 doi:10.20944/preprints202110.0087.v8

ON THE ADDITIVE AND SUBTRACTIVE REPRESENTATION OF EVEN NUMBERS FROM PRIMES

ALI SHEHU [1] AND JETMIRA UKA [2]

[1] Queen Mary University of London, United Kingdom
a.shehu@qmul.ac.uk
[2] Brunel University London, United Kingdom
jetmira.uka@brunel.ac.uk

ABSTRACT. We demonstrate a new quantitative method to the sieve of Eratosthenes, which is an alternative to the sieve of Legendre. In this method, every element of a given set is sifted out once only, and therefore, this method is free of the Möbius function and of the parity barrier. Using this method, we prove that every sufficiently large even number is the sum of two primes, and that every even number is the difference of two primes in infinitely many ways.

Keywords: Sieve of Eratosthenes; Goldbach conjecture; Polignac conjecture; Twin Prime conjecture

1. INTRODUCTION

1.1. **Representing even numbers from primes.** Since the set of prime numbers is infinite, and since all prime numbers ≥ 3 are odd numbers, then one knows immediately that there are infinitely many even numbers that can be represented as the sum of two primes, and, infinitely many even numbers that can be represented as the difference of two primes. Having said that, the following questions then naturally arise:

(i) Can every even number be represented as the sum of two primes?
(ii) Can every even number be represented as the difference of two primes?
(iii) Can any even number, or indeed all even numbers, be represented as the difference of two primes, in infinitely many ways?
(iv) If questions (ii) and (iii) are answered in the affirmative, can they also hold for consecutive primes, in representing even numbers ≥ 6 as the difference of two primes?

The earliest known record to have posed question (i), known as the Goldbach conjecture, dates back to 1742, in a correspondence between C. Goldbach and L. Euler, where it is propositioned that every even number > 2 can be represented as the sum of two primes [1]. The Goldbach conjecture has more lately become known as the binary Goldbach conjecture, or the strong Goldbach conjecture, in order to distinguish it from the ternary Goldbach conjecture, or the weak Goldbach conjecture, which states that every odd number > 5 can be represented as the sum of three primes. The binary Goldbach conjecture has to date been shown empirically to hold for every even number $\leq 4 \cdot 10^{18}$ [2], however, a rigorous proof, or disproof, remains elusive.

Nevertheless, some related theoretical results to the binary Goldbach conjecture have been achieved, the closest of which is due to J. R. Chen, who in 1973, proved that every sufficiently large even number can be represented as the sum of a prime and another prime or a semiprime, that is, the product of at most two primes [3] [4] [5] [6]. On the other hand, significant results

 doi:10.20944/preprints202110.0087.v8

have been achieved for the ternary Goldbach conjecture, culminating with a proof given by H. Helfgott in 2014 [7] [8] [9].

The earliest known record to have posed questions (ii)-(iv) are due to Alphonse de Polignac, who in 1849, conjectured that every even number can be represented as the difference of two consecutive primes, in infinitely many ways [10]. The most significant special case of Polignac's conjecture, is the so-called 'twin prime conjecture', which comprises of the number 2 being represented as the difference of two primes, in infinitely many ways. The twin prime conjecture is hugely supported by empirical data, where, over the past few decades, increasingly large twin prime pairs have been found to exist [11], with the current world record for a twin prime pair, found in the year 2016, standing at 388, 342 decimal digits long [12].

As with the binary Goldbach Conjecture, the closest theoretical result to Polignac's conjecture is given by J. R. Chen, who in the same article, proved that every even number can be represented as the difference of a prime and another prime or semiprime, that is, the product of at most two primes [6]. More recently, D. A. Goldston, J. Pintz, and C. Y. Yildirim, introduced the now known 'GPY method', which uses approximations to the prime k-tuples conjecture, to study small numbers that can be represented as the difference of two primes [13]. In 2013, Yitang Zhang built on the GPY method, thereby proving for the first time the existence of some even number $< 7 \cdot 10^7$, which can be represented as the difference of two primes in infinitely many ways [14]. Within a year after Zhang's result, J. Maynard presented an independent proof that lowered the bound to 600, which, by assuming the Elliott-Halberstam conjecture, could be further lowered down to 12 [15]. With some refinements to Zhang's method and a combination of that with Maynard's approach, the bound was lowered to 246 unconditionally, by an on-line collaborative project known as *Polymath 8*, organised by T. Tao [16].

1.2. **Sieve theory.** Sieve theory is a technique for distinguishing specific subsets of integers, amongst the set of natural numbers. Sieve theory began with Eratosthenes of Cyrene (276-194 B.C.), who constructed a method with which one could isolate the subset of prime numbers, from the set of natural numbers [17]. It starts by first crossing the multiples of 2 in the number line up to x, then the multiples of 3, then the multiples of 5, and then the multiples of all the primes $\leq \sqrt{x}$. If an integer $n > 1$ is not divisible by any prime $p \leq \sqrt{x}$, then n is necessarily a prime. Upon completion of the sieve, one has

$$\#\{P \in \mathbb{P} : P \leq x\} = \#\left\{n \leq x : P \nmid n, P \leq \sqrt{x}\right\}. \tag{1.1}$$

Having the means of constructing the complete subset of prime numbers, from the set of natural numbers up to x, one would be naturally interested in quantifying these primes. To this end, there are fundamentally two approaches that one could take, in order to quantify the set of primes generated by the sieve of Eratosthenes. We describe each approach, as follows.

(i) In the first approach, one quantifies the subset of the integers that are sifted out at each round of the sieve, as

$$\frac{\#\{\text{multiples of } P \text{ up to } x\}}{\#\{\text{integers up to } x\}}.$$

The easy part with this approach, is that one has no problem in defining quantitatively both the subset of the multiples of P up to x and the set of all integers up to x. However, the difficult part is due to the fact that those integers which have multiple prime factors, are necessarily counted at multiple rounds of the sieve, which must be accounted for. This is the approach taken by A. Legendre (1752-1833), who was the first to turn the sieve of Eratosthenes into a quantitative technique, and this has been the approach taken ever since. In the sieve of Legendre, one counts

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 August 2022 doi:10.20944/preprints202110.0087.v8

the integers that are crossed out at each round, and then one subtracts those that are counted at multiple rounds, according to the multiplicity of times that this has happened, as follows

$$[x] - \sum_{P\leq\sqrt{x}}\left[\frac{x}{P}\right] + \sum_{P_1<P_2\leq\sqrt{x}}\sum\left[\frac{x}{P_1P_2}\right] - \sum_{P_1<P_2<P_3\leq\sqrt{x}}\sum\sum\left[\frac{x}{P_1P_2P_3}\right] + - \cdots, \tag{1.2}$$

from which, one obtains Legendre's formula

$$\pi(x) - \pi\left(\sqrt{x}\right) + 1 = \sum_{\substack{d \\ P|d\rightarrow P\leq\sqrt{x}}} \mu(d)\left[\frac{x}{d}\right], \tag{1.3}$$

where $\mu(d)$ is the Möbius function, introduced by A. F. Möbius (1790–1868), and defined as

$$\mu(n) := \begin{cases} 1 & x \text{ is square-free and has an even number of prime factors,} \\ -1 & x \text{ is square-free and has an odd number of prime factors,} \\ 0 & x \text{ is not square-free,} \end{cases} \tag{1.4}$$

In efforts to evaluate the right-hand side of (1.3), one has

$$\pi(x) - \pi\left(\sqrt{x}\right) + 1 = x\sum_{d}\frac{\mu(d)}{d} + R = x\prod_{P=P_1}^{P\leq\sqrt{x}}\frac{P-1}{P} + R, \tag{1.5}$$

where the remainder R is

$$R = -\sum_{d}\mu(d)\left\{\frac{x}{d}\right\}, \tag{1.6}$$

which doubles at each round of the sieve, and thus quickly becomes larger than x. In modern sieve methods, one tries to replace $\mu(n)$ by a function $\Lambda = (\lambda_d)$, referred to as the "sieve weights", which mimics the $\mu(n)$ and gives satisfactory estimates on upper bounds, lower bounds, and asymptotics for smoother number sets such as almost primes, which leads to upper bounds for primes. However, obtaining lower bounds for primes has proved much more difficult, thus leaving Goldbach's and Polignac's conjectures out of reach. This has been due to the so-called "parity barrier", where the sieve struggles to distinguish integers with an odd number of prime factors from integers with an even number of prime factors. Although the parity barrier has been broken for certain specific sequences, it still remains the fundamental obstacle in the treatment of primes via this approach of sieve theory.

(ii) In the second approach, one quantifies the subset of the integers that are sifted out at each round of the sieve, as

$$\frac{\#\{\text{multiples of } P \text{ up to } x \text{ that survived the preceding rounds}\}}{\#\{\text{integers up to } x \text{ that survived the preceding rounds}\}}.$$

This means that every integer is sifted out according to its least prime factor, out of a set of integers whose least prime factors are greater than, or equal to, the least prime factors of the integers being sifted out at the given round, with the exception of the number 1. This is best illustrated by the following example: In the 1st round of the sieve, one quantifies the subset of the multiples of 2, as a ratio over the set of all integers, that is

$$\frac{\#\{2,4,6,8,10,12,14,16,18,20,22,24,26\}}{\#\{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26\}}.$$

 doi:10.20944/preprints202110.0087.v8

In the 2nd round, one quantifies the subset of the multiples of 3 that survived the preceding round of the sieve, as a ratio over the set of all the integers that survived the preceding round of the sieve, that is

$$\frac{\#\{3, 9, 15, 21\}}{\#\{1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25\}}.$$

In the 3rd round, one quantifies the subset of the multiples of 5 that survived the preceding rounds of the sieve, as a ratio over the set of all the integers that survived the preceding rounds of the sieve, that is

$$\frac{\#\{5, 25\}}{\#\{1, 5, 7, 11, 13, 17, 19, 23, 25\}}.$$

Since that completes the sieve, then the subset of integers that survives the sieve is

$$\{1, 7, 11, 13, 17, 19, 23\}$$

which consists of the number 1 and the complete set of primes in the interval $[\sqrt{26}, 26]$.

The easy part with this approach, is that every integer is sifted out once only, and so one does not have the problem of certain integers appearing at multiple rounds of the sieve, which one has with the first approach, as described above. However, the difficult part with this approach, is that at each round of the sieve, one finds it difficult to define quantitatively, both the subset of the multiples of P that survive the preceding rounds of the sieve and the subset of all the integers that survive the preceding rounds of the sieve. Nevertheless, since at each round of the sieve, the subset of the multiples of P that survive the preceding rounds of the sieve consists of the complete set of integers ($\leq x$) for which P is the least prime factor, and, the subset of all the integers that survive the preceding rounds of the sieve contains the complete set of primes in the interval $[\sqrt{x}, x]$, then one is able to define quantitatively the two subsets of integers, at least to the extent where one can then determine upper and/or lower bounds.

1.3. **Our results.** In the present paper, we take the second approach as outlined above, in order to describe quantitatively the sieve of Eratosthenes. On the condition that x is sufficiently large; at each round of the sieving of integers, we define quantitatively both the subset of the multiples of P up to x that survive the preceding rounds of the sieve and the subset of all the integers up to x that survive the preceding rounds of the sieve; to the extent where we are then able to determine an upper bound to the proportion of integers that are sifted out at each round of the sieve. To the best of our knowledge, we are the first to take this approach for these purposes.

Definition 1 (Additive representation)**.** *Let $p, q \in \mathbb{N}$, let $p \leq q$, and let x be an even number, then we say that $p + q$ is an additive representation of x, if $p + q = x$.*

Definition 2 (Subtractive representation)**.** *Let $p, q \in \mathbb{N}$, and let $2k$ be an even number, then we say that $q - p$ is a subtractive representation of $2k$, if $q - p = 2k$.*

We then extend our approach taken in the sieving of integers, to the sieving of representations, whereby we quantify the subset of the additive representations of x, and the subset of the subtractive representations of $2k$ up to x, that survive the sieve of Eratosthenes. As with the sieving of integers, we quantify the subset of representations that is sifted out at each round of the sieve, as

$$\frac{\#\{\text{representations, containing multiples of } P \text{ up to } x, \text{ that survived the preceding rounds}\}}{\#\{\text{representations, containing integers up to } x, \text{ that survived the preceding rounds}\}}.$$

This means that additive representations, and subtractive representations, are sieved out according to either p or q, depending on whose least prime factor is the smaller (or equal to) of the two.

 doi:10.20944/preprints202110.0087.v8

This allows for every representation that contains at least one multiple of P to be sifted out once only, while every representation that does not contain multiples of P, survives the sieve. Therefore, upon completion of the sieve, we have a subset of representations, where $p = 1$ or a prime in the interval $[\sqrt{x}, x]$ and q is a prime in the interval $[\sqrt{x}, x]$, with the quantity of representations where $p = 1$ being at most one.

However, contrary to the sieving of integers, in the sieving of additive representations, and in the sieving of subtractive representations, at each round of the sieve, the subset of representations containing multiples of P up to x that survive the preceding rounds of the sieve does not necessarily contain the complete set of integers for which P is the least prime factor. Likewise, the subset of representations containing all the integers up to x that survive the preceding rounds of the sieve does not necessarily contain the complete set of primes in the interval $[\sqrt{x}, x]$. Therefore, one does not have a set that one can define quantitatively at each round of the sieve, and therefore, one cannot determine an upper bound to the proportion of representations that are sifted out at each round of the sieve, in the same way that could be done with the sieving of integers, as described above. Nevertheless, as we demonstrate in this paper, by relating the sieving of representations to the sieving of integers for the same x, one can determine an upper bound to the proportion of additive representations that are sifted out at each round of the sieve, and an upper bound to the proportion of subtractive representations that are sifted out at each round of the sieve, which we do on the condition that x is sufficiently large. Following on from the upper bounds, we then determine a positive lower bound to the quantity of additive representations that survive the sieve, where both p and q are primes in the interval $[\sqrt{x}, x]$, and a positive lower bound to the quantity of subtractive representations that survive the sieve, where both p and q are primes in the interval $[\sqrt{x}, x]$.

Therefore, we prove the following:

Theorem 1. *Every sufficiently large even number is the sum of two primes.*

Theorem 2. *Every even number is the difference of two primes in infinitely many ways.*

Theorem 1 partially addresses question (i) as posed above, and, to date is the closest theoretical result to the binary Goldbach conjecture. Theorem 2 fully addresses questions (ii)-(iii) as posed above, fully establishes the twin prime conjecture, and, to date is the closest theoretical result to the Polignac conjecture.

2. Notation and preliminaries

$\mathbb{N}$ and $\mathbb{P}$: the set of natural numbers and the set of prime numbers respectively.
P_m and P_n : the m^{th} and the n^{th} prime numbers respectively. We have $P_{m+1} \leq P_n$.
Q_z : the z^{th} integer in the sequence $Q := \{a \in \mathbb{N} : 2 \nmid a, 3 \nmid a\}$.
x and $2k$: even numbers.
p : prime number.
w : a sufficiently large positive integer, not necessarily the same at every occurrence.
$\alpha_1, \alpha_2, \alpha_n, v_m, v_n, t_m, t_n$: positive real numbers.
$\beta_1, \beta_2, \beta_n, \gamma_1, \gamma_2, \gamma_n$: non-negative real numbers.
$\psi := \{p \in \mathbb{P} : 2 \leq p < P_n\}$.
$\omega := \left\{p \in \mathbb{P} : P_n < p \leq \sqrt{x}\right\}$.
$\pi(a) := |\{p \in \mathbb{P} : 2 \leq p \leq a\}|$.
$\pi[a, b] := |\{p \in \mathbb{P} : a \leq p \leq b\}|$.

 doi:10.20944/preprints202110.0087.v8

$g(x) := |\{p \in \mathbb{P} : (x-p) \in \mathbb{P}, \sqrt{x} \leq p \leq (x-p)\}|$.
$\pi_{2k}(x) := |\{p \in \mathbb{P} : (p+2k) \in \mathbb{P}, \sqrt{x} \leq p \leq (p+2k) \leq x\}|$.
$\pi_{2k}(x)' := |\{p \in \mathbb{P} : (p+2k) \in \mathbb{P}, 2 \leq p \leq (p+2k) \leq x\}|$.

$\mathcal{A} := \{a \in \mathbb{N} : a \leq x\}$.
$\mathcal{A}^p := \{a \in \mathcal{A} : \text{for all } p \in \psi, p \nmid a\}$.
$\mathcal{A}_p^p := \{a \in \mathcal{A}^p : p \mid a\}$.
$\mathcal{A}_{P_mP_n}^p := \{a \in \mathcal{A}^p : P_m \mid a \text{ and } P_n \mid a\}$.
$\mathcal{A}_\omega^p := \{a \in \mathcal{A}^p : \text{for all } p \in \omega, p \mid a\}$.
$S\left(\mathcal{A}^p, p, \sqrt{x}\right) := \left\{a \in \mathcal{A}^p : p \nmid a, p \leq \sqrt{x}\right\}$

$\mathcal{B} := \{b \in \mathbb{N} : b \leq (x-b)\}$.
$\mathcal{B}^p := \{b \in \mathcal{B} : \text{for all } p \in \psi, p \nmid b \text{ and } p \nmid (x-b)\}$.
$\mathcal{B}_p^p := \{b \in \mathcal{B}^p : p \mid b \text{ or } p \mid (x-b)\}$.
$\mathcal{B}_{P_mP_n}^p := \left\{b \in \mathcal{B}^p : \left(P_m \mid b \text{ or } P_m \mid (x-b)\right) \text{ and } \left(P_n \mid b \text{ or } P_n \mid (x-b)\right)\right\}$.
$S\left(\mathcal{B}^p, p, \sqrt{x}\right) := \left\{b \in \mathcal{B}^p : p \nmid b \text{ and } p \nmid (x-b), p \leq \sqrt{x}\right\}$

$\mathcal{C} := \left\{c \in \mathbb{N} : (c+2k) \leq x, 2k \leq \frac{x}{2}\right\}$.
$\mathcal{C}^p := \{c \in \mathcal{C} : \text{for all } p \in \psi, p \nmid c \text{ and } p \nmid (c+2k)\}$.
$\mathcal{C}_p^p := \{c \in \mathcal{C}^p : p \mid c \text{ or } p \mid (c+2k)\}$.
$S\left(\mathcal{C}^p, p, \sqrt{x}\right) := \left\{c \in \mathcal{C}^p : p \nmid c \text{ and } p \nmid (c+2k), p \leq \sqrt{x}\right\}$.

$\mathcal{D} := \left[d \hat{} e \in \mathbb{N} : d \leq x,\ 2k \leq \frac{x}{2},\ e = \begin{cases} 2, & \text{if } (1+2k) \leq d \leq (x-2k), \\ 1, & \text{otherwise,} \end{cases}\right]$; the square brackets are to mean that $\mathcal{D}$ is a multiset, with e being the multiplicity of d.
$\mathcal{D}_p := [d \in \mathcal{D} : p \mid d]$.
Any other notation used, is defined at the point of use.

We take the following two propositions to be evidently true:

Proposition 1.

$$\begin{cases} \text{if } p \mid x \text{ and } p \mid b, & \text{then } p \mid (x-b), \\ \text{if } p \nmid x \text{ and } p \mid b, & \text{then } p \nmid (x-b), \\ \text{if } p \nmid x \text{ and } p \mid (x-b), & \text{then } p \nmid b. \end{cases}$$

Proposition 2.

$$\begin{cases} \text{if } p \mid 2k \text{ and } p \mid c, & \text{then } p \mid (c+2k), \\ \text{if } p \nmid 2k \text{ and } p \mid c, & \text{then } p \nmid (c+2k), \\ \text{if } p \nmid 2k \text{ and } p \mid (c+2k), & \text{then } p \nmid c. \end{cases}$$

3. Sieve outline

3.1. **Sieving integers.** Prior to the 1st round of the sieve, we have

$$|\mathcal{A}_{P_1}| = \frac{\alpha_1}{P_1} \cdot |\mathcal{A}|, \tag{3.1}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 August 2022 doi:10.20944/preprints202110.0087.v8

and therefore, at the 1st round of the sieve, we have

$$|S\left(\mathcal{A}, P_1\right)| = |\mathcal{A}| - \frac{\alpha_1}{P_1} \cdot |\mathcal{A}| = |\mathcal{A}| \cdot \frac{P_1 - \alpha_1}{P_1} = |\mathcal{A}^{P_2}|. \tag{3.2}$$

Since $|\mathcal{A}| = x$, then

$$|\mathcal{A}^{P_2}| = x \cdot \frac{P_1 - \alpha_1}{P_1}. \tag{3.3}$$

Likewise, prior to the n^{th} round of the sieve, we have

$$|\mathcal{A}_{P_n}^{P_n}| = \frac{\alpha_n}{P_n} \cdot |\mathcal{A}^{P_n}|, \tag{3.4}$$

and therefore, at the n^{th} round of the sieve, we have

$$|S\left(\mathcal{A}^{P_n}, P_n\right)| = |\mathcal{A}^{P_n}| - \frac{\alpha_n}{P_n} \cdot |\mathcal{A}^{P_n}| = |\mathcal{A}^{P_n}| \cdot \frac{P_n - \alpha_n}{P_n} = |\mathcal{A}^{P_{n+1}}|, \tag{3.5}$$

and therefore

$$|\mathcal{A}^{P_{n+1}}| = x \cdot \frac{P_1 - \alpha_1}{P_1} \cdot \frac{P_2 - \alpha_2}{P_2} \cdot \dots \cdot \frac{P_n - \alpha_n}{P_n}. \tag{3.6}$$

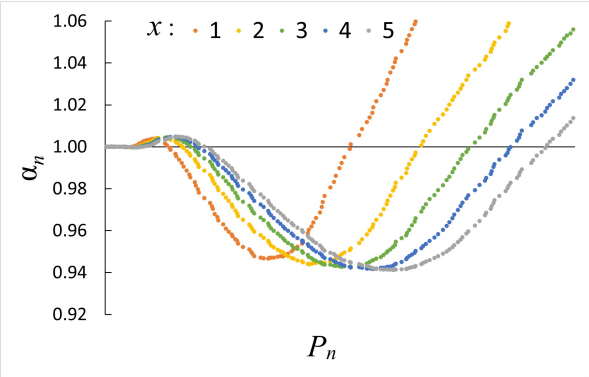
Since, upon completion of the sieve, we have

$$|S\left(\mathcal{A}^{P_n}, P_n, \sqrt{x}\right)| = \pi(x) - \pi\left(\sqrt{x}\right) + 1, \tag{3.7}$$

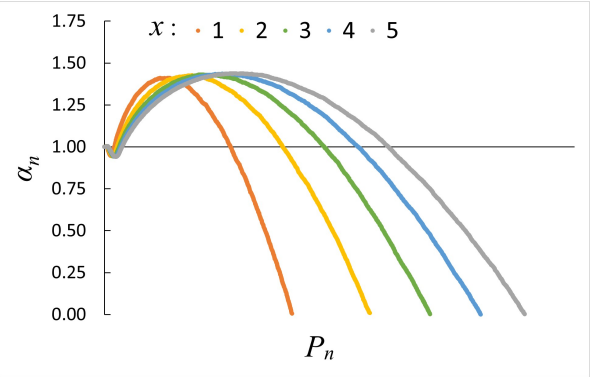
then

$$\pi(x) - \pi\left(\sqrt{x}\right) + 1 = x \cdot \prod_{P_n \le \sqrt{x}} \frac{P_n - \alpha_n}{P_n}. \tag{3.8}$$

Empirical data, as shown in Figure 1, suggest that α_n oscillates about 1, before tending to zero as $P_n \to \sqrt{x}$. In section 4, we determine an upper bound to α_n, for $P_n \le \sqrt{x}$ and $x \ge w$.



(a) All $P_n \le 1 \cdot 10^3$



(b) All $P_n \le \sqrt{x\,(\cdot 10^8)}$

FIGURE 1. $\alpha_n = \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot P_n$, for various $x\,\left(\cdot 10^8\right)$

 doi:10.20944/preprints202110.0087.v8

3.2. **Sieving additive representations.** Prior to the 1st round of the sieve, we have

$$|\mathcal{B}_{P_1}| = \frac{\beta_1}{P_1} \cdot |\mathcal{B}|, \tag{3.9}$$

and therefore, at the 1st round of the sieve, we have

$$|S\left(\mathcal{B}, P_1\right)| = |\mathcal{B}| - \frac{\beta_1}{P_1} \cdot |\mathcal{B}| = |\mathcal{B}| \cdot \frac{P_1 - \beta_1}{P_1} = |\mathcal{B}^{P_2}|. \tag{3.10}$$

Since $|\mathcal{B}| = \frac{x}{2}$, then

$$|\mathcal{B}^{P_2}| = \frac{x}{2} \cdot \frac{P_1 - \beta_1}{P_1}. \tag{3.11}$$

Likewise, prior to the n^{th} round of the sieve, we have

$$|\mathcal{B}_{P_n}^{P_n}| = \frac{\beta_n}{P_n} \cdot |\mathcal{B}^{P_n}|, \tag{3.12}$$

and therefore, at the n^{th} round of the sieve, we have

$$|S\left(\mathcal{B}^{P_n}, P_n\right)| = |\mathcal{B}^{P_n}| - \frac{\beta_n}{P_n} \cdot |\mathcal{B}^{P_n}| = |\mathcal{B}^{P_n}| \cdot \frac{P_n - \beta_n}{P_n} = |\mathcal{B}^{P_{n+1}}|, \tag{3.13}$$

and therefore

$$|\mathcal{B}^{P_{n+1}}| = \frac{x}{2} \cdot \frac{P_1 - \beta_1}{P_1} \cdot \frac{P_2 - \beta_2}{P_2} \cdot \ldots \cdot \frac{P_n - \beta_n}{P_n}. \tag{3.14}$$

Since, upon completion of the sieve, we have

$$|S\left(\mathcal{B}^{P_n}, P_n, \sqrt{x}\right)| = g\,(x) + u, \tag{3.15}$$

where $u = 1$ or 0, then

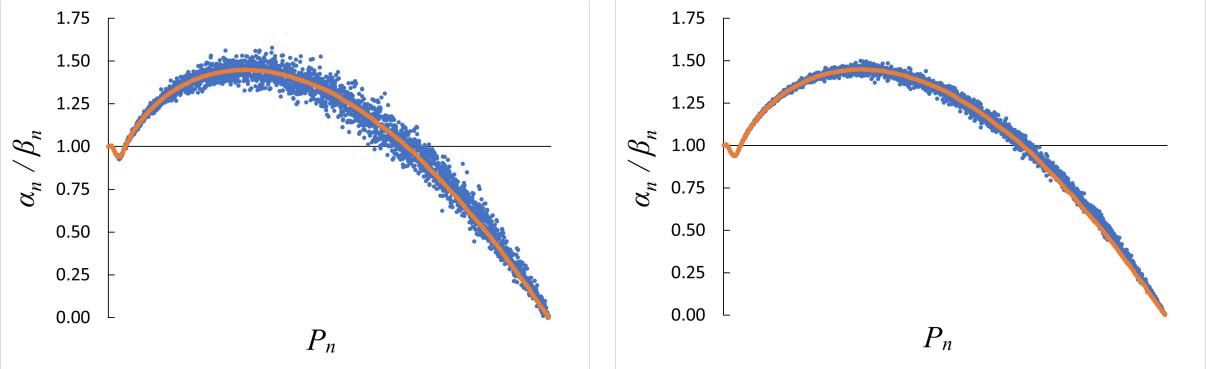
$$g\,(x) \geq \frac{x}{2} \cdot \prod_{P_n \leq \sqrt{x}} \frac{P_n - \beta_n}{P_n} - 1. \tag{3.16}$$

Empirical data for the sieving of additive representations, as shown in Figure 2 for all $P_n \leq \sqrt{x}$, suggest that

$$\frac{|\mathcal{B}_{P_n}^{P_n}|}{|\mathcal{B}^{P_n}|} \approx \begin{cases} 1 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|}, & \text{if } P_n \mid x, \\ 2 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|}, & \text{if } P_n \nmid x, \end{cases} \tag{3.17}$$

which, in section 5 we generalise for $P_n \leq \sqrt{x}$. Through (3.17), we are able to determine an upper bound to β_n, which we do in section 5 for $P_n \leq \sqrt{x}$ and $x \geq w$; which then allows us to determine a positive lower bound to (3.16) for $x \geq w$, which we do in section 7, and thereby we prove theorem 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 August 2022 doi:10.20944/preprints202110.0087.v8



(a) $x = 892,371,464$ (b) $x = 892,371,480$

FIGURE 2. $\alpha_n = \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot P_n$ (red) and $\beta_n = \begin{cases} \dfrac{1}{1} \cdot \dfrac{|\mathcal{B}_{P_n}^{P_n}|}{|\mathcal{B}^{P_n}|} \cdot P_n & \text{if } P_n \mid x, \\ \dfrac{1}{2} \cdot \dfrac{|\mathcal{B}_{P_n}^{P_n}|}{|\mathcal{B}^{P_n}|} \cdot P_n & \text{if } P_n \nmid x, \end{cases}$ (blue).

3.3. **Sieving subtractive representations.** Prior to the 1st round of the sieve, we have

$$|C_{P_1}| = \frac{\gamma_1}{P_1} \cdot |C|, \tag{3.18}$$

and therefore, at the 1st round of the sieve, we have

$$|S\left(C, P_1\right)| = |C| - \frac{\gamma_1}{P_1} \cdot |C| = |C| \cdot \frac{P_1 - \gamma_1}{P_1} = |C^{P_2}|. \tag{3.19}$$

Since $|C| = x - 2k$, then

$$|C^{P_2}| = (x - 2k) \cdot \frac{P_1 - \gamma_1}{P_1}. \tag{3.20}$$

Likewise, prior to the n^{th} round of the sieve, we have

$$|C_{P_n}^{P_n}| = \frac{\gamma_n}{P_n} \cdot |C^{P_n}|, \tag{3.21}$$

and therefore, at the n^{th} round of the sieve, we have

$$|S\left(C^{P_n}, P_n\right)| = |C^{P_n}| - \frac{\gamma_n}{P_n} \cdot |C^{P_n}| = |C^{P_n}| \cdot \frac{P_n - \gamma_n}{P_n} = |C^{P_{n+1}}|, \tag{3.22}$$

and therefore

$$|C^{P_{n+1}}| = (x - 2k) \cdot \frac{P_1 - \gamma_1}{P_1} \cdot \frac{P_2 - \gamma_2}{P_2} \cdot \ldots \cdot \frac{P_n - \gamma_n}{P_n}. \tag{3.23}$$

Since, upon completion of the sieve, we have

$$|S\left(C^{P_n}, P_n, \sqrt{x}\right)| = \pi_{2k}(x) + u, \tag{3.24}$$

where $u = 1$ or 0, then

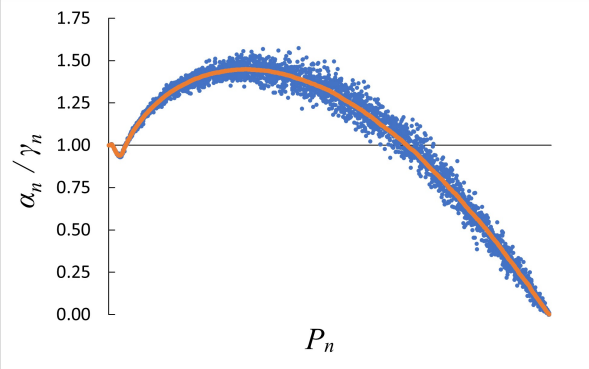
$$\pi_{2k}(x) \geq (x - 2k) \cdot \prod_{P_n \leq \sqrt{x}} \frac{P_n - \gamma_n}{P_n} - 1. \tag{3.25}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 August 2022 doi:10.20944/preprints202110.0087.v8

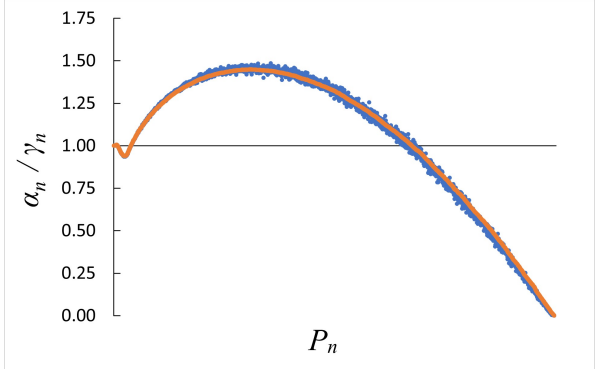
As with the sieving of additive representations, empirical data for the sieving of subtractive representations, as shown in Figure 3 for all $P_n \leq \sqrt{x}$, suggest that

$$\frac{|C_{P_n}^{P_n}|}{|C^{P_n}|} \approx \begin{cases} 1 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|}, & \text{if } P_n \mid 2k, \\ 2 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|}, & \text{if } P_n \nmid 2k, \end{cases} \tag{3.26}$$

which, in section 6, we generalise for $P_n \leq \sqrt{x}$. Through (3.26), we are able to determine an upper bound to γ_n, which we do in section 6 for $P_n \leq \sqrt{x}$ and $x \geq w$; which then allows us to determine a positive lower bound to (3.25) for $x \geq w$, which we do in section 7, and which then implies that $\pi_{2k}(x)' \to \infty$ as $x \to \infty$, and thereby we prove theorem 2.



(a) $x = 892,371,464$ and $2k = 446,185,732$



(b) $x = 892,371,464$ and $2k = 9,699,690$

FIGURE 3. $\alpha_n = \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot P_n$ (red) and $\gamma_n = \begin{cases} \dfrac{1}{1} \cdot \dfrac{|C_{P_n}^{P_n}|}{|C^{P_n}|} \cdot P_n & \text{if } P_n \mid 2k, \\ \dfrac{1}{2} \cdot \dfrac{|C_{P_n}^{P_n}|}{|C^{P_n}|} \cdot P_n & \text{if } P_n \nmid 2k, \end{cases}$ (blue).

4. An upper bound on the sifted out integers

In this section, we determine an upper bound to α_n, as follows.

Prior to the n^{th} round of the sieve, we have

$$\frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} = \frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}_{P_n}^{P_n}| + |\mathcal{A}_{\omega}^{P_n}| + \pi\left[P_n, x\right] + 1} \leq \frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}_{P_n}^{P_n}| + \pi\left[P_n, x\right]}. \tag{4.1}$$

Since $\mathcal{A}_{P_n}^{P_n}$ consists of the complete set of integers ($\leq x$) for which P_n is the least prime factor, then due to the Fundamental Theorem of Arithmetic, we have

$$|\mathcal{A}_{P_n}^{P_n}| = \pi\left[P_n, \frac{x}{P_n}\right] + O, \tag{4.2}$$

 doi:10.20944/preprints202110.0087.v8

where O is a non-negative integer. Substituting (4.2) into the right-hand side of (4.1), we have

$$\frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \leq \frac{\pi\left[P_n, \frac{x}{P_n}\right] + O}{\pi\left[P_n, \frac{x}{P_n}\right] + O + \pi\left[P_n, x\right]} \leq \frac{\pi\left[P_n, \frac{x}{P_n}\right]}{\pi\left[P_n, x\right]}. \tag{4.3}$$

Since

$$\frac{\pi\left[P_n, \frac{x}{P_n}\right]}{\pi\left[P_n, x\right]} \leq \frac{\pi\left[P_n, \frac{x}{P_n}\right] + \pi\left(P_{n-1}\right)}{\pi\left[P_n, x\right] + \pi\left(P_{n-1}\right)} = \frac{\pi\left(\frac{x}{P_n}\right)}{\pi\left(x\right)}, \tag{4.4}$$

then

$$\frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \leq \frac{\pi\left(\frac{x}{P_n}\right)}{\pi\left(x\right)}, \tag{4.5}$$

which we quantify as follows.

Let $f(a) := \dfrac{\pi\left(a\right)}{a}$, then

$$\frac{\pi\left(\frac{x}{P_n}\right)}{\pi\left(x\right)} = \frac{1}{P_n} \cdot \frac{f\left(\frac{x}{P_n}\right)}{f(x)}. \tag{4.6}$$

Since $P_n \leq \sqrt{x}$, then

$$\frac{f\left(\frac{x}{P_n}\right)}{f(x)} \leq \frac{f\left(\sqrt{x}\right)}{f(x)}. \tag{4.7}$$

Due to the Prime Number Theorem, we have

$$\lim_{x\to\infty} \frac{f\left(\sqrt{x}\right)}{f(x)} = 2, \tag{4.8}$$

and therefore, if $x \geq w$, then

$$\frac{f\left(\sqrt{x}\right)}{f(x)} \leq 2.15. \tag{4.9}$$

Therefore, if $P_n \leq \sqrt{x}$ and $x \geq w$, then due to (4.5), (4.6), and (4.9), we have

$$\frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \leq \frac{1}{P_n} \cdot 2.15. \tag{4.10}$$

5. Upper bounds on the sifted out additive representations

In this section, we determine upper bounds to β_n, as follows.

Prior to the 1st round of the sieve, we have

$$|\mathcal{B}| = \frac{1}{2} \cdot |\mathcal{A}|, \tag{5.1}$$

 doi:10.20944/preprints202110.0087.v8

and, due to Proposition 1, we have

$$|\mathcal{B}_{P_n}| = \begin{cases} \frac{1}{2} \cdot |\mathcal{A}_{P_n}|, & \text{if } P_n \mid x, \\ \frac{1}{1} \cdot |\mathcal{A}_{P_n}|, & \text{if } P_n \nmid x, \end{cases} \tag{5.2}$$

and therefore, we have

$$\frac{|\mathcal{B}_{P_n}|}{|\mathcal{B}|} = \begin{cases} 1 \cdot \frac{|\mathcal{A}_{P_n}|}{|\mathcal{A}|}, & \text{if } P_n \mid x, \\ 2 \cdot \frac{|\mathcal{A}_{P_n}|}{|\mathcal{A}|}, & \text{if } P_n \nmid x. \end{cases} \tag{5.3}$$

In the following, we describe the sieving process of additive representations of the set $\mathcal{B}$ at the m^{th} round of the sieve, in comparison to the sieving process of integers of the set $\mathcal{A}$, also at the m^{th} round of the sieve. The m^{th} round of the sieve precedes the n^{th} round of the sieve.

When sieving integers of the set $\mathcal{A}$, prior to the m^{th} round of the sieve, we have

$$|\mathcal{A}_{P_m}^{P_m}| = \frac{\nu_m}{P_m} \cdot |\mathcal{A}^{P_m}|, \tag{5.4}$$

and

$$|\mathcal{A}_{P_n}^{P_m}| = \frac{\nu_n}{P_n} \cdot |\mathcal{A}^{P_m}|. \tag{5.5}$$

Due to (5.4) and (5.5), we have

$$|\mathcal{A}_{P_mP_n}^{P_m}| \approx \frac{\nu_m}{P_m} \cdot \frac{\nu_n}{P_n} \cdot |\mathcal{A}^{P_m}|, \tag{5.6}$$

and therefore

$$|\mathcal{A}_{P_mP_n}^{P_m}| \approx \frac{\nu_n}{P_n} \cdot |\mathcal{A}_{P_m}^{P_m}| \approx \frac{\nu_m}{P_m} \cdot |\mathcal{A}_{P_n}^{P_m}|, \tag{5.7}$$

where the " $\approx$ " sign arises due to the uneven local distribution of primes amongst the natural numbers. Thus, at the m^{th} round of the sieve, due to (5.4) we have

$$|S\left(\mathcal{A}^{P_m}, P_m\right)| = |\mathcal{A}^{P_m}| \cdot \left(\frac{P_m - \nu_m}{P_m}\right) = |\mathcal{A}^{P_{m+1}}|, \tag{5.8}$$

and due to (5.7), we have

$$|S\left(\mathcal{A}_{P_n}^{P_m}, P_m\right)| \approx |\mathcal{A}_{P_n}^{P_m}| \cdot \left(\frac{P_m - \nu_m}{P_m}\right) = |\mathcal{A}_{P_n}^{P_{m+1}}|. \tag{5.9}$$

When sieving additive representations of the set $\mathcal{B}$, prior to the m^{th} round of the sieve, we have

$$|\mathcal{B}_{P_m}^{P_m}| = \begin{cases} \frac{t_m}{P_m} \cdot |\mathcal{B}^{P_m}|, & \text{if } P_m \mid x, \\ \frac{2t_m}{P_m} \cdot |\mathcal{B}^{P_m}|, & \text{if } P_m \nmid x, \end{cases} \tag{5.10}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 August 2022 doi:10.20944/preprints202110.0087.v8

and

$$|\mathcal{B}_{P_n}^{P_m}| = \begin{cases} \frac{t_n}{P_n} \cdot |\mathcal{B}^{P_m}|, & \text{if } P_n \mid x, \\ \frac{2t_n}{P_n} \cdot |\mathcal{B}^{P_m}|, & \text{if } P_n \nmid x, \end{cases} \tag{5.11}$$

where, the two-case scenarios arise due to Proposition 1. Due to (5.10) and (5.11), we have

$$|\mathcal{B}_{P_mP_n}^{P_m}| \approx \begin{cases} \frac{t_m}{P_m} \cdot \frac{t_n}{P_n} \cdot |\mathcal{B}^{P_m}|, & \text{if } P_m \mid x \text{ and } P_n \mid x, \\ \frac{t_m}{P_m} \cdot \frac{2t_n}{P_n} \cdot |\mathcal{B}^{P_m}|, & \text{if } P_m \mid x \text{ and } P_n \nmid x, \\ \frac{2t_m}{P_m} \cdot \frac{t_n}{P_n} \cdot |\mathcal{B}^{P_m}|, & \text{if } P_m \nmid x \text{ and } P_n \mid x, \\ \frac{2t_m}{P_m} \cdot \frac{2t_n}{P_n} \cdot |\mathcal{B}^{P_m}|, & \text{if } P_m \nmid x \text{ and } P_n \nmid x, \end{cases} \tag{5.12}$$

and therefore

$$|\mathcal{B}_{P_mP_n}^{P_m}| \approx \begin{cases} \frac{t_n}{P_n} \cdot |\mathcal{B}_{P_m}^{P_m}| \approx \frac{t_m}{P_m} \cdot |\mathcal{B}_{P_n}^{P_m}|, & \text{if } P_m \mid x \text{ and } P_n \mid x, \\ \frac{2t_n}{P_n} \cdot |\mathcal{B}_{P_m}^{P_m}| \approx \frac{t_m}{P_m} \cdot |\mathcal{B}_{P_n}^{P_m}|, & \text{if } P_m \mid x \text{ and } P_n \nmid x, \\ \frac{t_n}{P_n} \cdot |\mathcal{B}_{P_m}^{P_m}| \approx \frac{2t_m}{P_m} \cdot |\mathcal{B}_{P_n}^{P_m}|, & \text{if } P_m \nmid x \text{ and } P_n \mid x, \\ \frac{2t_n}{P_n} \cdot |\mathcal{B}_{P_m}^{P_m}| \approx \frac{2t_m}{P_m} \cdot |\mathcal{B}_{P_n}^{P_m}|, & \text{if } P_m \nmid x \text{ and } P_n \nmid x. \end{cases} \tag{5.13}$$

Thus, at the m^{th} round of the sieve, due to (5.10) we have

$$|S\left(\mathcal{B}^{P_m}, P_m\right)| = \begin{cases} |\mathcal{B}^{P_m}| \cdot \left(\frac{P_m - t_m}{P_m}\right) = |\mathcal{B}^{P_{m+1}}|, & \text{if } P_m \mid x, \\ |\mathcal{B}^{P_m}| \cdot \left(\frac{P_m - 2t_m}{P_m}\right) = |\mathcal{B}^{P_{m+1}}|, & \text{if } P_m \nmid x, \end{cases} \tag{5.14}$$

and due to (5.13), we have

$$|S\left(\mathcal{B}_{P_n}^{P_m}, P_m\right)| \approx \begin{cases} |\mathcal{B}_{P_n}^{P_m}| \cdot \left(\frac{P_m - t_m}{P_m}\right) = |\mathcal{B}_{P_n}^{P_{m+1}}|, & \text{if } P_m \mid x, \\ |\mathcal{B}_{P_n}^{P_m}| \cdot \left(\frac{P_m - 2t_m}{P_m}\right) = |\mathcal{B}_{P_n}^{P_{m+1}}|, & \text{if } P_m \nmid x. \end{cases} \tag{5.15}$$

Therefore, if prior to the m^{th} round of the sieve, we have

$$\frac{|\mathcal{B}_{P_n}^{P_m}|}{|\mathcal{B}^{P_m}|} \approx \frac{|\mathcal{A}_{P_n}^{P_m}|}{|\mathcal{A}^{P_m}|}, \tag{5.16}$$

then, due to (5.8), (5.9), (5.14), and (5.15), post the m^{th} round of the sieve, we have

$$\frac{|\mathcal{B}_{P_n}^{P_{m+1}}|}{|\mathcal{B}^{P_{m+1}}|} \approx \frac{|\mathcal{A}_{P_n}^{P_{m+1}}|}{|\mathcal{A}^{P_{m+1}}|}. \tag{5.17}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 August 2022 doi:10.20944/preprints202110.0087.v8

Therefore, due to (5.3) we have

$$\frac{|\mathcal{B}_{P_n}^{P_n}|}{|\mathcal{B}^{P_n}|} \approx \begin{cases} 1 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|}, & \text{if } P_n \mid x, \\ 2 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|}, & \text{if } P_n \nmid x, \end{cases} \tag{5.18}$$

and thus, we have generalised (3.17).

Due to (5.18), we have

$$\frac{|\mathcal{B}_{P_n}^{P_n}|}{|\mathcal{B}^{P_n}|} = \begin{cases} 1 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot h_n, & \text{if } P_n \mid x, \\ 2 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot h_n, & \text{if } P_n \nmid x, \end{cases} \tag{5.19}$$

where h_n is a positive real number. Since h_n is an effect of the local distribution of primes amongst the natural numbers, then $h_n \to 1$ as $|\mathcal{B}_{P_n}^{P_n}| \to \infty$. However, since $|\mathcal{B}_{P_n}^{P_n}| \to 0$ as $P_n \to \sqrt{x}$, then h_n can be significantly larger than 1. Therefore, it is necessary to determine an upper bound to (5.19) for $P_n \leq \sqrt{x}$, which we do as follows.

Let

$$\frac{|\mathcal{B}_{P_n}^{P_n}|}{|\mathcal{B}^{P_n}|} \cdot P_n = \beta_n > 0. \tag{5.20}$$

Let $x \to \infty$ and let $|\mathcal{B}_{P_n}^{P_n}|$ remain constant, then $|\mathcal{B}| \to \infty$ and therefore $|\mathcal{B}_{P_n}^{P_n}| \to 0$. Since $P_n \to \sqrt{x}$ as $|\mathcal{B}_{P_n}^{P_n}| \to 0$ and $\beta_n \to 0$ as $P_n \to \sqrt{x}$, then $\beta_n \to 0$ as $|\mathcal{B}_{P_n}^{P_n}| \to 0$. Therefore, there exists $x \geq w$, such that $|\mathcal{B}_{P_n}^{P_n}| \geq w$, such that $h_n \leq 1.15$, and such that $\beta_n \leq 2.15$, where 2.15 is the upper bound of α_n, as determined in section 4. Therefore, if $P_n \leq \sqrt{x}$ and $x \geq w$, then

$$\frac{|\mathcal{B}_{P_n}^{P_n}|}{|\mathcal{B}^{P_n}|} \leq \begin{cases} 1 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot 1.15, & \text{if } P_n \mid x, \\ 2 \cdot \dfrac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot 1.15, & \text{if } P_n \nmid x, \end{cases} \tag{5.21}$$

and therefore

$$\frac{|\mathcal{B}_{P_n}^{P_n}|}{|\mathcal{B}^{P_n}|} \leq \frac{5}{P_n}. \tag{5.22}$$

Furthermore, empirical data show that

$$\frac{|\mathcal{B}_{P_1}|}{|\mathcal{B}|} \leq \frac{1}{P_1}, \tag{5.23}$$

and

$$\frac{|\mathcal{B}_{P_2}^{P_2}|}{|\mathcal{B}^{P_2}|} \leq \frac{2}{P_2} + 2, \tag{5.24}$$

and

$$\frac{|\mathcal{A}_{P_3}^{P_3}|}{|\mathcal{A}^{P_3}|} \leq \frac{1}{P_3} + 1. \tag{5.25}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 August 2022 doi:10.20944/preprints202110.0087.v8

Due to (5.21) and (5.25), we have

$$\frac{|\mathcal{B}_{P_3}^{P_3}|}{|\mathcal{B}^{P_3}|} \leq 2 \cdot \left(\frac{1}{P_3} + 1\right) \cdot 1.15 \leq \frac{3}{P_3} + 3. \tag{5.26}$$

6. Upper bounds on the sifted out subtractive representations

In this section, we determine upper bounds to γ_n, as follows.

Prior to the 1st round of the sieve, we have

$$|C| = \frac{1}{2} \cdot |\mathcal{D}|, \tag{6.1}$$

and, due to Proposition 2, we have

$$|C_{P_n}| = \begin{cases} \frac{1}{2} \cdot |\mathcal{D}_{P_n}|, & \text{if } P_n \mid 2k, \\ \frac{1}{1} \cdot |\mathcal{D}_{P_n}|, & \text{if } P_n \nmid 2k, \end{cases} \tag{6.2}$$

and therefore, we have

$$\frac{|C_{P_n}|}{|C|} = \begin{cases} 1 \cdot \frac{|\mathcal{D}_{P_n}|}{|\mathcal{D}|}, & \text{if } P_n \mid 2k, \\ 2 \cdot \frac{|\mathcal{D}_{P_n}|}{|\mathcal{D}|}, & \text{if } P_n \nmid 2k. \end{cases} \tag{6.3}$$

Since

$$\frac{|\mathcal{D}_{P_n}|}{|\mathcal{D}|} = \frac{|\mathcal{A}_{P_n}|}{|\mathcal{A}|}, \tag{6.4}$$

then

$$\frac{|C_{P_n}|}{|C|} = \begin{cases} 1 \cdot \frac{|\mathcal{A}_{P_n}|}{|\mathcal{A}|}, & \text{if } P_n \mid 2k, \\ 2 \cdot \frac{|\mathcal{A}_{P_n}|}{|\mathcal{A}|}, & \text{if } P_n \nmid 2k. \end{cases} \tag{6.5}$$

By the same procedure as in (5.4) through (5.18), we can find that due to (6.5) we have

$$\frac{|C_{P_n}^{P_n}|}{|C^{P_n}|} \approx \begin{cases} 1 \cdot \frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|}, & \text{if } P_n \mid 2k, \\ 2 \cdot \frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|}, & \text{if } P_n \nmid 2k, \end{cases} \tag{6.6}$$

and thus, we have generalised (3.26). If $P_n \leq \sqrt{x}$ and $x \geq w$, then by the same procedure as in (5.19) through (5.21), we can find that due to (6.6) we have

$$\frac{|C_{P_n}^{P_n}|}{|C^{P_n}|} \leq \begin{cases} 1 \cdot \frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot 1.15, & \text{if } P_n \mid 2k, \\ 2 \cdot \frac{|\mathcal{A}_{P_n}^{P_n}|}{|\mathcal{A}^{P_n}|} \cdot 1.15, & \text{if } P_n \nmid 2k, \end{cases} \tag{6.7}$$

 doi:10.20944/preprints202110.0087.v8

and therefore

$$\frac{|C_{P_n}^{P_n}|}{|C^{P_n}|} \leq \frac{5}{P_n}. \tag{6.8}$$

Furthermore, empirical data show that

$$\frac{|C_{P_1}^{P_1}|}{|C^{P_1}|} \leq \frac{1}{P_1}, \tag{6.9}$$

and

$$\frac{|C_{P_2}^{P_2}|}{|C^{P_2}|} \leq \frac{2}{P_2} + 2, \tag{6.10}$$

and

$$\frac{|\mathcal{A}_{P_3}^{P_3}|}{|\mathcal{A}^{P_3}|} \leq \frac{1}{P_3} + 1. \tag{6.11}$$

Due to (6.7) and (6.11), we have

$$\frac{|C_{P_3}^{P_3}|}{|C^{P_3}|} \leq 2 \cdot \left(\frac{1}{P_3} + 1\right) \cdot 1.15 \leq \frac{3}{P_3} + 3. \tag{6.12}$$

7. Lower bounds on the surviving representations

Let $x \geq w$, then due to the upper bounds for β_n in (5.22), (5.23), (5.24), and (5.26), we have the following lower bound for (3.16):

$$g(x) \geq \left(\frac{x}{2} \cdot \frac{(P_1 - 1)}{P_1} \cdot \frac{(P_2 - 2)}{P_2} \cdot \frac{(P_3 - 3)}{P_3} \cdot \prod_{P_n = P_4}^{P_n \leq \sqrt{x}} \frac{(P_n - 5)}{P_n}\right) - 6, \tag{7.1}$$

which we quantify as follows. Since

$$\prod_{P_n = P_4}^{P_n \leq \sqrt{x}} \frac{(P_n - 5)}{P_n} \geq \prod_{P_n = P_4}^{P_n = P_{70}} \frac{(P_n - 5)}{P_n} \cdot \prod_{P_n = P_{71}}^{P_n \leq \sqrt{x}} \frac{(P_n - 6)}{P_n}, \tag{7.2}$$

and

$$\frac{(P_1 - 1)}{P_1} \cdot \frac{(P_2 - 2)}{P_2} \cdot \frac{(P_3 - 3)}{P_3} \cdot \prod_{P_n = P_4}^{P_n = P_{70}} \frac{(P_n - 5)}{P_n} \geq \frac{79}{500,000}, \tag{7.3}$$

then

$$g(x) \geq \left(x \cdot \frac{79}{1,000,000} \cdot \prod_{P_n = P_{71}}^{P_n \leq \sqrt{x}} \frac{(P_n - 6)}{P_n}\right) - 6. \tag{7.4}$$

Since $P_{71} = Q_{118}$, then

$$\prod_{P_n = P_{71}}^{P_n \leq \sqrt{x}} \frac{(P_n - 6)}{P_n} \geq \prod_{Q_z = Q_{118}}^{Q_z \leq \sqrt{x}} \frac{(Q_z - 6)}{Q_z}, \tag{7.5}$$

and therefore

$$g(x) \geq \left(x \cdot \frac{79}{1,000,000} \cdot \prod_{Q_z = Q_{118}}^{Q_z \leq \sqrt{x}} \frac{(Q_z - 6)}{Q_z}\right) - 6. \tag{7.6}$$

 doi:10.20944/preprints202110.0087.v8

Since

$$\prod_{Q_z=O_{118}}^{Q_z\le\sqrt{x}} \frac{(Q_z-6)}{Q_z} \ge \frac{Q_{116}\cdot Q_{117}}{x}, \tag{7.7}$$

where $Q_{116} = 347$ and $Q_{117} = 349$, then

$$g(x) \ge \left(\frac{79\cdot 347\cdot 349}{1,000,000}\right) - 6 \ge 1. \tag{7.8}$$

This completes the proof for theorem 1.

Let $x \ge w$ and let $2k \le \frac{x}{2}$, then due to the upper bounds for γ_n in (6.8), (6.9), (6.10), and (6.12), we have the following lower bound for (3.25):

$$\pi_{2k}(x) \ge \left(\frac{x}{2}\cdot\frac{(P_1-1)}{P_1}\cdot\frac{(P_2-2)}{P_2}\cdot\frac{(P_3-3)}{P_3}\cdot\prod_{P_n=P_4}^{P_n\le\sqrt{x}}\frac{(P_n-5)}{P_n}\right) - 6, \tag{7.9}$$

and therefore, as with (7.8), we have

$$\pi_{2k}(x) \ge \left(\frac{79\cdot 347\cdot 349}{1,000,000}\right) - 6 \ge 1. \tag{7.10}$$

Since the elements of the set $\pi_{2k}(x)$ consist of primes in the interval $\left[\sqrt{x}, x\right]$, then $\pi_{2k}(x)' \to \infty$ as $x \to \infty$. This completes the proof for theorem 2.

Acknowledgements

In constructing the proofs presented in this paper, we were constantly guided by observations of empirical data, such as those presented in section 3. To this end, we give very special thanks to our dear friends, Abdul Karim and Zhuen Xie, with the assistance of whom we were able to obtain the data in a much shorter time frame than we would have otherwise. We're also very grateful for the free access to the following websites, which we've made extensive use of: `oeis.org/` and `onlinenumbertools.com/calculate-prime-factors`.

References

[1] L. E. Dickson. Goldbach's Empirical Theorem: Every Integer is a Sum of Two Primes, in *History of the Theory of Numbers, Vol. 1: Divisibility and Primality*. Dover Publications, New York, 1952, pp. 421-424.

[2] T. O. Silva, S. Herzog, and S. Pard. Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4\cdot 10^{18}$. *Math. Comp.* **83** (2014), 2033–2060.

[3] V. Brun. The Sieve of Eratosthenes and the Theorem of Goldbach, in *Goldbach Conjecture*. World Scientific, Singapore, 2002, pp. 99–136.

[4] H. L. Montgomery and R. C. Vaughan. The exceptional set in Goldbach's problem. in *Acta. Arith.* **27** (1975), 353–370.

[5] O. Ramare. On Šnirel'man's constant. *Ann. Scuola. Norm. Sup. Pis.* **22** (1995), 645–706.

[6] J. R. Chen. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica.* **16** (1973), 157–176.

[7] I. M. Vinogradov. Representations of an odd integer as a sum of three primes. in *Goldbach Conjecture*. World Scientific, Singapore, 2002, pp. 61–64.

[8] T. Tao. Every odd number greater than 1 is the sum of at most five primes. *Preprint*, `arxiv.org/abs/1201.6656`, July 2012.

[9] H. Helfgott. The ternary Goldbach conjecture is true. *Preprint*, `arxiv.org/abs/1312.7748`, December 2013.

[10] L. E. Dickson. Theorems Analogous to Goldbach's. in *History of the Theory of Numbers, Vol. 1: Divisibility and Primality*. Dover Publications, New York, 1952, pp. 424-425.

[11] H. Rezgui. Conjecture of twin primes (Still unsolved problem in Number Theory). An expository essay. *Sur. Math. App.* **12** (2017), 229–252.
[12] L. Bethune. Found: New World Record Twin Primes!, 2016: available at: `https://www.epcc.ed.ac.uk/blog/2016/09/27/found-new-world-record-twin-prime`
[13] D. A. Goldston, J. Pintz, and C. Y. Yıldırım. Primes in tuples. I. *Ann. of Math.* **170**, No. 2, (2009), 819–862.
[14] Y. Zhang. Bounded gaps between primes. *Ann. Math.* **179** (2014), 1121–1174.
[15] J. Maynard. Small gaps between primes. *Ann. Math.* **181** (2015), 1–31.
[16] D. H. J. Polymath. The "bounded gaps between primes" Polymath project - a retrospective. *Preprint*, `arxiv.org/abs/1409.8361`, September 2014.
[17] S. Horsley. The Sieve of Eratosthenes. Being an account of his method of finding all the Prime Numbers. *Phil. Trans.* **62** (1772), 327–347.

About authors

Ali Shehu has a background in Chemistry / Materials Science from London Metropolitan University, UK, and University College London, UK. He was awarded the PhD degree in Chemistry, from Queen Mary University of London, UK, in April 2018.

Jetmira Uka has a background in Chemistry / Chemical Engineering from the University of Tirana, Albania, and she is currently a Physics Teacher at Herschel Grammar School, UK. She is also a PhD student in Mechanical and Aerospace Engineering, at Brunel University, UK.