NUMERICAL EXPERIMENTS WITH NONCOTOTIENTS

DAVID NG AND MICHAEL JOHN JACOBSON, JR.

Abstract. Let $\varphi(n)$ denote Euler's function. A positive integer m is called a noncototient if the equation $n-\varphi(n) = m$ has no solution for any n. Using an improved version of Yang's algorithm [3] for enumerating noncototients, we have extended the computation of noncototients to 10^{12}, providing further evidence in support of the conjectural density proposed by Pollack and Pomerance [5]. We also present data on classes of noncototients and sequences of consecutive even noncototients, that suggest a number of open conjectures.

1. Introduction

Euler's function $\varphi(n)$ is defined as the number of integers k in the range $1 \leq k \leq n$ for which the greatest common divisor $\gcd(n, k) = 1$. Let $s_\varphi(n) := n - \varphi(n)$. A number that can be expressed as $s_\varphi(n)$ is known as a *cototient*. A positive integer m is called a *noncototient* if the equation $s_\varphi(n) = m$ has no solution for any n. For instance, the first 8 noncototients are $10, 26, 34, 50, 52, 58, 86,$ and 100. Assuming that the strong Goldbach conjecture holds (every even integer $n > 6$ is the sum of 2 distinct primes p and q), then all noncototients are even.

In 1974, Erdős [4, B36] asked if a positive proportion of even numbers are noncototients. Browkin and Schinzel [1] proved in 1995 that there are infinitely many noncototients, a problem first posed by Erdős and Sierpínski [4, B36]. However, it was not until 2013 that Pomerance and Yang detected results that suggested a positive asymptotic density by extending their enumeration of the noncototients up to 10^8 [8] from Noe's computation to 10^4 found in the Online Encyclopedia of Integer Sequences [9]. Luca and Pomerance proved that a positive density exists in [5] and in 2016, Pollack and Pomerance [7] extended the enumeration further to 10^{10}. This computational lead to their proposed conjectural density

$$\Delta_\varphi := \lim_{x\to\infty} \frac{1}{\log x} \sum_{a \leq x}^{*} \frac{1}{a} e^{-a/s_\varphi(a)}, \tag{1.1}$$

2010 *Mathematics Subject Classification.* Primary 11Y55; Secondary 11B05,11A25.
Key words and phrases. noncototient, Euler's totient function, asymptotic density.

which rounds to 0.0908721526. at $x = 10^{12}$. Their data shows that up to $x = 10^{10}$ the observed density of even noncototients rounds to 0.1130, and that the densities appear to be decreasing as x increases.

In this paper, we describe a revised version of the second stage of Pomerance and Yang's algorithm [10] that yields a constant improvement in running time, and present results from using this algorithm to extend the enumeration of noncototients to 10^{12}. Our data provides further evidence that the density of even noncototients appears to be converging to the conjectural bound from [7]. In addition, we present a variety of data related to the distribution of noncototients, including the frequency of noncototient pairs, leading to a number of open conjectures.

The paper is organized as follows. In Section 2, we recall the previous enumeration algorithm. Our improved algorithm, along with an asymptotic analysis, is presented in Section 3. Our numerical results are given in Section 4, followed by a open problems and conjectures.

2. Pomerance and Yang's Algorithm

To determine whether a given number m is a noncototient, it suffices to ensure that m is not a solution to $n - \varphi(n)$ for all $n \leq 2m$. To see this, first note that from [1], for all even n,

$$n - \varphi(n) \geq \frac{n}{2},$$

with the terms being equal only when $n = 4$. Thus, if m is a cototient, where $m = n - \varphi(n)$, then it is necessarily the case that $n \leq 2m$. Therefore, m is a noncototient in the case that there is no such n for which equality is achieved.

To enumerate all the noncototients $\leq m$, one must find every even number $l \leq m$ that it is not a solution to $n - \varphi(n)$ for all $n \leq 2l$. A naive solution would run, for each even number l, a loop from 2 to $2l$ to determine whether each l is of the form $n - \varphi(n)$ for some $n \leq 2l$. It is clear, however, that this algorithm is computationally expensive, as it requires repeated calculation of $n - \varphi(n)$ without storing the information in some manner for later use. Instead, Yang's algorithm makes use of a sieve to mark off all those even numbers m that are of the form $n - \varphi(n)$ for some $n \leq 2m$. This guarantees that all the even numbers left are noncototients.

To implement this algorithm, an array is allocated to store every even number $\leq m$. Since an even number is either a cototient or a noncototient, it suffices to sieve through the entire array to mark off the former. The following two relations are used to calculate all even cototients $\leq m$:

$$s_\varphi(2k) = 2k - \varphi(k), \tag{2.1}$$

$$s_\varphi(2^{j+1}k) = 2s_\varphi(2^j k)\ . \tag{2.2}$$

Here, k is an odd integer and $j \geq 1$ is an integer. Since $s_\varphi(n) \equiv n \pmod 2$ when $n > 2$, all even cototients are of the form $s_\varphi(2^j k)$. Thus, the algorithm first computes $\varphi(k)$ for all odd numbers $1 \leq k \leq m$, using a pre-computed table in practice. For each $\varphi(k)$, the first relation above is applied to determine that $2k - \varphi(k)$ is the cototient of $2k$, and the entry corresponding to $2k - \varphi(k)$ in the array is thus marked as a cototient. The second relation is then used repeatedly to calculate the remaining even cototients of the form $s_\varphi(2^j k)$ until the result exceeds m, and each such even cototient is also marked off in the array. The unmarked array elements left after this process are all even noncototients.

The subproblem of tabulating $\varphi(k)$ for all odd integers $1 \leq k \leq m$ must also be considered to compute noncototients efficiently. Instead of directly computing $\varphi(k)$ for all $1 \leq k \leq m$, an array is created to store the value of the function evaluated at each k for constant lookup. A modified Sieve of Eratosthenes is used in conjunction with Euler's product formula

$$\varphi(k) = k \prod_{p|k} \left(1 - \frac{1}{p}\right),$$

which states that $\varphi(k)$ is equal to n multiplied by $\left(1 - \frac{1}{p}\right)$ for all prime factors p of k:

(1) First initialize an array A such that each element in the array stores its own index. This serves to distinguish primes as the array is processed.
(2) Loop for $p = 2$ to n: If $A[p] = p$, then p is a prime number, so set $A[p] = p - 1$. In the case that $A[p] = p$, update the value of each index i which is a multiple of p by multiplying $A[i]$ with $\left(1 - \frac{1}{p}\right)$.

3. Improvement to the Algorithm

In our improved version of the algorithm, the calculation of $\varphi(k)$ for all $1 \leq k \leq m$ is performed as described above. To count the noncototients $\leq m$ more efficiently, note that the second cototient relation used in Yang's algorithm, $s_\varphi(2^{j+1}k) = 2s_\varphi(2^j k)$, only serves to identify additional cototients. In other words, this method of noncototient calculation entails enumeration of all cototients, so that set negation gives the noncototients less than some bound.

However, it is possible to reduce the calculations required when enumerating noncototients. The following proposition forms the basis of the improved algorithm.

Proposition 3.1. *All even noncototients are of the form* $2d$*, where* d *is an odd number or* d *is a noncototient.*

Proof. Suppose to the contrary that there exists an even noncototient of the form $2d$, where d is even and d is a cototient. Because $s_\varphi(n) \equiv n \pmod 2$ when $n > 2$, even cototients can be expressed in the form $s_\varphi(2^j m)$ where m and j are positive integers, with m odd. Thus, $d = s_\varphi(2^j m)$ and so $2d = 2s_\varphi(2^j m)$, which can be written in the form $s_\varphi(2^{j+1}m)$. This is a cototient, contradicting the assumption that $2d$ is a noncototient. □

A *class (of noncototients)* is therefore defined to be the noncototients generated by repeatedly multiplying an odd integer by two until the result is no longer a noncototient. The first integer in a class, necessarily twice an odd number, is called a *principal noncototient*. For instance, 10 is a principal noncototient and the sole member of its noncototient class, as it is generated by the odd base 5, and 20 (twice of 10) is not a noncototient. This class contains one term, whereas the class generated by the odd base 13 (consisting of the principal noncototient 26 followed by 52) has two terms, and the class of 509203 contains infinitely many terms, as was shown by Browkin and Schinzel [1].

Instead of traversing the entire search space to mark off all cototients, a faster approach is to check whether a number belongs to a particular noncototient class. By doing so, we can immediately terminate the search of noncototients in a particular class once multiplication by two results in a cototient. We designate this last noncototient of a non-infinite class the *terminal noncototient*. It is clear that, given a terminal noncototient t, then $2t$ is a cototient of the form $2t = s_\varphi(2k)$ for some odd integer k. It cannot be of the form $2t = s_\varphi(2^{j+1}k)$ for $j \geq 1$, as this would imply that $t = s_\varphi(2^j k)$, contradicting the assumption that t is a noncototient.

In summary, the idea of our algorithm is to use the first cototient property (2.1) to find all even cototients of the form $s_\varphi(2k)$ for every odd integer $1 \leq k \leq m$. Then, rather than applying (2.2) to enumerate all cototients, we apply Proposition 3.1 to enumerate every class of noncototients, using the previously-found cototients of the form $s_\varphi(2k)$ to mark the end of each class.

Our algorithm proceeds as follows. As before, we first compute a table of φ-values for all odd integers $1 \le k \le m$, and an array to mark even cototients $\le m$ is allocated. Also as before, the first cototient property (2.1) is applied to all odd integers k until $2k > 2m$, and the resulting even cototients $s_\varphi(2k)$ are all marked off in the array. We then enumerate all noncototients by enumerating the noncototient class of each odd integer $d \le m$ as follows:

(1) For every integer i such that $0 \le i \le m/4$:
   (a) Set $t = 4i + 2$.
   (b) While $t \le m$ and entry t in the array is not marked
      (i) Identify (and count) t as a noncototient.
      (ii) Set $t = 2t$.

Note that if t hits a marked value in the array, then dividing t by two gives the terminal noncototient in the class.

Both Yang's algorithm and our version require the same pre-computation of $\varphi(k)$ values for all odd $k \le m$. Using the Sieve of Eratosthenes, this requires time in $O(m \log m)$, and dominates the overall cost in both cases. The remaining parts of both the original and our improved version of the algorithm, given the $\varphi(k)$ values, require time linear in m. However, we expect a constant improvement with the new version, since in essence the new algorithm enumerates noncototients directly instead of the cototients, and there are fewer noncototients.

4. Numerical Results

The two algorithms for calculating noncototients were implemented in `C++`. The computations were performed on the Helix cluster [6] at the University of Calgary. Helix is a computing cluster installed in March 2016, and currently acts as a test environment for exploring the use of large-memory general purpose compute nodes in research. When the computations in this paper were done, it included two large-memory servers, each with 64 CPU cores running at 2.2 GHz and 2 TB of RAM.

The time to compute $\varphi(k)$ for all odd $k \le 4 \cdot 10^{11}$ was just under 10.78 hours. This pre-computation dominates the running time, and is common to both Yang's algorithm and our improved version. Given the table of φ values, the runtime required to enumerate all noncototients less than or equal to $4 \cdot 10^{11}$ was just under 4.24 hours using the original algorithm, and

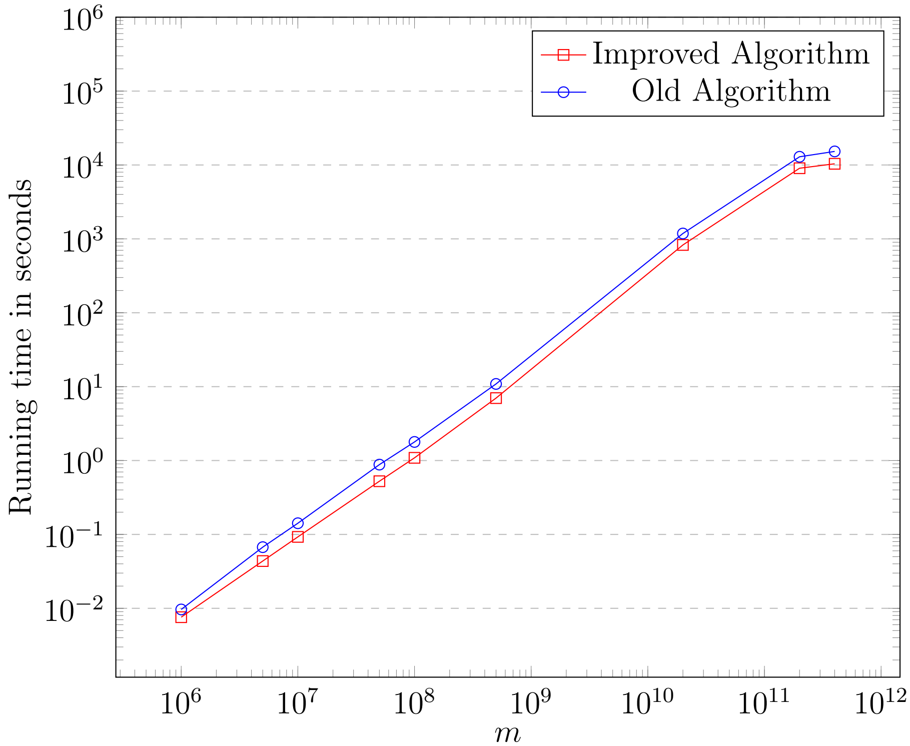

FIGURE 1. Runtime comparison of algorithms evaluated at various bounds m.

was just under 2.89 hours using the improved algorithm. The runtimes of both algorithms, not including the time to compute the table of φ values, are plotted against logarithmically-scaled bounds in Figure 1. The improved algorithm consistently finished in around 68% of the time required by the original algorithm, yielding roughly a 32% speedup, slightly more than predicted by the analysis in the previous section. The runtime data supports the observation that the noncototient enumeration algorithms have cost linear in m, but pre-computation of $\varphi(k)$ effectively increases this to $O(m \log m)$.

The storage required for computing the $\varphi(k)$ values with a sieve is one of the main bottlenecks in the algorithm, but of course can be segmented in order to work with whatever computer memory is available. In order to push our results a bit further, we implemented a segmented version of the improved algorithm. To squeeze out a bit more savings, the array representing the noncototients was also implemented as a bit-array. This implementation was used to compute all noncototients up to 10^{12}. The computation took about 6.5 days on an Intel Xeon X7560 core running at 2.27 GHz with 256 GB of RAM.

The data files containing the list of the noncototients are approximately 1.4 TB in size, and are available upon request.

4.1. **Density Results.** Our results, comparing the conjectural density with our data, are summarized in Table 1, where $N(x)$ denotes the number of noncototients less than or equal to x, $D(x) = N(x)/x$ denotes the density of noncototients less than or equal to x, and $\Delta_\varphi(x)$ denotes the evaluation of the conjectural density (1.1) at the bound x. Figure 2 plots both the observed and conjectured densities at various bounds. The data suggests that the observed density gradually decreases, possibly to the limit Δ_φ.

TABLE 1. Number of noncototients $\leq x$ for various bounds $x \leq 10^{12}$, along with their corresponding densities.

x	$N(x)$	$D(x)$	$\Delta_\varphi(x)$
10000	963	0.0963000000	0.0875145554
100000	10527	0.1052700000	0.0885211904
1000000	110786	0.1107860000	0.0891928567
10000000	1128160	0.1128160000	0.0896726519
100000000	11355049	0.1135504900	0.0900325019
1000000000	113482572	0.1134825720	0.0903123855
10000000000	1129598504	0.1129598504	0.0905362923
100000000000	11223107307	0.1122310731	0.0907194888
1000000000000	111422520897	0.1114225209	0.0908721526

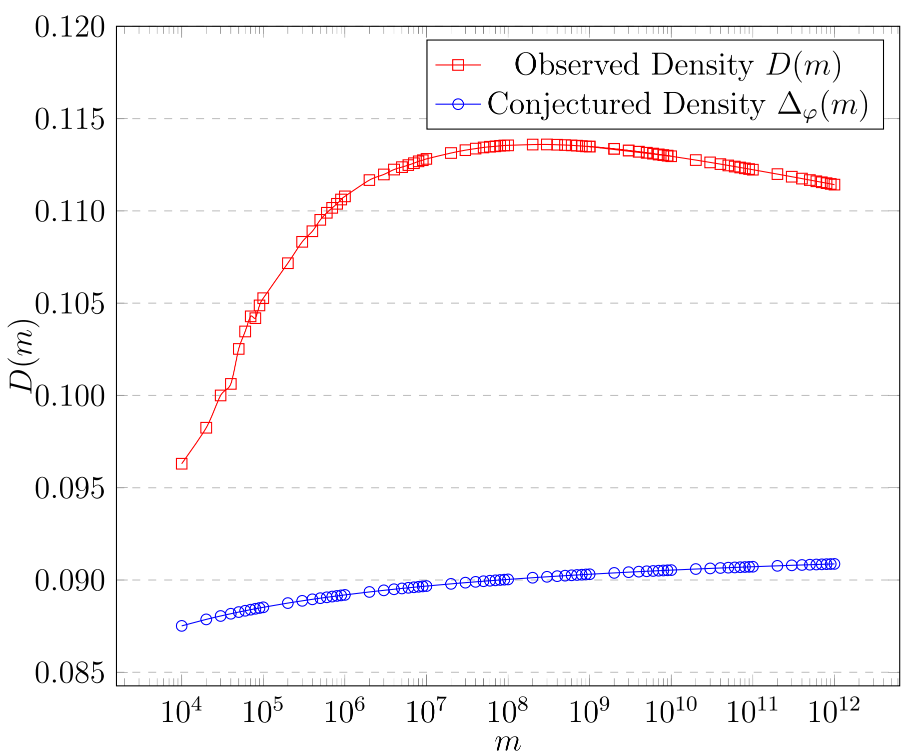

FIGURE 2. Density of noncototients $D(m)$ vs. m for $10^4 \leq x \leq 10^{12}$.

4.2. **Results on the Distribution of Noncototients.** Table 2 gives the number of classes of the specified length, where "Unknown / Infinite" indicates that the class is either infinite or has yet to terminate at 10^{12}, as well as the smallest principal noncototient ("First Appearance") generating a class of the given length. In general, it appears that classes of shorter length are more numerous that classes of greater length. We found examples of classes with all lengths up to 28 except 27; it is unknown whether for every positive integer i, there exists a class of noncototients containing i terms. Similarly, it is not known whether any relationship exists between the length of the class and the search bound required to reach said class.

TABLE 2. Number of classes of length l for noncototients to 10^{12}.

Class Length (l)	First Appearance	Number of Classes
Unknown / Infinite	1286	55587470339
1	10	2144082157
2	26	1115413259
3	58	451275610
4	134	235917405
5	634	86147965
6	86	48435254
7	1018	18965826
8	218	9505884
9	202	3501329
10	482	2110628
11	1498	706390
12	2102	368511
13	1462	132298
14	2006	85466
15	2134	25242
16	4082	15549
17	10882	5042
18	1466	3496
19	778	878
20	3074	537
21	1522	154
22	1886	145
23	11586	21
24	2870	21
25	2626	3
26	2486	4
27	0	0
28	3554	1

We call a sequence of consecutive even noncototients a *chain*. Table 3 summarizes the results regarding the distribution of chains of various lengths. It is clear that consecutive noncototient pairs must necessarily alternate between principal and non-principal noncototients, as the values alternate between 0 mod 4 and 2 mod 4. It is not known if there exists an upper bound on the length of chains. For noncototients less than 10^{12}, chains of every length up to 20 except 19 have been found. It is unknown whether we have not found a chain of length 19 because the current search bound is too small or because these chains do not exist.

TABLE 3. Number of consecutive noncototient chains of length l for noncototients to 10^{12}.

Chain Length (l)	First Appearance	Number of Chains
1	10	64217003593
2	50	15408230556
3	532	3755744247
4	2314	896000615
5	4628	219769433
6	22578	52947013
7	115024	12660551
8	221960	3014671
9	478302	703130
10	3340304	163297
11	22527850	39589
12	117335136	9497
13	1118736102	2316
14	1564578508	524
15	6121287812	121
16	7515991946	38
17	470344908044	2
18	300899994422	1
19	0	0
20	234063318774	1

By sifting through the output files listing the noncototients, it is clear that even noncototients which differ by two occur with some level of frequency. Again, of each pair, one is necessarily a principal noncototient, while the other is not. Chains of length z contain $z-1$ noncototient pairs. The density of twin noncototients can be similarly defined as the number of noncototient pairs divided by the particular search bound. The results can be seen in Figure 3, plotting the observed density of noncototient pairs against the search bound, in which it appears that the observed density may be approaching an asymptotic density similarly to the count of noncototients.

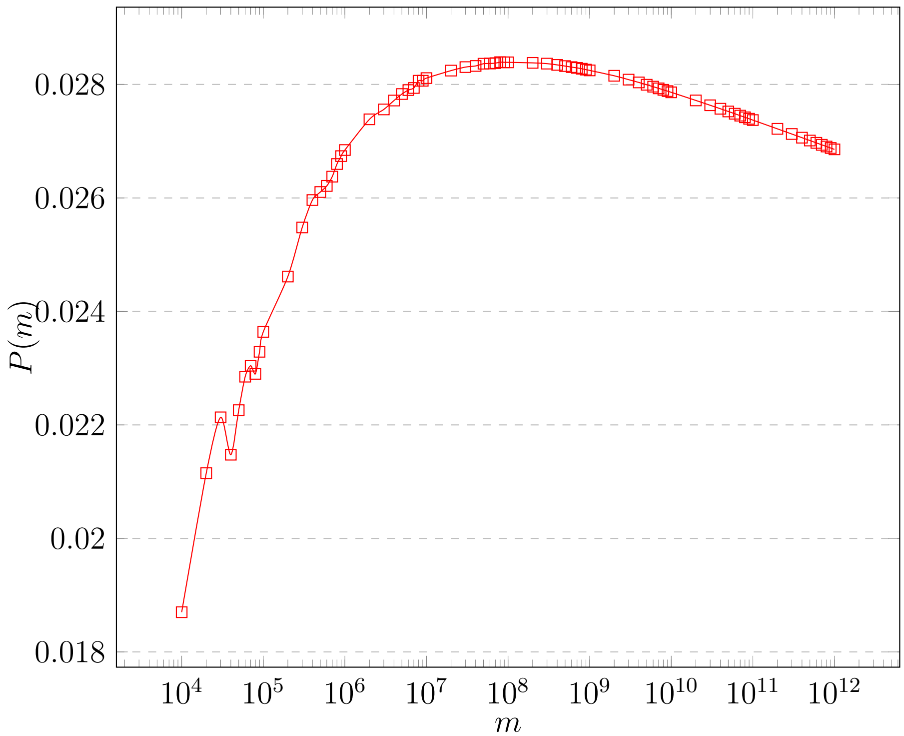

FIGURE 3. Density of noncototients pairs $P(m)$ vs. m for $10^4 \leq m \leq 10^{12}$.

5. Open Problems

Our data on the density of even cototients, as illustrated in Figure 2, provides more evidence that the asymptotic density of even noncototients does in fact approach the conjectural asymptotic density given by Pollack and Pomerance [7]. In addition, our observations lead to a number of conjectures.

Chains of consecutive even noncototients of various lengths occur, with shorter chains occurring more frequently. As shown in Table 3, we have found chains of every length up to 20 except for 19.

Conjecture 5.1. *There exist noncototient chains of every integer length* $l \geq 1$.

Classes of noncototients were used in our improved algorithm for enumerating even noncototients. Browkin and Schinzel [1] and Flammenkamp and Luca [2] proved that there exist classes of infinite length. As shown in Table 2, we have found noncototient classes of every length up to 28 except 27, but it is unknown whether there exist classes of every positive integer length.

Conjecture 5.2. *There exist noncototient classes of every integer length* $l \geq 1$.

Grytczuk and Mędryk [3, Theorem 1], following [2], give necessary and sufficient conditions for an odd prime p to generate an infinite class, i.e. that $2p$ is a principal noncototient:

- $2p$ must be a noncototient,
- p is not a Mersenne prime,
- p is a Riesel number, i.e. $2^k p - 1$ is composite for every positive integer $k \geq 1$.

Primes p that fail the third condition generate noncototient classes of length $k-1$ when $k \geq 1$ is the smallest integer such that $2^k p - 1$ is prime. Thus, data from the search for Riesel numbers, specifically from confirmed non-Riesel numbers, would almost certainly yield examples of classes of many more lengths. This connection to Riesel numbers itself may provide a direction for proving the existence of noncototient classes of every integer length.

Although there are infinitely many noncototients, it is unknown whether there are infinitely many principal noncototients, i.e. noncototients that are twice an odd integer.

Conjecture 5.3. *There exist infinitely many principal noncototients.*

Grytczuk and Mędryk [3, Theorem 3] prove that $2p$ is a noncototient for p an odd prime if and only if there are distinct odd primes $p_1, \ldots, p_r$, $r \geq 1$ such that $p = p_1 p_2 \cdots p_r - \frac{1}{2}(p_1 - 1)(p_2 - 1) \cdots (p_r - 1)$, giving one possible direction to prove that an infinitude of principal noncototients exists. Alternatively the conjecture is also implied by the following two conjectures.

Conjecture 5.4. *There exist infinitely many terminal noncototients.*

Conjecture 5.5. *There exist infinitely many noncototient pairs.*

In addition to an infinitude of noncototient pairs, our data in Figure 3 suggests that noncototient pairs may also have an asymptotic density, possibly somewhere between 0.025 and 0.026.

Conjecture 5.6. *There is a positive asymptotic density of noncototient pairs.*

We hope that our data and observations will inspire further work on these questions.

Acknowledgements. The first author was supported by awards from the University of Calgary PURE and the NSERC USRA programs. The second author's research is supported by NSERC Discovery Grant RGPIN-2016-04545.

Both authors would like to thank Doug Phillips and Carl Pomerance for their help and suggestions.

References

1. J. Browkin and A. Schinzel, *On integers not of the form $n - \phi(n)$*, Colloq. Math. **68** (1995), no. 1, 55–58. MR 1311762
2. A. Flammenkamp and F. Luca, *Infinite families of noncototients*, Colloq. Math. **86** (2000), no. 1, 37–41. MR 1799887
3. A. Grytczuk and B. Mędryk, *On a result of Flammenkamp-Luca concerning noncototient sequence*, Tsukuba J. Math. **29** (2005), no. 2, 533–538. MR 2177025
4. R. K. Guy, *Unsolved problems in number theory*, third ed., Problem Books in Mathematics, Springer-Verlag, New York, 2004. MR 2076335
5. F. Luca and C. Pomerance, *The range of the sum-of-proper-divisors function*, Acta Arith. **168** (2015), no. 2, 187–199. MR 3339454
6. University of Calgary High Performance Computing, *Helix Quickstart Guide*, `http://hpc.ucalgary.ca/quickstart/helix`, last accessed: 2017-10-05.
7. P. Pollack and C. Pomerance, *Some problems of Erdős on the sum-of-divisors function*, Trans. Amer. Math. Soc. Ser. B **3** (2016), 1–26. MR 3481968
8. C. Pomerance and H.-S. Yang, *Variant of a theorem of Erdős on the sum-of-proper-divisors function*, Math. Comp. **83** (2014), no. 288, 1903–1913. MR 3194134
9. N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at `https://oeis.org/A005278`, last accessed: 2017-10-06, Sequence A005278.
10. H.-S. Yang, *Unitary untouchable numbers*, Honor's Thesis, Dartmouth College, July 2012.

Department of Computer Science, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4

E-mail address: `david.ng2@ucalgary.ca`

Department of Computer Science, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4

E-mail address: `jacobs@ucalgary.ca`