Investigating the Consecutive Bias of the Products of $k$ Primes

Jacob William Ferrier

Advisor: Prof. Arindam Roy
Reader: Profs. William Brian and Xingjie Li

Spring 2021

A thesis submitted in partial fulfilment of the requirements for the Mathematics Honors Program.
Acknowledgements

I would like to gratefully acknowledge my advisor Professor Arindam Roy for his guidance, support, wisdom and dedication. I would also like to thank my readers, Professor William Brian and Professor Xingjie Li.
Investigating the Consecutive Bias of the Products of $k$ Primes
Jacob William Ferrier

Abstract

In this thesis, I examine the consecutive bias of the products of $k$ primes. The Chebyshev bias of primes, quasiprimes and products of $k$ primes is well studied. Furthermore, the consecutive bias of primes is well studied too, yet not in the case of the quasiprimes or more generally, the products of $k$ primes. In this thesis, we make three main conjectures that support our data. We also develop an algorithm for generating/verifying products of $k$ primes. We use multiple ceilings and values of $k$ as well as modulate our results by multiple modulos. Our thesis expands upon the work of Lemke Oliver and Soundararajan in their breakthrough paper on the consecutive bias. However, we developed conjectures that are distinct from theirs.
1 Introduction

1.1 Preliminaries

An integer \( p \) whose only divisors are \( \pm p \) and \( \pm 1 \) is said to be a member of the set of prime numbers (and each member \( p \) is called prime), a collection of integers central to the field of mathematics. The primes have been extensively studied for thousands of years and as a result, much is know about them. One such concept of interest for mathematicians is the distribution of the primes; that is, where do the primes lie within the integers in relation to each other? Understanding the distribution of the primes in fundamentally important to understanding the nature of the primes themselves. Much is known already about this prime distribution such as the infinitude of the primes, as proven by Euclid. Another result of great importance is Dirichlet’s theorem on the infinitude of the primes in the arithmetic expression

\[
\begin{align*}
a, a + q, a + 2q, & \ldots
\end{align*}
\]

where \( a, q \) are relatively prime, positive integers.

The prime numbers and their distribution largely forms the foundation of the field of number theory. Despite our wealth of knowledge regarding the primes, we know scarcely little compared to what there is to know about them. Such is why much of our knowledge concerning the primes lies on the fringe of speculation or conjecture opposed to in the realm of rigorous proof. The most famous unsolved question in mathematics is the Riemann Hypothesis, which itself is concerned with the distribution of the primes. Though still unsolved in the infinite case, the vast majority of the mathematics community agrees that it is “probably” true. Another unsolved problem regarding the primes is the Goldbach Conjecture, as well as Euler’s extension of it.[2] Though no definitive proof has been given for either, both are believed to be probably true as well.
For many of the unsolved problems regarding the primes and their distribution, efforts to work towards proofs has proven to be fruitful, yet much work remains. However, one piece of formally infeasible evidence that has proven to be immensely useful to the mathematics community over recent years does provide invaluable insight into understanding these problems: computers. The advent of computing technologies have revolutionized mathematics alongside the rest of the world. Computers offer unparalleled levels of precision, accuracy and speed and because of computers, “probably” true can be reinforced with massive volumes of numerical evidence. Alternatively, numerical evidence offers a new method for disproving conjecture. The Riemann Hypothesis could be proven to be false if a single, non-trivial zero to the Riemann zeta-function was found to be outside of the critical strip, or off the critical line, by Riemann’s further assertion. But also for the affirmative, by seeing if conjecture holds up at very large heights or with very large sets of numbers, one can be further sure of the fact that a proposition is “probably” true.

Another example of the usefulness of computers in solving difficult number theory problems is the case of the Mertens Conjecture, as presented in [2]. In 1897, Mertens conjectured that \( \sum_{k=1}^{n} \mu(k) < \sqrt{n} \) for \( n > 1 \), where \( \mu \) is the Möbius \( \mu \)-function. Mertens himself suggested that this was “very probable”. Many years later, in 1963, the Mertens Conjecture was computationally verified up to all \( n < 10 \) billion. However, it was in 1984 when Andrew Odlyzko and Herman te Riele demonstrated that the conjecture was in fact false, by use of a computer. Not only does this exemplify a case where a “very probably true” conjecture was disproved computationally, but also a case where computational verification appeared misleading, thus only reaffirming the need for continuous computational verification at very large heights.
1.2 Chebyshev’s Bias

Consider the function $\pi(x; q, a)$ which denotes the number of primes in the arithmetic progression (1) up to $x$ where $q$ is some fixed positive integer and $a$ is some positive integer relatively prime to $q$ (coprime, or GCD($a, q$) = 1). Note that neither $q$ nor $a$ need be prime themselves. As shown by [4], for all residues $a$ coprime to $q$,

$\pi(x; q, a) \sim \frac{x}{\phi(q) \log x}$

where $\phi(q)$ is the Euler phi function, i.e. the number of positive integers up to $q$ that are relatively prime to $q$. As shown by [8], more precisely, if the Generalized Riemann Hypothesis is true then

$\pi(x; q, a) = \text{li}(x) \phi(q) + O(x^{1/2+\epsilon})$

where $\text{li}(x)$ is the logarithmic integral; $\text{li}(x) = \int_{2}^{x} \frac{dt}{\log t}$. That is to say, regardless of the value of $a$, all $\pi$ functions of a fixed $q$ and $x$ should asymptotically similar.

This is where the bias comes into play. Discovered by Chebyshev in 1853 [3], $\pi(x; 4, 3) \geq \pi(x; 4, 1)$ for both large and small values of $x$. That is, there are actually more primes of the form $4n + 3$ than $4n + 1$ up to a given $x$ for some positive integer $n$. As shown by [7], $\pi(x; 4, 3) < \pi(x; 4, 1)$ for the first time at $x = 26,861$. Furthermore, consider the function shown by Ford and Sneed [5]

$\Delta(x; q, a, b) = \pi(x; q, a) - \pi(x; q, b)$

where $a, b$ are distinct residues that are both relatively prime to $q$. In the case of $\Delta(x; 4, 3, 1)$, the sign changes infinitely many times, as shown by [9], despite the seeming tendency for the sign to be negative as suggested by [5].
To exemplify this effect, consider the following table provided by [6].

<table>
<thead>
<tr>
<th>x</th>
<th>Number of primes $4n$ + 3 up to x</th>
<th>Number of primes $4n$ + 1 up to x</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>200</td>
<td>24</td>
<td>21</td>
</tr>
<tr>
<td>300</td>
<td>32</td>
<td>29</td>
</tr>
<tr>
<td>400</td>
<td>40</td>
<td>37</td>
</tr>
<tr>
<td>500</td>
<td>50</td>
<td>44</td>
</tr>
<tr>
<td>600</td>
<td>57</td>
<td>51</td>
</tr>
<tr>
<td>700</td>
<td>65</td>
<td>59</td>
</tr>
<tr>
<td>800</td>
<td>71</td>
<td>67</td>
</tr>
<tr>
<td>900</td>
<td>79</td>
<td>74</td>
</tr>
<tr>
<td>1000</td>
<td>87</td>
<td>80</td>
</tr>
<tr>
<td>10000</td>
<td>619</td>
<td>609</td>
</tr>
<tr>
<td>20000</td>
<td>1136</td>
<td>1125</td>
</tr>
<tr>
<td>50000</td>
<td>2583</td>
<td>2549</td>
</tr>
<tr>
<td>100000</td>
<td>4808</td>
<td>4783</td>
</tr>
</tbody>
</table>

### 1.3 Bias of Prime-Like Numbers

Chebyshev’s bias is well studied in the primes however, it is somewhat less so in other prime-like sequences. One example of prime-like numbers would be the quasiprimes (frequently referred to as quasi-primes, biprimes or semiprimes), where a quasiprime $q = p_1p_2$ where $p_1$ and $p_2$ are both primes and not necessarily distinct: thus, $p_1p_1$ is a valid quasiprime. It is worth noting that, due to to the infinitude of the primes, there are infinitely many quasiprimes. In the table below, provided by [5], the quasiprimes up to 100 are grouped into residue classes modulo 4.
\[
p_1 p_2 \equiv 1 \pmod{4} \\
p_1 p_2 \equiv 3 \pmod{4}
\]

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>15</td>
<td>21</td>
<td>35</td>
</tr>
<tr>
<td>25</td>
<td>39</td>
<td>33</td>
<td>51</td>
</tr>
<tr>
<td>49</td>
<td>55</td>
<td>57</td>
<td>87</td>
</tr>
<tr>
<td>65</td>
<td>91</td>
<td>69</td>
<td>95</td>
</tr>
<tr>
<td>77</td>
<td></td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The products of \( k \) primes are prime-like numbers that are similar to the quasiprimes. An integer \( n \) is said to be the product of \( k \) primes if \( \Omega(n) = k \). A quasiprime, for instance, can be defined as the product of 2 primes as for any quasiprime \( q \), \( \Omega(q) = 2 \). The presence of Chebyshev’s bias is less well-studied in the products of \( k \) primes than in even the quasiprimes, though it has been shown to exist by [10].

### 1.4 Consecutive Bias

The previous sections have demonstrated the existence of the Chebyshev bias for primes numbers as well as prime-like numbers. Within the last decade, a breakthrough occurred in mathematics as presented by [8]: not only does the Chebyshev bias numerically exist within the prime and prime-like numbers, but also in the consecutive primes. As demonstrated in Lemke Oliver and Soundararajan’s study, and using their notation and definitions, consider the consecutive sequences of residues class \( a_1 \pmod{q}, a_2 \pmod{q}, a_3 \pmod{q}, \ldots \) and let \( \vec{a} = (a_1, a_2, \ldots, a_r) \) denote the \( r \)-tuple of reduced residue classes \( \pmod{q} \) for some \( r \geq 1 \). Furthermore, define

\[
\pi(x; q, \vec{a}) = \# \{ p_n \leq x : p_{n+i-1} \equiv a_i \pmod{q} \text{ for each } 1 \leq i \leq r \} 
\]

where \( p_n \) is the sequence of primes in ascending order. This \( \pi(x; q, \vec{a}) \) function, defined by Lemke Oliver and Soundararajan, “counts the number of occurrences of the pattern \( \vec{a} \)
(mod $q$) among $r$ consecutive primes the least of which is below $x$.” Using this formula, Lemke Oliver and Soundararajan demonstrate that, though it would be expected that $\pi(x; q, \vec{a}) \sim \pi(x) / \phi(q)^r$ (where $\pi(x)$ denotes the number of primes up to $x$ and $\phi(q)$ denotes Euler’s $\phi$-function, the number reduced residue classes of $q$). However, this is not the case, as demonstrated in the following table provided by Lemke Oliver and Soundararajan for the first million primes greater than 3.

$$\pi(x_0; 3, (1, 1)) = 215,873$$
$$\pi(x_0; 3, (1, 2)) = 283,957$$
$$\pi(x_0; 3, (2, 1)) = 283,957$$
$$\pi(x_0; 3, (2, 2)) = 216,213$$

Though all of the values should be 250,000, this is clearly not the case. Lemke Oliver and Soundararajan also demonstrate that this bias exists for other, larger $x$ as well as different values of $q$. See the following table, provided by Lemke Oliver and Soundararajan that shows the values of $\pi(x; q, \vec{a})$ for the first hundred million primes modulo 10, where each of the 16 pairs should have values of 6.25 million (with $\pi(x_0) = 10^8$).

<table>
<thead>
<tr>
<th>$a$</th>
<th>$b$</th>
<th>$\pi(x; q, (a, b))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4,623,042</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>7,429,438</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>7,504,612</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>5,442,345</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6,010,982</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4,442,562</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>7,043,695</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>7,502,896</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>6,373,981</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>6,755,195</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>4,439,355</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>7,431,870</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>7,991,431</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>6,372,941</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>6,012,739</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>4,622,916</td>
</tr>
</tbody>
</table>
As noted by Lemke Oliver and Soundararajan, the diagonal classes \((a, a)\) are underrepresented and that, according to Chebyshev’s bias modulo 10, the residue classes 3 and 7 modulo 10 should contain more primes than classes 1 and 9 modulo 10 but quite the opposite is true for their data, possibly suggesting “that a different phenomenon is at play here.” Lemke Oliver and Soundararajan, based on the Hardy-Littlewood prime \(k\)-tuples conjecture, go on to develop a heuristic to explain the demonstrated, consecutive biases. The difficulty presented by the lack of proof for the Hardy-Littlewood conjecture lends to the need for numerical analysis as a stepping stone towards proof. It is important to note that for this consecutive bias, numerical evidence is once again demonstrated as the furthest step taken towards proof.
2 Results, Observations and Main Conjectures

Using the methods described in section 3, the products of \( k \) primes for \( k = 2, 3, 4, 5 \) were generated up to three different ceilings: \( 10^5, 10^6 \) and \( 10^7 \). The following figure shows the number of products of \( k \) primes found under these conditions:

![Figure 1: The number of products of \( k \) primes generated up to various ceilings.](image)

Furthermore, the products of \( k \) primes that were generated were individually modulated by \( 2, 3, \ldots, 10 \) to investigate the bias. In this section, the bias of these products of \( k \) primes is demonstrated as well as the consecutive bias of products of \( k \) prime pairs \((a, b)\). The results are not necessarily consistent at all ceilings however only the largest ceiling of \( 10^7 \) will be shown in this thesis for brevity. All results will be made available freely online at [https://github.com/Jacob-Ferrier](https://github.com/Jacob-Ferrier). Although initial results indicate that the biases found are consistent at all ceilings, data generation and testing at higher ceilings is an ongoing effort and will be made available online at a later date.
2.1 The products of $k$ primes bias

2.1.1 Products of 2 primes

Figure 2: The number of products of 2 primes generated up to a ceiling of $10^7$, with the results of their modulations by 2, 3, ..., 10. As in Figure 1, the number of products of 2 primes less than $10^7$ is 1,904,324.
2.1.2 Products of 3 primes

Figure 3: The number of products of 3 primes generated up to a ceiling of $10^7$, with the results of their modulations by 2, 3, ..., 10. As in Figure 1, the number of products of 3 primes less than $10^7$ is 2,444,359.
2.1.3 Products of 4 primes

Figure 4: The number of products of 4 primes generated up to a ceiling of $10^7$, with the results of their modulations by 2, 3, ..., 10. As in Figure 1, the number of products of 4 primes less than $10^7$ is 2,050,696.
2.1.4 Products of 5 primes

Figure 5: The number of products of 5 primes generated up to a ceiling of $10^7$, with the results of their modulations by 2, 3, ..., 10. As in Figure 1, the number of products of 5 primes less than $10^7$ is 1,349,779.
2.2 The products of $k$ primes consecutive bias

2.2.1 Products of 2 primes races

Figure 6: The number of products of 2 primes pairs $(a, b)$, modulated by 2, up to a ceiling of $10^7$.

Figure 7: The number of products of 2 primes pairs $(a, b)$, modulated by 3, up to a ceiling of $10^7$. 
Figure 8: The number of products of 2 primes pairs \((a, b)\), modulated by 4, up to a ceiling of \(10^7\).

Figure 9: The number of products of 2 primes pairs \((a, b)\), modulated by 5, up to a ceiling of \(10^7\).
Figure 10: The number of products of 2 primes pairs \((a, b)\), modulated by 6, up to a ceiling of \(10^7\).

Figure 11: The number of products of 2 primes pairs \((a, b)\), modulated by 7, up to a ceiling of \(10^7\).
Figure 12: The number of products of 2 primes pairs \((a, b)\), modulated by 8, up to a ceiling of \(10^7\).

Figure 13: The number of products of 2 primes pairs \((a, b)\), modulated by 9, up to a ceiling of \(10^7\).
Figure 14: The number of products of 2 primes pairs \((a, b)\), modulated by 10, up to a ceiling of \(10^7\).
2.2.2 Products of 3 primes races

Figure 15: The number of products of 3 primes pairs \((a, b)\), modulated by 2, up to a ceiling of \(10^7\).

Figure 16: The number of products of 3 primes pairs \((a, b)\), modulated by 3, up to a ceiling of \(10^7\).
Figure 17: The number of products of 3 primes pairs \((a, b)\), modulated by 4, up to a ceiling of \(10^7\).

Figure 18: The number of products of 3 primes pairs \((a, b)\), modulated by 5, up to a ceiling of \(10^7\).
Figure 19: The number of products of 3 primes pairs \((a, b)\), modulated by 6, up to a ceiling of \(10^7\).

Figure 20: The number of products of 3 primes pairs \((a, b)\), modulated by 7, up to a ceiling of \(10^7\).
Figure 21: The number of products of 3 primes pairs \((a, b)\), modulated by 8, up to a ceiling of \(10^7\).

Figure 22: The number of products of 3 primes pairs \((a, b)\), modulated by 9, up to a ceiling of \(10^7\).
Figure 23: The number of products of 3 primes pairs \((a, b)\), modulated by 10, up to a ceiling of \(10^7\).
2.2.3 Products of 4 primes races

Figure 24: The number of products of 4 primes pairs \((a, b)\), modulated by 2, up to a ceiling of \(10^7\).

Figure 25: The number of products of 4 primes pairs \((a, b)\), modulated by 3, up to a ceiling of \(10^7\).
Figure 26: The number of products of 4 primes pairs \((a, b)\), modulated by 4, up to a ceiling of \(10^7\).

Figure 27: The number of products of 4 primes pairs \((a, b)\), modulated by 5, up to a ceiling of \(10^7\).
Figure 28: The number of products of 4 primes pairs \((a, b)\), modulated by 6, up to a ceiling of \(10^7\).

Figure 29: The number of products of 4 primes pairs \((a, b)\), modulated by 7, up to a ceiling of \(10^7\).
Figure 30: The number of products of 4 primes pairs \((a, b)\), modulated by 8, up to a ceiling of \(10^7\).

Figure 31: The number of products of 4 primes pairs \((a, b)\), modulated by 9, up to a ceiling of \(10^7\).
Figure 32: The number of products of 4 primes pairs \((a, b)\), modulated by 10, up to a ceiling of \(10^7\).
2.2.4 Products of 5 primes races

Figure 33: The number of products of 5 primes pairs $(a, b)$, modulated by 2, up to a ceiling of $10^7$.

Figure 34: The number of products of 5 primes pairs $(a, b)$, modulated by 3, up to a ceiling of $10^7$. 
Figure 35: The number of products of 5 primes pairs \((a, b)\), modulated by 4, up to a ceiling of \(10^7\).

Figure 36: The number of products of 5 primes pairs \((a, b)\), modulated by 5, up to a ceiling of \(10^7\).
Figure 37: The number of products of 5 primes pairs \((a, b)\), modulated by 6, up to a ceiling of \(10^7\).

Figure 38: The number of products of 5 primes pairs \((a, b)\), modulated by 7, up to a ceiling of \(10^7\).
Figure 39: The number of products of 5 primes pairs \((a, b)\), modulated by 8, up to a ceiling of \(10^7\).

Figure 40: The number of products of 5 primes pairs \((a, b)\), modulated by 9, up to a ceiling of \(10^7\).
Figure 41: The number of products of 5 primes pairs \((a, b)\), modulated by 10, up to a ceiling of \(10^7\).
2.3 Conjecture

In the previous section, we have shown multiple consecutive products of $k$ primes “races” - to echo the language of [6] - for $k = 2, 3, 4, 5$, each modulated by $2, 3, ..., 10$ and up to a ceiling of $10^7$. Here, we will create three main conjectures from this data. It should be noted, as before, that the trends in data is not necessarily consistent at all ceilings. More work will be done to understand why this is in the future. Thus, these conjectures only apply to the data up to a ceiling of $10^7$. All of the races presented here show a clear consecutive bias of the products of $k$ primes. Furthermore, there appears to be distinct “leading sets” in each race and for the products of 2 primes races, a consisting winning pair.

**Definition 1.** For any given products of $k$ primes race modulated by $m$, the leading set, denoted as $L_{k,m}$ is defined by taking all of the members of the reduced residue class $m$. Call these members $a_1, a_2, ..., a_{\phi(m)}$. For all of these members, consider the corresponding set of pairs in the race $\{(a_i, b_j) \mid 1 \leq j < m\}$ for all $i = 1, 2, ..., \phi(m)$. Now, for all values of $i$, let $b_i$ denote the value of $b_j$ that is the max value the corresponding set. Thus, $(a_1, b_1), (a_2, b_2), ..., (a_{\phi(m)}, b_{\phi(m)})$ constitutes the leading set.

For example, consider the products of 2 primes race modulo 8. As 1, 3, 5 and 7 are reduced residues of 8, the leading set will be made up of pairs $(1, b_1), (3, b_3), (5, b_5), (7, b_7)$. As demonstrated in the corresponding figure in the previous section, the pairs $(1, 3), (3, 5), (5, 7)$ and $(7, 1)$ have the greatest quantity for their corresponding groups. Thus, $L_{2,8} = \{(1, 3), (3, 5), (5, 7), (7, 1)\}$.

**Conjecture 1.** For all consecutive products of $k$ primes modulated by any modulo $m$, there is a strong consecutive bias.

**Conjecture 2.** For all consecutive products of $k$ primes modulated by any modulo $m$, $\sum_{(a,b) \in L_{k,m}} b \equiv 0 \pmod{m}$.
Conjecture 3. For the products of 2 primes modulated by any modulo \( m \), there is not only a strong consecutive bias, but also always a clear winning pair: \((-1, 1) \mod m\).
3 Methods

In this thesis, we are investigating the consecutive bias of the products of \( k \) primes. Although there exist some repositories of quasiprimes, they are relatively small and do not include information on the products of \( k \) primes when \( k > 2 \). \(^1\) Thus, for this thesis, we had develop an algorithm for generating a list of products of \( k \) primes. Despite the lack of data for the products of \( k \) primes, there is no shortage of pre-generated primes. In our generation algorithm, we utilized a list of the first 2 billion primes accessed at \texttt{http://www.primos.mat.br/} \[1\]. Initially, we attempted to develop a method for generating the quasiprimes from a finite subset of the primes. However, this method proved to be unhelpful for generating a complete list of quasiprimes. The following section exemplifies this problem.

3.1 Quasiprime generation problem

Given a set of \( r \) prime numbers, say \( P = \{p_1, p_2, \ldots, p_r\} \), it is possible to generated \( m = \binom{r}{2} + r \) unique quasiprimes. Defined the resultant set as \( Q_P = \{p_ip_k \mid 1 \leq i \leq r, i \leq k \leq r\} \). This set can also be defined as \( Q_P = \{a \cdot b \mid (a, b) \in P \times P\} \), where \( P \times P \) is the Cartesian product of \( P \) with itself.

For example, let \( P = \{2, 3, 5, 7\} \). Thus, \( Q_P = \{4, 6, 10, 14, 9, 15, 21, 25, 35, 49\} \) which has expected cardinality of 10. The algorithm to generate this set operates in row-major ordering, thus a list in order of results would be:

\[
2 \times \{2, 3, 5, 7\}, 3 \times \{3, 5, 7\}, 5 \times \{5, 7\}, 7 \times \{7\} \rightarrow \\
(4, 6, 10, 14, 9, 15, 21, 25, 35, 49).
\]

The first problem with this approach arises here; the resultant list is not in ascending

\(^1\)Two such databases include the A001358 entry at OEIS (http://oeis.org/A001358) and here: https://prime-numbers.info/list/semiprimes#table.
order. A computationally expensive secondary operation is required, the SORT opera-

\[(4, 6, 10, 14, 9, 15, 21, 25, 35, 49) \xrightarrow{\text{SORT}} (4, 6, 9, 10, 14, 15, 21, 25, 35, 49).\]

There exist yet another problem with this approach; though the largest quasiprime in \(Q_P\) is 49, this set is not equivalent to the set of all quasiprimes up to, and including, 49. That set is so:

\[Q_{\leq 49} = \{4, 6, 8, 9, 10, 12, 14, 15, 18, 21, 22, 25, 26, 33, 34, 35, 38, 39, 46, 49\}.\]

This set so happens to be a subset of \(Q_P\), where \(P_* = \{2, 3, 5, 7, 11, 13, 17, 19, 23\}\).

This exemplifies a major problem as for any finite \(P\) where \(|P| \geq 3\), \(Q_P \subsetneq Q_{\leq \text{MAX}(Q_P)}\), thus the set \(P\) would not be useful to render a list of consecutive quasiprimes with no skips.

3.2 Products of \(k\) primes verification algorithm

The previous example showed that for any finite set of primes numbers, it is impossible to generate a consecutive quasiprime list with no skips, and to make the resultant, inadequate list ordered, an expensive SORT operation must be employed. Rather than generating the products of \(k\) primes from a set of primes, we instead developed an algorithm for verifying the products of \(k\) primes \textit{ab initio} the integers. This method is not a true ”generation” rather a ”verification” of the product of \(k\) primes status.

3.2.1 Implementation and Performance

The algorithm was implemented in GO, a object-oriented program language developed roughly ten years ago. The code is available online at \url{https://github.com/Jacob-Ferrier}.

For computational efficiency, the method of generating the product of \(k\) primes up to an integer height of \(n\) utilizes a divide-and-conquer approach. That is, before any calls to the algorithm are made, the list of integers up to \(n\) is divided into smaller lists to be verified.
The algorithm is then called for each individual list and deployed by a light-weight thread called a *goroutine*. These *goroutines* run concurrently and thus much more efficiently in terms of computational space and time complexity.

Once developed, the program was ran on the UNC Charlotte supercomputing cluster: a high performance computing environment. The program has been consistently ran on multi-core and high volume memory computing node, allowing for rapid computations. For the datasets used in this analysis, the table below shows the CPU times for the verification algorithm.

<table>
<thead>
<tr>
<th>$k$ ceiling num products of $k$ primes</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 $10^5$</td>
<td>23378</td>
</tr>
<tr>
<td>2 $10^6$</td>
<td>210035</td>
</tr>
<tr>
<td>2 $10^7$</td>
<td>1904324</td>
</tr>
<tr>
<td>3 $10^5$</td>
<td>25556</td>
</tr>
<tr>
<td>3 $10^6$</td>
<td>250853</td>
</tr>
<tr>
<td>3 $10^7$</td>
<td>2444359</td>
</tr>
<tr>
<td>4 $10^5$</td>
<td>18744</td>
</tr>
<tr>
<td>4 $10^6$</td>
<td>198062</td>
</tr>
<tr>
<td>4 $10^7$</td>
<td>2050696</td>
</tr>
<tr>
<td>5 $10^5$</td>
<td>11185</td>
</tr>
<tr>
<td>5 $10^6$</td>
<td>124465</td>
</tr>
<tr>
<td>5 $10^7$</td>
<td>1349779</td>
</tr>
</tbody>
</table>
4 Conclusion

In this thesis, we have developed an algorithm for generating a list of products of $k$ primes up to a desired integer ceiling. Furthermore, we have provided numerical evidence for three main conjectures corresponding to data up to a specific ceiling of $10^7$. We are far from proving these conjectures and we provide no mathematical explanation as of now; however, these conjectures are believed to be true given the numerical evidence. As previously discussed, strong conjectures such as the Riemann Hypothesis and $k$-tuple conjecture do not necessarily verify the conjectures of Lemke Oliver and Soundararajan in [8] so more work needs to be done to prove their conjectures for just the consecutive primes. It should also be noted that our conjectures are distinct from the conjectures of Lemke Oliver and Soundararajan. As counting the products of $k$ primes is somehow connected to counting the primes, in order to prove the conjectures that we presented here, one would need to also consider the conjectures from Lemke Oliver and Soundararajan. In the future, we plan to further verify our conjectures numerically at larger ceilings. In doing so, we also would like to improve the efficiency of our algorithm. Most importantly, we plan to work toward providing mathematical justification for our conjectures.
References


