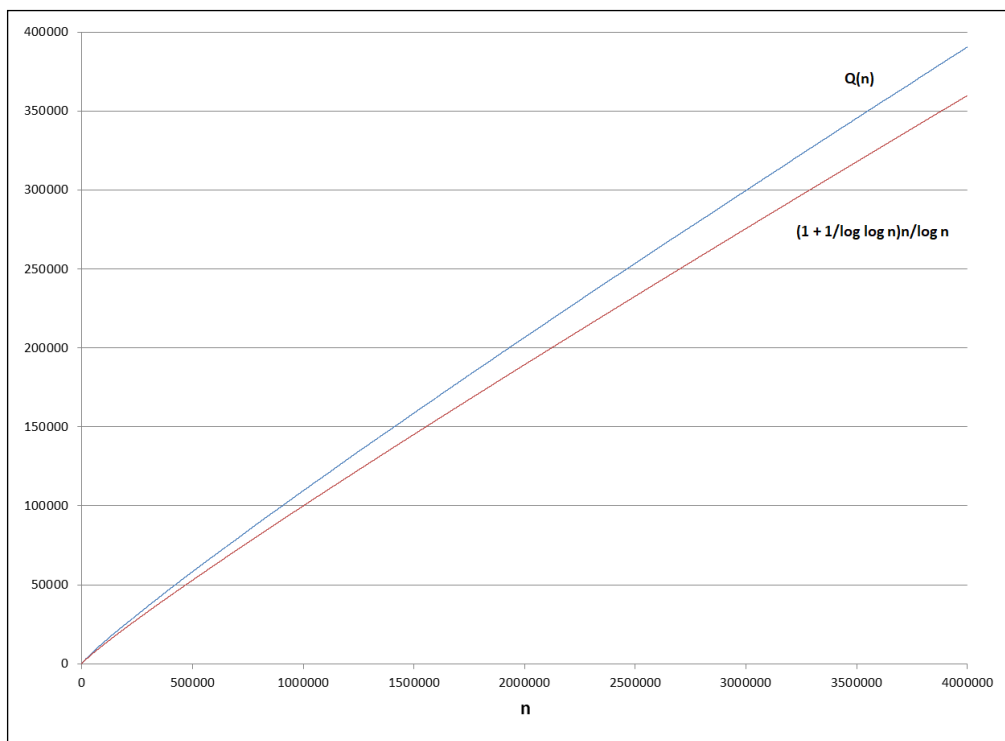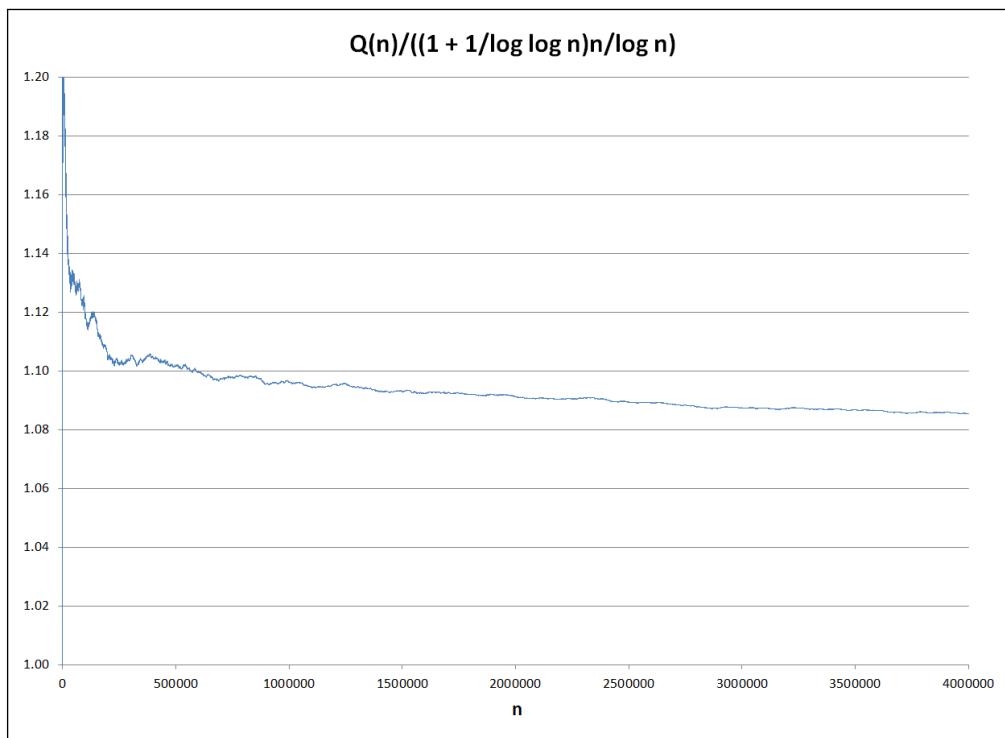Figure 8. Distributions of $Q(n)$ and $(1 + 1/\log \log n)n/\log n$ for $2 \le n \le 4000000$

Figure 9. Distribution of the ratio $Q(n)/((1 + 1/\log \log n)n/\log n)$ for $2 \le n \le 4000000$

7

Corollary 4.4. *If Q(n) ≈ (1 + 1/log log n)n/log n as n → ∞, then L_n → ∞ as n → ∞.*

Proof. Since (1 + 1/log log n)n/log n → ∞ as n → ∞ and Q(n) ≈ (1 + 1/log log n)n/log n as n → ∞, replace (1 + 1/log log n)n/log n with Q(n), so Q(n) → ∞ as n → ∞. The result means there are infinitely many Goldbach steps, which obviously implies L_n → ∞ as n → ∞. Thus the corollary holds.

Corollary 4.5. *If the limit of Q(n)/((1 + 1/log log n)n/log n) as n → ∞ is 1, then L_n → ∞ as n → ∞.*

Proof. Since (1 + 1/log log n)n/log n → ∞ as n → ∞ and Q(n)/((1 + 1/log log n)n/log n) → 1 as n → ∞, replace (1 + 1/log log n)n/log n with Q(n), so Q(n) → ∞ as n → ∞. The result means there are infinitely many Goldbach steps, which obviously implies L_n → ∞ as n → ∞. Thus the corollary holds.

Remark 4.6. If Proposition 4.3 is proven then there are infinitely many Goldbach steps to imply L_n → ∞ as n → ∞ by Corollary 4.4 or Corollary 4.5, therefore, Goldbach's conjecture is true.

5. Conclusion

In this paper, we presented studies on distributions of $L_n/2P_n$ and $L_n/(P_n + n \log n)$, distribution of the ratio of width to height for Goldbach steps and distribution of Goldbach steps using numerical evidence for the first 4000000 primes. We obtained Proposition 2.3 which is a generalization of previous Proposition 3.2 (Zhou 2017). If any of the five approaches derived from Proposition 2.3 is proven, then Goldbach's conjecture is also proven true. By ratio of width to height for Goldbach steps, we assumed the limit of $B = ((n_2 − n_1)\log L_n)/L_n$ as n → ∞ is 0, which will be equivalent to the limit of $L_n/2P_n$ as n → ∞ being 1 and the limit of $L_n/(P_n + n \log n)$ as n → ∞ being 1 in Proposition 2.3, if it is proven that the limit of $A = ((n_2 − n_1)\log 2P_n)/2P_n$ as n → ∞ is 0 and the limit of $C = ((n_2 − n_1)\log (P_n + n \log n))/(P_n + n \log n)$ as n → ∞ is 0 but A and C are lower and upper bounds of the set {A, B, C}. By distribution of Goldbach steps, we discovered a pattern that $Q(n) ≈ (1 + 1/\log \log n)n/\log n$ to support the existence of the limit of $Q(n)/((1 + 1/\log \log n)n/\log n)$ as n → ∞ being 1, which means there may be infinitely many Goldbach steps to imply Goldbach's conjecture.

Acknowledgments

References

Li, H. (1999). The Exceptional Set of Goldbach Numbers. *Quart J Math Oxford(2)*, *50*, 471-482. https://doi.org/10.1093/qjmath/50.200.471

Lu, W. C. (2010). Exceptional set of Goldbach number. *J Number Theory, 130*(10), 2359-2392. https://doi.org/10.1016/j.jnt.2010.03.017

Montgomery, H. L., & Vaughan, R. C. (1975). The exceptional set in Goldbach's problem. *Acta Arith, 27*, 353-370. https://doi.org/10.4064/aa-27-1-353-370

Zhou, P. (2017). Strong Goldbach Number in Goldbach's Problem. *Journal of Mathematics Research, 9*(6), 95-105. https://doi.org/10.5539/jmr.v9n6p95