ON THE EQUATIONS $\phi(\mathbf{n}) = \phi(\mathbf{n}+\mathbf{k})$ AND $\phi(\mathbf{p}-\mathbf{1}) = \phi(\mathbf{q}-\mathbf{1})$

KIM, SUNGJIN

Abstract. We prove that there exists positive even integer k such that $\phi(n) = \phi(n+k)$ holds for infinitely many n. We also prove various estimates on number of solutions to $\phi(p-1) = \phi(q-1)$ for distinct primes p and q.

1. Introduction

Let k be a positive integer. We study the equations $\phi(n+k) = \phi(n)$ and $\phi(p-1) = \phi(q-1)$ in this paper. Here, ϕ is the Euler's totient function, and p, q are distinct primes. The former equation has been studied by many authors. In 1958, A. Schinzel [24] proved that if p and $2p-1$ are primes not dividing an even number k, and $n = (2p-1)k$, then

$$\phi(n) = \phi(n+k). \tag{1}$$

In 1972, Lal and Gillard [15, Table 1] searched for solutions to (1) for $k \leq 30$ and $n \leq 10^5$, and observed that the equation (1) has more solutions for even k than for odd k. In 1999, Graham, Holt, and Pomerance [9, Theorem 1] showed that, if j and $j+k$ have the same prime factors with $g = (j, j+k)$, and $jr/g + 1$, $(j+k)r/g + 1$ are both primes not dividing j, then $\phi(n) = \phi(n+k)$ holds for

$$n = j\left(\frac{(j+k)r}{g} + 1\right). \tag{2}$$

Using this result and assuming a quantitative form of Dickson's prime k-tuple conjecture (see §2), they proved that for even k, the number $P_0(k;x)$ of solutions to (1) of the form (2) with $n \leq x$ satisfies $P_0(k;x) \sim c_k x/(\log x)^2$ for some positive constant c_k. Moreover, they proved that for fixed k and sufficiently large x, the number $P_1(k;x)$ of solutions to (1) not of the form (2) with $n \leq x$ satisfies

$$P_1(k;x) \ll x\exp(-(\log x)^{1/3}).$$

Let $P(k;x)$ be the number of solutions to (1) with $n \leq x$. Then we have $P(k;x) = P_0(k;x) + P_1(k;x)$. For odd numbers k, it is widely known that $P(k;x) = P_1(k;x)$. In 2013, Pollack, Pomerance, and Trevino [22] showed that

$$P_1(k;x) \ll x\exp(-(\log x)^{1/3}) \text{ uniformly for } k \leq \exp((\log x)^{1/3}).$$

In 2017, Yamada [27, Corollary 1.1] refined (2) and showed that for each odd integer k,

$$P(k;x) \ll x\exp(-(2^{-1/2} + o(1))(\log x \log\log\log x)^{1/2})).$$

On the opposite end, Schinzel [24] proved that there are at least two solutions to (1) for any positive even integer $k \leq 8 \cdot 10^{47}$. Schinzel conjectured that there are infinitely many solutions to (1) for any positive even integer k. In 2003, Holt [14] proved that there are at least two solutions to (1) for any positive even integer k. In this paper, we prove that there exist positive even integers k such that (1) has infinitely many solutions. Moreover, we are able to obtain quantitative lower bounds for the number of such k's. The main ingredients are [9, Corollary 2] and Maynard-Tao theorem (see Zhang's breakthrough paper on prime gaps [28], Maynard's results on prime gaps by an application of multidimensional Selberg sieve [16], also Polymath8b's collaborative work on reducing the upper bound of $\liminf(p_{n+m} - p_n)$ for each $m \geq 1$ [23, Theorem 16(vi)]) for admissible sets of linear forms. Throughout this paper, $m\#$ for a positive integer m denotes the primorial $\prod_{p\leq m} p$. We use symbols $A(x) \ll B(x)$, $A(x) \gg B(x)$, and $A(x) = o(B(x))$ as their usual meaning $A(x) = O(B(x))$, $B(x) = O(A(x))$, and $\lim_{x\to\infty} |A(x)/B(x)| = 0$, respectively. Denote by $P(n)$ the largest prime factor of n.

Theorem 1.1. *Let $C \geq 1$, $j = 50\#$, and $q = 49$. For sufficiently large x, there exists $k_x \in \{j, 2j, \ldots, qj\}$ such that*

$$P_0(k_x; x) \gg \frac{x}{(\log x)^{50}}.$$

As a direct consequence, we obtain that there exists a positive even integer $k \in \{j, 2j, \ldots, qj\}$ such that $\phi(n) = \phi(n+k)$ has infinitely many solutions. The following theorem and corollary address answers to Paul Pollack's questions raised in a private e-mail conversation. In the following theorem, we obtain infinitely many sets of consecutive integers so that at least some of them have equal ϕ-value.

Theorem 1.2. *For any $m \geq 1$, there exist infinitely many n such that at least $m+1$ members of the set*

$$\{n, n+1, \ldots, n + (k_m - 1) \cdot k_m\#\}$$

have equal ϕ value. Here, $k_m \geq B \exp(4m)$ for an absolute effective constant B.

As a consequence of the above theorem, we obtain a lower bound of the number of $k \leq K$ yielding infinitely many solutions to (1).

Corollary 1.1. *For any $K \geq K_0$, the number $N(K)$ of positive integers $k \leq K$ such that the equation (1) has infinitely many solutions satisfies*

$$N(K) \gg \log \log K.$$

After the author's initial submission of this paper, Kevin Ford's arXiv paper [8] appeared and a stronger result $N(K) \gg K$ is obtained in the direction of Corollary 1.1. Moreover, explicit examples of such k are obtained (multiples of 442720643463713815200).

In July 2019, a question on Mathematics Stack Exchange (MSE) [26] by a user *coffeemath* asked if there are infinitely many pairs (a, b) with $\phi(\phi(a)) = \phi(\phi(b))$ and $\phi(a) \neq \phi(b)$. This original version of the problem is answered by the author considering the numbers with prime factors 2 and 3. Through a further discussion in the comments to the question, the user *coffeemath* was interested in special cases where a and b admit primitive roots, especially, when a and b are powers of odd prime numbers. This sparked the question whether the equality

$$\phi(p-1) = \phi(q-1), \tag{3}$$

for distinct pair of primes (p, q) holds infinitely often. A conditional proof of the infinitude of solutions to (3) is given in MSE by the author. If there are infinitely many k satisfying $p = 4k+1$, $q = 6k+1$ are primes (Dickson's prime tuple conjecture), and $(6, k) = 1$, we have infinitely many nontrivial pairs of primes p and q such that $\phi(p-1) = \phi(q-1)$. The author then solved this question unconditionally at MSE by applying the Maynard-Tao theorem (see [16] and [23, Theorem 16(i)]). Hereby, we present a quantitative estimate on the number of solutions. The main ingredients are an Erdős-Kac type result [3] on the number of prime factors of $\phi(p-1)$ and the Cauchy-Schwarz (Hölder) inequality.

Theorem 1.3. *Let $\mathcal{P}$ be the set of all prime numbers. Let*

$$Q(x) = \#\{(p, q) \in \mathcal{P}^2 \mid p < q \leq x,\ \phi(p-1) = \phi(q-1)\}.$$

Then we have for sufficiently large x,

$$Q(x) \geq x \exp\left((1 + o(1))(\log \log x)^2 \log \log \log x\right).$$

More generally for $m \geq 2$, let

$$Q_m(x) = \#\{(p_1, \ldots, p_m) \in \mathcal{P}^m \mid p_1 < \cdots < p_m \leq x,\ \phi(p_1 - 1) = \cdots = \phi(p_m - 1)\}.$$

Then for sufficiently large x,

$$Q_m(x) \geq x \exp\left((m - 1 + o(1))(\log \log x)^2 \log \log \log x\right).$$

The normal order with an asymptotic formula is not available in [3]. Usually, Erdős-Kac type estimates (e.g. Berry-Esseen and several theorems in [4]) come as an asymptotic formula with main term $\pi(x)\Phi(z)$ and error term of rough size $\pi(x)/\sqrt{\log \log x}$, or an upper bound estimate (e.g. Hardy-Ramanujan). Thus, if we consider the number of $\{p \leq x \mid \omega(\phi(p-1)) \geq N(x)\}$ with $N(x) \geq (\log \log x)^2$, then the main term $\pi(x)\Phi(z)$ and the error term becomes larger than the main term. The author believes that these lower estimates are best possible with this method. To push these estimates further, we take a completely different conditional approach. The main ingredients of the second method are the argument by Erdős presented in [21, Section 6], and Maynard's theorem [17, Theorem 3.1]. This requires assuming a version of Bombieri-Vinogradov inequality for spaced moduli (see Conjecture 5.1).

Theorem 1.4 (Conditional)**.** *Let $Q(x)$ and $Q_m(x)$ be defined as in Theorem 1.3. Assuming Conjecture 5.1, we have an absolute constant $\alpha > 0$ such that for sufficiently large x,*

$$Q(x) \geq x^{1+\alpha+o(1)}, \quad \text{and } Q_m(x) \geq x^{1+\frac{m\alpha}{2}+o(1)}.$$

Paul Pollack [19, Theorem 1] proved that the set $\{m/n |\ \sigma(m) = \sigma(n)\}$ is dense in the set of positive real numbers. Applying his method, we are able to obtain analogous results for ϕ-values of shifted primes.

Theorem 1.5. *Let $\lambda \geq 1$ and $\epsilon > 0$. Then for some $K > 0$, there is $N \in (0, K] \cap \mathbb{N}$ such that*

$$Q(\lambda, \epsilon, x) = \#\left\{(p, q) \in \mathcal{P}^2 \mid p < q \leq x,\ \phi(p-1) = \phi(q-1),\ \lambda \leq \left(\frac{q-1}{p-1}\right)^{\frac{1}{N}} < \lambda + \epsilon\right\} \gg_{\lambda,\epsilon} \frac{x}{(\log x)^{50}}.$$

A direct corollary is that there is a number $K \in \mathbb{N}$ such that

$$[0, \infty) \subseteq \bigcup_{N \leq K} \overline{\left\{\frac{1}{N} \log \frac{q-1}{p-1} \,\middle|\, (p, q) \in \mathcal{P}^2,\ p < q,\ \phi(p-1) = \phi(q-1)\right\}}.$$

On the opposite end, applying a result [20, Corollary 2], we have the following upper bounds of $Q(x)$ and $Q_m(x)$, which are much stronger than the trivial bounds $\ll \frac{x^2}{(\log x)^2}$, and $\ll \frac{x^m}{(\log x)^m}$.

Theorem 1.6. *Let $Q(x)$ and $Q_m(x)$ be defined as in Theorem 1.3. Let $L(x) = \exp(\log x \log\log\log x/\log\log x)$. For sufficiently large x,*

$$Q(x) \leq \frac{x^2}{L(x)^{2-o(1)}}, \quad \text{and } Q_m(x) \leq \frac{x^m}{L(x)^{m-o(1)}}.$$

We also investigated consecutive primes p and q with $\phi(p-1) = \phi(q-1)$. It seems that there are many solutions too (see Table 1 for a code, Table 2 for an accumulated data).

Conjecture 1.1. *Define*

$$Q^*(x) = \#\{(p, q) \in \mathcal{P}^2 \mid p < q \leq x,\ \phi(p-1) = \phi(q-1),\ q \text{ is the next prime to } p\}.$$

Then we have

$$\lim_{x\to\infty} Q^*(x) = \infty.$$

Assuming Bateman-Horn conjecture, a stronger form of Conjecture 1.1 holds.

Theorem 1.7 (Conditional)**.** *Assuming the special case of Bateman-Horn conjecture, we have for sufficiently large x,*

$$Q^*(x) \gg \frac{x}{(\log x)^4}.$$

2. Estimates on the Numbers of Prime k-Tuples

We begin with this equation since the lemmas developed for this one eventually help treating the equation $\phi(n) = \phi(n+k)$. A set of m-tuple of linear forms $\{a_1x + b_1, \ldots a_mx + b_m\}$ is said to be admissible if for any prime p there is $x_p \in \mathbb{Z}$ such that $p \nmid \prod_{i=1}^m (a_ix_p + b_i)$. We consider the tuples with

$$\prod_i a_i \neq 0 \quad \text{and} \quad \prod_{i<j} (a_ib_j - a_jb_i) \neq 0.$$

The following is a special case of Bateman-Horn conjecture (a quantitative estimate on Dickson's conjecture).

Conjecture 2.1 (Bateman-Horn)**.** *Let $m \geq 2$ and $A_m = \{a_1x+b_1, \dots, a_mx+b_m\}$ be an admissible set of linear forms. Then for sufficiently large x, the number $R_m(x)$ of $r \leq x$ such that a_ix+b_i, $1 \leq i \leq m$ are all prime satisfies*

$$R_m(x) \gg_{A_m} \frac{x}{(\log x)^m}.$$

Substantial progress toward this conjecture was made by the Maynard-Tao theorem. We state the quantitative form of the Maynard-Tao theorem for an admissible set of linear forms. The proof requires slight modifications of [23]. Note that the following is unconditional.

Lemma 2.1. *Let $A_{50} = \{a_1r+b_1, \dots, a_{50}r+b_{50}\}$ be an admissible set of linear forms. Then for sufficiently large x, the number $R(A_{50}, 2, x)$ of $r \leq (x - \max_i b_i)/\max_i a_i$ such that at least two of the linear forms are primes satisfies*

$$R(A_{50}, 2, x) \gg_{A_{50}} \frac{x}{(\log x)^{50}}.$$

Let $m \geq 1$ and $A_{k_m} = \{a_1r+b_1, \dots, a_{k_m}r+b_{k_m}\}$ be an admissible set of linear forms. Moreover, assume that $k_m \geq B\exp(4m)$. Then for sufficiently large x, the number $R(A_{k_m}, m+1, x)$ of $r \leq (x - \max_i b_i)/\max_i a_i$ such that at least $m+1$ of the linear forms are primes satisfies

$$R(A_{k_m}, m+1, x) \gg_{A_{k_m}} \frac{x}{(\log x)^{k_m}}.$$

Sketch of Proof. In [23, Lemma 18], we replace $\theta(n+h_i)$ and all subsequent uses of $n+h_i$ by $\theta(a_in+b_i)$ and a_in+b_i. The definition of $\tilde{S}(d_1, \dots, d_{k-1}, d'_1, \dots, d'_{k-1})$ is modified as

$$\tilde{S}(d_1, \dots, d_{k-1}, d'_1, \dots, d'_{k-1}) = \sum_{\substack{x<n\leq 2x \\ n\equiv b\ (W) \\ a_in+b_i\equiv 0\ ([d_i,d'_i])\ \forall i=1,\dots,k-1}} \theta(a_kn+b_k).$$

Here, b is chosen so that $(a_ib+b_i, W) = 1$ for all $i = 1, \dots, k$. Then we write the multiple congruences in the summation as a single congruence modulo $q = a_kW\prod_{i=1}^{k-1}[d_i, d'_i]$ for which the following hold

$$\begin{aligned} n' &= a_kn+b_k, \\ n' &\equiv a_kb+b_k\ (W), \text{and} \\ n' &\equiv a_k(-b_ia_i^{-1})+b_k\ ([d_i, d'_i]). \end{aligned}$$

Here, a_i^{-1} is the inverse of a_i modulo $[d_i, d'_i]$. Then we rewrite the sum over n as a sum over n' under this single congruence. The lower estimates in x are due to Remark 32 of [23, Page 27].

By [23, Theorem 23(xi)], we have the result for $k_m \geq B\exp(4m)$. For admissible set of monic linear forms, we are able to apply MPZ$[\varpi, \delta]$ to reduce $B\exp(4m)$ to $B\exp((4-28/157)m)$. However, for the general admissible set of linear forms, we cannot directly apply MPZ$[\varpi, \delta]$ since the moduli may not be square-free because of the extra a_k. With $k_m \geq B\exp(4m)$, the proof only requires the Bombieri-Vinogradov theorem, and expressions in the theorems are simpler. □

Remark. It had been widely known that Maynard-Tao theorem can be applied to admissible sets of linear forms (see [10] and [2]). This method of rewriting the sum over n as a sum over n' has been explained in [18] with the admissible set $\{n, 6n+1, 12n+1\}$ of linear forms. Applying pigeon hole principle, we are able to obtain the following version of Lemma 2.2. This version will be more helpful toward proving Theorem 1.1.

Lemma 2.2. *Let $A_{50} = \{a_1r+b_1, \dots, a_{50}r+b_{50}\}$ be an admissible set of linear forms. For sufficiently large x, there exists $T_2(x) = \{i, j\} \subseteq \{1, \dots, 50\}$ with $i < j$ for which the number $R'(A_{50}, T_2, x)$ of $r \leq (x - \max_i b_i)/\max_i a_i$ such that the two linear forms $\{a_ir+b_i, a_jr+b_j\} \subseteq A_{50}$ are both primes satisfies*

$$R'(A_{50}, T_2(x), x) \gg_{A_{50}} \frac{x}{(\log x)^{50}}.$$

Let $m \geq 1$ and $A_{k_m} = \{a_1r+b_1, \dots, a_{k_m}r+b_{k_m}\}$ be an admissible set of linear forms. Moreover, assume that $k_m \geq B\exp(4m)$. For sufficiently large x, there exists $T_{m+1}(x) = \{i_1, \dots, i_{m+1}\} \subseteq \{1, \dots, k_m\}$ with

$i_1 < \cdots < i_{m+1}$ for which the number $R'(A_{k_m}, T_{m+1}, x)$ of $r \le (x - \max_i b_i)/\max_i a_i$ such that the $m+1$-tuple $\{a_{i_1}r + b_{i_1}, \dots a_{i_{m+1}}r + b_{i_{m+1}}\} \subseteq A_{k_m}$ of the linear forms are all primes satisfies

$$R'(A_{k_m}, T_{m+1}(x), x) \gg_{A_{k_m}} \frac{x}{(\log x)^{k_m}}.$$

Remark. We are not able to remove the dependence of T_2, T_{m+1} on x. However, due to the finiteness of the sets of linear forms, we can say that there is at least a pair (or an $m+1$-tuple) which are all primes infinitely often. A celebrated result by Heath-Brown [13] is that at least one of 2, 3, or 5 are primitive roots modulo p for infinitely many primes p. My question to him was whether we are able to say that $P_a(x) = \#\{p \le x \mid a \text{ is a primitive root modulo } p\} \gg x/(\log x)^2$ for some $a = 2$, 3, or 5. The conclusion in my question was indeed an incorrect application of Pigeon-hole principle. Heath-Brown's answer to this question was that we are only able to say $P_{a(x)}(x) \gg x/(\log x)^2$ where $a(x) \in \{2, 3, 5\}$.

3. The equation $\phi(n) = \phi(n+k)$

For the proof of Theorem 1.1, a modified version of [9, Theorem 4] plays a central role.

Lemma 3.1. *Suppose that j, $j+k$, ..., $j+qk$ all have the same prime factors. Define $B = \prod_{i=0}^{q}(j+ik)$. For $i = 0, \dots q$, let*

$$b_i = \frac{B}{j+ik}, \; g = \gcd(b_0, b_1, \dots, b_q),$$

$$a_i = \frac{b_i}{g} = \frac{B}{(j+ik)g}.$$

Suppose that for some positive integer r, and for some $0 \le i_1 < i_2 \le q$,

$$a_{i_1}r + 1, \; a_{i_2}r + 1 \text{ are both primes that do not divide } j.$$

Then for $n = j(a_0 r + 1) = \frac{Br}{g} + j$, we have

$$\phi(n + i_1 k) = \phi(n + i_2 k).$$

Proof. The proof is essentially identical to [9, Theorem 4]. Note that for any $0 \le i \le q$,

$$n + ik = \frac{Br}{g} + j + ik = (j+ik)\left(\frac{b_i r}{g} + 1\right) = (j+ik)(a_i r + 1).$$

Then we have

$$\phi(n+i_1k) = \phi(j+i_1k)\frac{Br}{(j+i_1k)g} = \phi(j+i_2k)\frac{Br}{(j+i_2k)g} = \phi(n+i_2k).$$

Here, $\frac{\phi(j+i_1k)}{j+i_1k} = \frac{\phi(j+i_2k)}{j+i_2k}$ since their prime factors are identical. □

We can directly apply Lemma 3.1 to prove Theorem 1.1.

Proof of Theorem 1.1. Let $j = k = 50\#$ and $q = 49$ in Lemma 3.1. The positive integers $a_0, \dots, a_q$ are distinct. Thus, $\{a_0 r + 1, \dots, a_q r + 1\}$ is an admissible set of linear forms. By Lemma 2.2, for sufficiently large x, there is a pair (i_1, i_2) depending on x, $0 \le i_1 < i_2 \le q$ such that the number of $r \le x$ for which $a_{i_1}r + 1$ and $a_{i_2}r + 1$ are both primes is $\gg x/(\log x)^{50}$. Then the number of $n \le x$ such that $\phi(n + i_1 k) = \phi(n + i_2 k)$ is $\gg x/(\log x)^{50}$. Thus, for some $0 \le i_1 < i_2 \le q$, not depending on x, we have $\phi(n + i_1 k) = \phi(n + i_2 k)$ for infinitely many n. □

The existence of $m+1$ primes in an admissible set of linear forms implies Theorem 1.2.

Proof of Theorem 1.2. Let $j = k = k_m\#$ and $q = k_m - 1$. Since $\{a_0 r + 1, \dots, a_q r + 1\}$ is an admissible set of linear forms, there are at least $m+1$ members of $\{a_0 r + 1, \dots, a_q r + 1\}$ are prime. Lemma 2.2 implies that for sufficiently large x, there is an $m+1$-tuple $(i_1, \dots, i_{m+1}) \in \{0, \dots, q\}^{m+1}$ depending on x such that $\{a_{i_1}r + 1, \dots, a_{i_{m+1}}r + 1\}$ are all primes for $\gg x/(\log x)^{k_m}$ values of $r \le x$. By Lemma 3.1, we have

$$\phi(n + i_1 k) = \cdots = \phi(n + i_{m+1} k)$$

for such values of r. Thus, there must be a tuple $(i_1, \dots, i_{m+1}) \in \{0, \dots, q\}^{m+1}$ not depending on x such that all members in $\{a_{i_1}r+1, \dots, a_{i_{m+1}}r+1\}$ are primes infinitely many r. Thus, the result follows. $\square$

We have Corollary 1.1 as a direct consequence of Theorem 1.2.

Proof of Corollary 1.1. With $j = k = k_m\#$ and $q = k_m - 1$, we observed in the proof of Theorem 1.2 that there is an $m+1$-tuple $(i_1, \dots, i_{m+1}) \in \{0, \dots, q\}^{m+1}$ with $i_1 < \cdots < i_{m+1}$ such that

$$\phi(n + i_1 k) = \cdots = \phi(n + i_{m+1}k)$$

holds for infinitely many n. Then for any $k' \in \{(i_2 - i_1)k, \dots, (i_{m+1} - i_1)k\}$,

$$\phi(n) = \phi(n + k')$$

for infinitely many n.

Let $K \geq K_0$ and $(k_m - 1)k_m\# \leq K < (k_{m+1} - 1)k_{m+1}\#$. Then the number of $k' \leq K$ such that $\phi(n) = \phi(n + k')$ has infinitely many solutions is at least m. Applying $k_m = \lceil B\exp(4m)\rceil$ in Lemma 2.3, we have $m \gg \log\log K$. The result follows. $\square$

4. The equation $\phi(p-1) = \phi(q-1)$

Let us say that a set of linear forms $\{a_1r + b_1, \dots, a_mr + b_m\}$ $(m \geq 2)$ is strongly admissible if for any prime p, there is $x_p \not\equiv 0 \bmod p$ such that $p \nmid \prod_{i=1}^{m}(a_ix_p + b_i)$. We remark that for any integer $t \geq 2$, there are infinitely many positive integers z so that $\phi(x) = z$ has exactly t distinct solutions. This is a celebrated result by Ford [7, Theorem 1]. Thus, we are able to find a set $\{a_i\}_{i=1}^{t}$ of positive integers with $\phi(a_i) = z$ for all $1 \leq i \leq t$. Moreover, an algorithm of finding solutions to $\phi(x) = z$ for fixed z is provided at [25]. The following procedure is to obtain a strongly admissible set of linear forms $\{a_1r + 1, \dots, a_mr + 1\}$ so that all $\phi(a_i)$, $1 \leq i \leq m$ are equal.

Procedure S(m).

Step 1. Let $u \geq m$ be an odd number and $t = 2^{\pi(u)}u$. By [7, Theorem 1], we begin with a set M of positive integers $a_1, \dots, a_t$ of the equal ϕ-value.

Step 2. Take a prime $p \leq u$. Find the number N_p of integers in M for which are 1 mod p. If $N_p < t/2$, then discard those N_p integers from M. If $N_p \geq t/2$, then multiply p to those N_p integers and discard remaining $|M| - N_p$ integers. Repeat Step 2 with a different prime $p' \leq u$.

Step 3. Repeat Step 2 until all primes $p \leq u$ are used.

Step 4. After Step 3, we have $|M| \geq t/2^{\pi(u)} = u$ and none of them are 1 mod p for any prime $p \leq u$. Moreover, all elements in M have the equal ϕ-value. Take m elements from M and discard all others. Let $M = \{a_1, \dots, a_m\}$ after relabeling. Then $\{a_1r+1, \dots, a_mr+1\}$ is a set of strongly admissible linear forms.

The following lemma and its idea of proof are similar to those of [2, Theorem 1].

Lemma 4.1. *Let $\{a_1r + 1, \dots, a_mr + 1\}$ be a set of strongly admissible linear forms. Then there exists positive integers q, y_0 such that $(\prod_i a_i, qy+y_0) = 1$ for any $y \in \mathbb{Z}$. Furthermore, $\{a_1(qy+y_0)+1, \dots, a_m(qy+y_0)+1\}$ is a set of admissible linear forms.*

Proof. Since $\{a_ir + 1\}_{i=1}^{m}$ is strongly admissible, for each prime $p|q := \mathrm{LCM}(\prod_i a_i, (m+1)\#)$, there exists $x_p \not\equiv 0 \bmod p$ such that $p \nmid \prod_i(a_ix_p + 1)$. By Chinese remainder theorem, there is $y_0 \in \mathbb{Z}$ such that $qy+y_0 \equiv x_p \bmod p$ for each $p|q$ and for all $y \in \mathbb{Z}$. Since $(q, y_0) = 1$, we have $(\prod_i a_i, qy+y_0) = 1$ for any $y \in \mathbb{Z}$. Note that the primorial $(m+1)\#$ in the definition of q ensures that $\{a_1(qy + y_0) + 1, \dots, a_m(qy + y_0) + 1\}$ is admissible. $\square$

Let $N(t)$ be the number of solutions to the equation $\phi(x) = t$. Pollack [20, Corollary 2] showed that for sufficiently large x,

$$\sum_{t \le x} N(t)^2 \le \frac{x^2}{L(x)^{2-o(1)}}.$$

Note that a slight modification of the argument in [20, Corollary 2] yields that for $m \ge 2$,

$$\sum_{t \le x} N(t)^m \le \frac{x^m}{L(x)^{m-o(1)}}.$$

Let $\phi_1 < \phi_2 < \cdots < \phi_\ell$ be the distinct numbers in $\{\phi(p-1) | p \le x,\ p \in \mathcal{P}\}$ written in an increasing order. Let m_i be the number of solutions $p \in \mathcal{P}$ with $\phi(p-1) = \phi_i$ and $p \le x$. Then $m_i \le N(\phi_i)$ for each $i \le \ell$. Then we may write

$$Q(x) = \sum_{\substack{i \le \ell \\ m_i \ge 2}} \binom{m_i}{2}.$$

With this set-up, the proof of Theorem 1.6 is relatively short.

Proof of Theorem 1.6. We have

$$Q(x) = \sum_{\substack{i \le \ell \\ m_i \ge 2}} \binom{m_i}{2} \le \sum_{i \le \ell} m_i^2 \le \sum_{t \le x} N(t)^2 \le \frac{x^2}{L(x)^{2-o(1)}}.$$

For $m \ge 2$, we have

$$Q_m(x) = \sum_{\substack{i \le \ell \\ m_i \ge m}} \binom{m_i}{m} \le \sum_{i \le \ell} m_i^m \le \sum_{t \le x} N(t)^m \le \frac{x^m}{L(x)^{m-o(1)}}.$$

□

To obtain a lower estimate of $Q(x)$, we need the following Erdös-Kac type result of Bassily, Katai, and Wijsmuller [3] for $\omega(\phi(p-1))$. See also [6] for the results on $\omega(\phi(n))$.

Lemma 4.2 (Bassily, Katai, and Wijsmuller). *Let $\omega(n) = \sum_{p|n} 1$ be the number of prime factors of n counted without multiplicity, and $\pi(x) = \sum_{p \le x} 1$ be the prime counting function. For sufficiently large x, we have for any $u \in \mathbb{R}$,*

$$\lim_{x \to \infty} \frac{1}{\pi(x)} \# \left\{ p \le x \middle| \frac{\omega(\phi(p-1)) - \frac{1}{2}(\log\log x)^2}{\frac{1}{\sqrt{3}}(\log\log x)^{\frac{3}{2}}} \le u \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{y^2}{2}}\, dy.$$

Remark. Lemma 4.2 shows us that $\phi(p-1)$ has many prime factors normally. More precisely, we have

$$\omega(\phi(p-1)) = (1/2 + o(1))(\log\log x)^2$$

for $(1 + o(1))\pi(x)$ primes $p \le x$. But, there are not so many numbers $n \le x$ with this many prime factors. Therefore, ϕ function maps a large subset of shifted primes $\{p-1|\ p \le x\}$ into a small subset of $\{\phi(p-1)|\ p \le x\}$. Recall that $\phi_1 < \phi_2 < \cdots < \phi_\ell$ are the distinct numbers in $\{\phi(p-1) | p \le x,\ p \in \mathcal{P}\}$ written in an increasing order.

Lemma 4.3. *Let A_x be the set of primes $p \le x$ with $|\omega(\phi(p-1)) - \frac{1}{2}(\log\log x)^2| \le (\log\log x)^{4/3}$. Let $\{i_1, i_2 \ldots, i_{\ell_A}\} \subseteq \{1, \ldots, \ell\}$ so that $i_j < i_{j+1}$ for each $j < \ell_A$, and $\phi_{i_1} < \phi_{i_2} < \cdots < \phi_{i_{\ell_A}}$ are distinct numbers in $B_x = \{\phi(p-1) | p \in A_x\}$ written in an increasing order, so $|B_x| = \ell_A \le \ell$. Then for sufficiently large x,*

$$\ell_A \ll x \exp\left(-\left(1 + o(1)\right)(\log\log x)^2 \log\log\log x\right).$$

Proof. We have $|A_x| = (1+o(1))\pi(x)$ by Lemma 4.2, and any $n \in B_x$ satisfies

$$\omega(n) = \frac{1}{2}(\log\log x)^2 + O((\log\log x)^{4/3}).$$

We count the number of $n \le x$ with this property by applying Hardy-Ramanujan inequality [12, (2.21)]. There is an absolute constant $C > 0$ such that for any $k = 1, 2, \ldots,$

$$\sum_{n\le x,\ \omega(n)=k} 1 < \frac{Cx}{\log x}\frac{(\log\log x + C)^{k-1}}{(k-1)!}.$$

We sum this up over k with $|k - \frac{1}{2}(\log\log x)^2| = O((\log\log x)^{4/3})$. If X is Poisson random variable with parameter $\lambda > 0$, then for any $y > \lambda$, Chernoff's bound yields

$$P(X \ge y) \le \frac{(e\lambda)^y e^{-\lambda}}{y^y}.$$

Applying this for Poisson random variable with parameter $\lambda = \log\log x + C$ and $y = (\frac{1}{2} + o(1))\lambda^2$, we obtain

$$\#\left\{n \le x \middle| \left|\omega(n) - \frac{1}{2}(\log\log x)^2\right| = O((\log\log x)^{4/3})\right\} \ll x\exp\left(-(1+o(1))(\log\log x)^2\log\log\log x\right).$$

Therefore, the result follows. □

We are now ready to prove Theorem 1.3 for $Q(x)$ and $Q_m(x)$. We apply Cauchy-Schwarz inequality for $Q(x)$, Hölder inequality for $Q_m(x)$.

Proof of Theorem 1.3. We consider the numbers ϕ_{i_j} as in Lemma 4.3. In addition to those, consider m_{i_j} too. By Lemma 4.2,

$$\sum_{j\le \ell_A} m_{i_j} = |A_x| = (1+o(1))\pi(x).$$

Thus, by Cauchy-Schwarz inequality, we have

$$\sum_{j\le \ell_A} m_{i_j} \le \left(\sum_{j\le \ell_A} 1^2\right)^{1/2}\left(\sum_{j\le \ell_A} m_{i_j}^2\right)^{1/2}.$$

It follows that

$$(1+o(1))\pi(x) \le \ell_A^{1/2}\left(\sum_{j\le \ell_A} m_{i_j}^2\right)^{1/2}.$$

Then by Lemma 4.3,

$$\frac{((1+o(1))x/\log x)^2}{x\exp\left(-(1+o(1))(\log\log x)^2\log\log\log x\right)} \le \sum_{j\le \ell_A} m_{i_j}^2.$$

Therefore,

$$x\exp\left((1+o(1))(\log\log x)^2\log\log\log x\right) \le \sum_{j\le \ell_A} m_{i_j}^2 \le \sum_{i\le \ell} m_i^2.$$

On the other hand,

$$Q(x) = \sum_{j\le \ell}\binom{m_i}{2} = \sum_{j\le \ell}\frac{1}{2}(m_i^2 - m_i) \ge x\exp\left((1+o(1))(\log\log x)^2\log\log\log x\right) + O(\pi(x)),$$

where $O(\pi(x))$ and $\frac{1}{2}$ can be absorbed into $\exp(o((\log\log x)^2\log\log\log x))$. Thus, we have the result

$$x\exp\left((1+o(1))(\log\log x)^2\log\log\log x\right) \le Q(x).$$

To treat $Q_m(x)$, we will replace Cauchy-Schwarz inequality with Hölder inequality. We have

$$\sum_{j\le \ell_A} m_{i_j} \le \left(\sum_{j\le \ell_A} 1^{\frac{m}{m-1}}\right)^{1-\frac{1}{m}} \left(\sum_{j\le \ell_A} m_{i_j}^m\right)^{\frac{1}{m}}.$$

It follows that

$$(1+o(1))\pi(x) \le \ell_A^{1-\frac{1}{m}} \left(\sum_{j\le \ell_A} m_{i_j}^m\right)^{\frac{1}{m}}.$$

Thus,

$$\frac{((1+o(1))\pi(x))^m}{\ell_A^{m-1}} \le \sum_{j\le \ell_A} m_{i_j}^m.$$

Then by Lemma 4.3,

$$x\exp\left((m-1+o(1))(\log\log x)^2\log\log\log x\right) \le \sum_{j\le \ell_A} m_{i_j}^m.$$

We split the sum as

$$\sum_{j\le \ell_A,\ m_{i_j}\le m} m_{i_j}^m + \sum_{j\le \ell_A,\ m_{i_j}>m} m_{i_j}^m.$$

The former sum is bounded by $m^m\ell_A$. Since $m(m_{i_j}-m+1)\ge m_{i_j}$ provided that $m_{i_j}\ge m$, the latter sum satisfies

$$\sum_{j\le \ell_A,\ m_{i_j}>m} m_{i_j}^m \le \sum_{j\le \ell_A} m^m m! \binom{m_{i_j}}{m} \le \sum_{i\le \ell} m^m m! \binom{m_i}{m} = m^m m! Q_m(x).$$

Since the constants and $O(\ell_A)$ are absorbed into $\exp\left(o((\log\log x)^2\log\log\log x)\right)$, we obtain

$$x\exp\left((m-1+o(1))(\log\log x)^2\log\log\log x\right) \le Q_m(x).$$

□

For Theorem 1.5, we begin with the following lemma which is analogous to [19, Lemma 1].

Lemma 4.4. *Let $A\in\mathbb{N}$, $m\ge 2$, and $0<\alpha<\beta$. Then there are infinitely many m-tuples of pairwise relatively prime square-free integers $a_1,\dots,a_m$ and $1\le K_i\le k_m-1$ with $\sum_{i<m} K_i \le k_m-1$ such that*

$$(A, a_1\cdots a_m) = 1,\quad \phi(a_1) = \cdots = \phi(a_m),$$

and

$$e^{K_i\alpha} \le \frac{a_{i+1}}{a_i} \le e^{K_i\beta} \quad \textit{for}\quad 1\le i<m.$$

Proof. Let $\alpha<\alpha'<\beta'<\beta$. The sum $\sum_{p>A}\log(1-\frac{1}{p})$ is divergent for any $A\in\mathbb{N}$, and $1/p\to 0$ as $p\to\infty$. Thus, the set $\{b/\phi(b) \mid b \text{ is square-free},\ (A,b)=1\}$ is dense in $[1,\infty)$. We can find pairwise relatively prime square-free integers with $(A,b_i)=1$ for each i, so that

$$b_1<\cdots<b_{k_m},\quad \phi(b_1)<\cdots<\phi(b_{k_m}),$$

$$\frac{b_1}{\phi(b_1)}<\cdots<\frac{b_{k_m}}{\phi(b_{k_m})},\quad \text{and } \alpha'<\log\frac{b_{i+1}}{\phi(b_{i+1})}-\log\frac{b_i}{\phi(b_i)}<\beta'.$$

Let $B=\prod_{i\le k_m} b_i$ and $a_i = b_i\left(\frac{\phi(B)}{\phi(b_i)}r+1\right)$. By Lemma 2.1, m expressions $p_j=\frac{\phi(B)}{\phi(b_{i_j})}r+1>B$, $j=1,\dots m$, $1\le i_1<\cdots<i_m\le k_m$ are simultaneously primes for infinitely many r. For such r, we have $\phi(a_{i_j})=\phi(B)r$ for all $j=1,\dots,m$. Here, we have $a_1<\cdots<a_{k_m}$ and as $r\to\infty$,

$$\frac{a_{i_{j+1}}}{a_{i_j}} \to \left(\frac{b_{i_{j+1}}}{\phi(b_{i_{j+1}})}\right) \Big/ \left(\frac{b_{i_j}}{\phi(b_{i_j})}\right).$$

Thus, for sufficiently large r,

$$(i_{j+1}-i_j)\alpha \le \log \frac{a_{i_{j+1}}}{a_{i_j}} \le (i_{j+1}-i_j)\beta$$

and $1 \le i_{j+1}-i_j \le k_m - 1$. We relable $(a_{i_j})_{j\le m}$ as $(a_i)_{i\le m}$ and $(i_{j+1}-i_j)_{j<m}$ as $(K_i)_{i<m}$. It is possible to generate infinitely many such tuples $(a_1,\dots,a_m)$ by this method. Since there are at most $\binom{k_m}{m}$ tuples $(K_i)_{i<m}$, there is a fixed tuple $(K_i)_{i<m}$ such that the conclusion of Lemma 4.4 works for infinitely many $(a_1,\dots,a_m)$. □

Now, proof of Theorem 1.5 follows similarly as in [19, Theorem 1].

Proof of Theorem 1.5. Let $\lambda \ge 1$, $\epsilon > 0$, $K \in [2,\infty)\cap\mathbb{N}$, $e^{K\alpha} = \lambda + \epsilon/3$, and $e^{K\beta} = \lambda + 2\epsilon/3$. We apply Lemma 4.4 repeatedly for $m = 2^{\pi(51)}\cdot 51$. Then we have a fixed tuple (K_j) with $1 \le K_j \le k_m - 1$, $\sum_j K_j \le k_m - 1$, and pairwise relatively prime square-free numbers $\{a_{ij}\}$ with the following properties.

1. For each i, $\phi(a_{i1}) = \cdots = \phi(a_{im})$.
2. For each i, j, $e^{K_j\alpha} \le \frac{a_{i,j+1}}{a_{ij}} \le e^{K_j\beta}$.

For each $k \ge 2$, consider $\{(\prod_{i\le k} a_{ij})r + 1\}_{j\le m}$. We apply Procedure $S(50)$ to obtain a set of strongly admissible linear forms $\{(P_k\prod_{i\le k} a_{ij})r + 1\}_{j\le 50}$. Here, P_k is a square-free integer relatively prime to $\prod_{i\le k} a_{ij}$ which is composed of extra primes multiplied in Step 2 of Procedure $S(50)$. For each j, the values $\phi(P_k\prod_{i\le k} a_{ij})$ are identical. By Lemma 4.1 and Lemma 2.2, for sufficiently large x, the number of $r \ll x$ with $(P_k\prod_{j\le 50}\prod_{i\le k} a_{ij}, r) = 1$ such that at least two ($j=j_1$, $j=j_2$) of the linear forms $\{(P_k\prod_{i\le k} a_{ij})r + 1\}_{j\le 50}$ are primes is $\gg x/(\log x)^{50}$. Let p, q be primes obtained by such r and $1 \le j_1 < j_2 \le 50$ (the choice of one pair (j_1,j_2) may depend on x), then $\phi(p-1) = \phi(q-1)$. Moreover, by 2, for such primes p, q, there is $1 \le N_k = N_k(x) \le k_m - 1$ such that

$$e^{kN_k\alpha} \le \frac{q-1}{p-1} = \frac{P_k(\prod_{i\le k} a_{ij_2})r}{P_k(\prod_{i\le k} a_{ij_1})r} \le e^{kN_k\beta}.$$

Thus, for $k = K$ and $N = N_K$,

$$\lambda < \left(\frac{q-1}{p-1}\right)^{\frac{1}{N}} < \lambda + \epsilon.$$

□

5. Conditional Theorems

Lemma 2.1 and 2.2 are applicable when the admissible set of linear forms is fixed. Maynard [17, Theorem 3.1] proved that under well-distributed hypothesis [17, Hypothesis 1], the coefficients of the linear forms are bounded by a fixed power of x. We need the following special case of Maynard's theorem.

Lemma 5.1 (Maynard)**.** *Let $\mathcal{P}$ be the set of all primes. Let $\mathcal{A}(x) = [x,2x]\cap\mathbb{N}$, $\mathcal{A}(x;q,a) = \{n \in \mathcal{A}(x) \mid n \equiv a\ (q)\}$, and $\mathcal{L} = \{L_i\}_{i\le k}$ be an admissible set of linear forms, and all coefficients a_i, b_i satisfy $a_i, b_i \le x^\alpha$ for some $\alpha > 0$. For any $L(n) = a_i n + b_i \in \mathcal{L}$, denote by $\phi_L(q) = \phi(a_i q)/\phi(a_i)$, $\mathcal{P}_L(x) = L(\mathcal{A}(x))\cap\mathcal{P}$, and $\mathcal{P}_L(x;q,a) = L(\mathcal{A}(x;q,a))\cap\mathcal{P}$. Assume that there are constants C and δ depending on α and θ such that $\delta > (\log k)^{-1}$ and the following hold for $k \ge C$ and $L \in \mathcal{L}$.*

(1) $\sum_{q\le x^\theta} \max_a \left|\#\mathcal{A}(x;q,a) - \frac{\#\mathcal{A}(x)}{q}\right| \ll \frac{\#\mathcal{A}(x)}{(\log x)^{100k^2}}$.

(2) $\sum_{q\le x^\theta} \max_{(L(a),q)=1} \left|\#\mathcal{P}_L(x;q,a) - \frac{\#\mathcal{P}_L(x)}{\phi_L(q)}\right| \ll \frac{\#\mathcal{P}_L(x)}{(\log x)^{100k^2}}$.

(3) For any $q \le x^\theta$, we have $\#\mathcal{A}(x;q,a) \ll \frac{\mathcal{A}(x)}{q}$.

(4) $\frac{1}{k}\sum_{L\in\mathcal{L}} \frac{\phi(a_i)}{a_i}\#\mathcal{P}_L(x) \ge \delta\frac{\#\mathcal{A}(x)}{\log x}$.

Then we have

$$\#\{n \in \mathcal{A}(x) \mid \{L_1(n), \ldots, L_k(n)\} \cap \mathcal{P} \geq C^{-1}\delta \log k\} \gg \frac{\#\mathcal{A}(x)}{(\log x)^k \exp(Ck)}.$$

The conditions (1) and (3) are easy to prove, but (2) and (4) are often difficult in general. If a_i, b_i are bounded by a fixed power of $\log x$, then (2) with $\theta = 1/2 - \epsilon$ follows by Bombieri-Vinogradov inequality, and (4) follows by Siegel-Walfisz theorem. If a_i and b_i are bounded by a fixed power of x, then a version of Bombieri-Vinogradov inequality for spaced moduli implies (2) and (4).

Conjecture 5.1. *Let $0 < \rho < \theta = 1/2 - \epsilon$. For any $A > 0$ and $d \leq x^\rho$, we have*

$$\sum_{\substack{q \leq x^\theta \\ d|q}} \max_{(a,q)=1} \left| \pi(x; q, a) - \frac{\pi(x)}{\phi(q)} \right| \ll \frac{x}{\phi(d)(\log x)^A}.$$

Currently available unconditional results are obtained by Elliott [5] and Guo [11]. Elliott proved that the above is true when $d \leq x^{1/3} \exp(-(\log \log x)^3)$ is a power of a fixed number. Guo improved the range of d to $d \leq x^{2/5} \exp(-(\log \log x)^3)$. Clearly, Conjecture 5.1 implies (2) and (4) of Lemma 5.1.

Proof of Theorem 1.4. By a similar argument in [21, Section 6], there is $v \leq x$ such that there are $x^{1-1/\theta+o(1)}$ solutions $a \leq x$ with $\phi(a) = v$. Let $y = \log x$ and $n = \lfloor y^\theta/(\log y)^{O(1)} \rfloor$ so that there are n primes $p \leq y^\theta$ with $P(p-1) \leq y$. By [1, Theorem 1], this is possible with $\theta = 3.377$. Let $w = \lfloor \log x/\log(y^\theta) \rfloor$ and consider $a \leq x$ formed by the product of w primes chosen from those n primes. Enumerate them as $\{a_1, \ldots, a_M\}$ so that $a_i < a_{i+1}$ and $M = \binom{n}{w} = x^{1-1/\theta+o(1)}$. Note that a_i's are square-free with the same number of prime factors. Let $\rho, \epsilon > 0$ be arbitrarily small. Further following [21, Section 6], the set B of all pairs (a_i, a_j), $i < j \leq M$ with $\omega(a_j/\gcd(a_i, a_j)) < (1-\epsilon)w$ satisfies $|B| \leq x^{(2-\epsilon)(1-1/\theta)+o(1)}$. We say that pairs in B bad pairs.

Let k be a fixed odd number so that $C^{-1}\delta \log k \geq 2$ in Lemma 5.1. We shall obtain a pair of primes (p, q) with $p < q$ by following the steps.

1. Select $t = 2^{\pi(k)}k$ members from $\{a_1, \ldots, a_M\}$ so that the collection does not contain any bad pair.

2. Apply Procedure $S(k)$ to obtain a strongly admissible set of linear forms $\{P_k a_{i_1} r + 1, \ldots, P_k a_{i_k} r + 1\}$. Here, $P_k = O(1)$ is the product of primes multiplied in Step 2 of Procedure $S(k)$, and all $\phi(P_k a_{i_s})$, $s \leq k$ are identical.

3. Apply Lemma 4.1 and Lemma 5.1 (with assuming Conjecture 5.1) to obtain a pair of primes $p = P_k a_{i_s}(Qy + y_0) + 1$, $q = P_k a_{i_t}(Qy + y_0) + 1$, $1 \leq s < t \leq k$ with $(Qy + y_0, a_{i_s} a_{i_t}) = 1$ and $y \leq x^\rho$. For such pairs, $\phi(p-1) = \phi(q-1)$.

Now, we count the number of pairs of primes that could be obtained from these steps. There are at least $\binom{M}{t} - \binom{M-2}{t-2}|B| = x^{t(1-1/\theta)+o(1)}$ selections in Step 1. We obtain $\asymp \phi(Q)$ possible values of y_0 in Lemma 4.1. From the admissible set of linear forms in Step 3, we obtain the pairs of primes for $\gg x^\rho/(\log x)^k$ values of $y \leq x^\rho$. The pair $\{a_{i_s}, a_{i_t}\}$ in Step 3 may be obtained in at most $\binom{M-2}{t-2} = x^{(t-2)(1-1/\theta)+o(1)}$ other selections in Step 1. Moreover, for any pair (a_i, a_j) belonging to a selection in Step 1, and (a_u, a_v) belonging to any selection in Step 1, $a_j/a_i = a_v/a_u$ holds for at most $x^{\epsilon(1-1/\theta)+o(1)}$ times. Therefore, the number of pairs obtained by these three steps is

$$\gg x^{(2-\epsilon)(1-1/\theta)+o(1)} \frac{x^\rho}{(\log x)^k} \phi(Q) = x^{(2-\epsilon)(1-1/\theta)+\rho+o(1)} Q.$$

The size of these primes is $\ll xQx^\rho = x^{1+\rho}Q$. Finally, note that the ratio of exponents of these numbers with base x is arbitrarily close to

$$\frac{2(1-1/\theta) + \frac{\log Q}{\log x}}{1 + \frac{\log Q}{\log x}} \geq \frac{2(1-1/\theta) + k}{1+k} = 1 + \alpha$$

with $\alpha > 0$ is an absolute constant. The estimate for $Q(x)$ now follows.

To treat $Q_m(x)$ for $m \geq 2$, recall the definition of m_{i_j} and ℓ_A. By the Hölder inequality,

$$\sum_{j\leq \ell_A} m_{i_j}^2 \leq \left(\sum_{j\leq \ell_A} 1^{\frac{m/2}{m/2-1}}\right)^{1-\frac{2}{m}} \left(\sum_{j\leq \ell_A} m_{i_j}^{2\cdot\frac{m}{2}}\right)^{\frac{2}{m}}.$$

We have

$$\left(\frac{x^{1+\alpha+o(1)}}{\ell_A^{1-2/m}}\right)^{\frac{m}{2}} \leq \sum_{j\leq \ell_A} m_{i_j}^m.$$

Applying the upper bound of ℓ_A obtained in the proof of Theorem 1.3, we obtain

$$x^{1+\frac{m\alpha}{2}+o(1)} \leq \sum_{j\leq \ell_A} m_{i_j}^m.$$

Therefore, the estimate of $Q_m(x)$ follows as in Theorem 1.3. □

Through a computer search (see below), there are several solutions even when p and q are consecutive primes. Define

$$Q^*(x) = \#\{(p,q) \in \mathcal{P}^2 \mid p < q \leq x,\ \phi(p-1) = \phi(q-1),\ q \text{ is the next prime to } p\}.$$

```
from sympy.ntheory import totient
from sympy import primerange, nextprime
file2 = open('phiconsecutive.txt', 'w+')
N=10000000000
s=0
for p in list(primerange(1,N+1)):
    if totient(p-1)==totient(nextprime(p)-1):
        file2.write(str(p))
        file2.write(' ')
        file2.write(str(nextprime(p)))
        file2.write(' ')
        file2.write(str(totient(p-1)))
        file2.write(' ')
        file2.write(str(s+1))
        file2.write('\n')
        s+=1
file2.write(str(s))
file2.close()
```

Table 1. A python code calculating $Q^*(x)$

The following is a part of the output of this program. If we take x from the second column, then the number in the fourth column is $Q^*(x)$. It was not possible to finish running this code due to a memory error. However, it terminated with $x = 4299327457$ with output file created up to that point.

```
999856811 999856841 399942720 50400
999887843 999887849 499943920 50401
999923363 999923369 499961680 50402
999931333 999931357 333310440 50403
999976163 999976169 499988080 50404
1000019107 1000019113 333339696 50405
1000032493 1000032499 333344160 50406
1000049917 1000049929 333349968 50407
1000055347 1000055359 333351780 50408
```

```
1000094917 1000094923 333364968 50409
1000115437 1000115449 333371808 50410
1000132297 1000132327 333377424 50411
...................................
4299143237 4299143249 2149571616 167891
4299152483 4299152489 2149576240 167892
4299232517 4299232529 2149616256 167893
4299283057 4299283063 1433094336 167894
4299327433 4299327457 1433109120 167895
```

Table 2. Some parts of output of the code in Table 3

Upon this data, we conjecture (Conjecture 1.1) that $Q^*(x)$ tends to infinity as x. Assuming Bateman-Horn conjecture, we are able to prove that $Q^*(x) \gg x/(\log x)^4$. To see this, we need to combine the approaches to both equations considered in this paper.

Proof of Theorem 1.7. By $q = 1$ case of Lemma 3.1, if j and $j + k$ have the same prime factors with $g = (j, j+k)$, and $jr/g+1$, $(j+k)r/g+1$ are both primes that do not divide j, then we have $\phi(n) = \phi(n+k)$ with

$$n = j\left(\frac{(j+k)r}{g} + 1\right).$$

If further

$$p = \frac{j(j+k)r}{g} + j + 1, \quad q = \frac{j(j+k)r}{g} + j + k + 1 \text{ are both primes,}$$

then we also have $\phi(p-1) = \phi(q-1)$. Thus, we consider the four linear forms $jr/g+1$, $(j+k)r/g+1$, $j(j+k)r/g+j+1$, and $j(j+k)r/g+j+k+1$. With $j = k = g = 6$, we have an admissible set of linear forms

$$\{r+1,\ 2r+1,\ 12r+7,\ 12r+13\}.$$

Following the argument in [2, Theorem 1], for the number 11, we find a prime $q = 7$ which do not divide any of $11-12$, $11-6$, $11-7$, and $11-13$. We find an integer $a = 2$ so that

$$12a + 11 \equiv 0\ (7).$$

Then we have

$$12a+12 \not\equiv 0,\ 12a+6 \not\equiv 0,\ 12a+7 \not\equiv 0,\ 12a+13 \not\equiv 0\ (7)$$

Let r have the form $7y + 2$. Then we have an admissible set of linear forms

$$\{7y+3,\ 14y+5,\ 84y+31,\ 84y+37\}$$

which is obtained by $\{7y+a+1,\ 2(7y+a)+1,\ 12(7y+a)+7,\ 12(7y+a)+13\}$. Thus, for all $t \in \{8, 9, \ldots, 12\}$ and sufficiently large y, $12(7y+a)+t$ is not prime. By Conjecture 2.1, the number of $y \le x$ such that the above are simultaneously prime is $\gg x/(\log x)^4$. For such y, the primes $p = 84y+31$ and $q = 84y+37$ must be consecutive primes, and $\phi(p-1) = \phi(q-1)$. Moreover, the primes p and q are sexy primes. □

Acknowledgement.
The author thanks Paul Pollack for addressing [3] and suggesting ways to improve Theorem 1.3. The author thanks the referee for also suggesting ways to improve Theorem 1.3 from its earlier version. The author also thanks Tristan Freiberg, Terence Tao, and Roger Heath-Brown for helpful conversations.

References

[1] R. C. Baker, G. Harman, Shifted primes without large prime factors, *Acta Arith.* **83**(1998) 331-361.
[2] W. D. Banks, T. Freiberg, C. L. Turnage-Butterbaugh, Consecutive Primes in Tuples, *Acta Arith.* **167**(2015) 261-266.
[3] N. L. Bassily, I. Kátai, M. Wijsmuller, Number of Prime Divisors of $\phi_k(n)$, where ϕ_k Is the k-fold Iterate of ϕ, *J. Number Theory* **65**(1997) 226-239.
[4] P. D. T. A. Elliott, Probabilistic Number Theory II: Central Limit Theorems, *Springer* 1980.
[5] P. D. T. A. Elliott, Primes in progressions to moduli with a large power factor, *Ramanujan J.* **13**(2007) 241-251.
[6] P. Erdős, C. Pomerance, On the Normal Order of Prime Factors of $\phi(n)$, *Rocky Mountain J. Math.* **15** (1985) 343-352.
[7] K. Ford, The Number of Solutions of $\phi(x) = m$, *Ann. Math.* **150**(1999) 281-312.

[8] K. Ford, Solutions of $\phi(n) = \phi(n+k)$ and $\sigma(n) = \sigma(n+k)$, *Int. Math. Res. Not.*, rnaa218, available at `https://doi.org/10.1093/imrn/rnaa218`
[9] S. W. Graham, J. J. Holt, C. Pomerance, On the Solutions to $\phi(n) = \phi(n+k)$, *Number Theory in Progress, eds. K. Gyory, H. Iwaniec, and J. Urbanowicz, eds.*, Vol. 2 (de Gruyter, Berlin, 1999), pp. 867-882.
[10] A. Granville, Primes in Intervals of Bounded Length, *Bull. Amer. Math. Soc.* **52**(2015) 171-222.
[11] R. Guo, Primes in Arithmetic Progressions to Moduli with a Large Power Factor, *Adv. Pure Math.* **3**(2013) 25-32.
[12] G. H. Hardy, S. Ramanujan, The Normal Number of Prime Factors of a Number n, *Quart. J. Math.* **48**(1917) 76-92.
[13] D. R. Heath-Brown, Artin's Conjecture for Primitive Roots, *Quart. J. Math.* **37**(1986) 2738.
[14] J. Holt, The Minimal Number of Solutions to $\phi(n) = \phi(n+k)$, *Math. Comput.* **72**(2003) 2059-2061.
[15] M. Lal, P. Gillard, On the Equation $\phi(n) = \phi(n+k)$, *Math. Comput.* **26**(1972) 579-583.
[16] J. Maynard, Small Gaps Between Primes, *Ann. Math.* **181**(2015) 383-413.
[17] J. Maynard, Dense Clusters of Primes in Subsets, *Composito Math.* **152**(2016) 1517-1554.
[18] R. Murty, A. Vatwani, A Remark on the Lang-Trotter and Artin Conjectures, *Proc. Amer. Math. Soc.* **146**(2018) 3191-3202.
[19] P. Pollack, Remarks on Fibers of Sum-of-Divisors Function, *Analytic Number Theory: In Honor of Helmut's 60th Birthday*, eds. M. Rassias and C. Pomerance (Springer, 2015) pp. 305-320.
[20] P. Pollack, Popular Subsets for Euler's ϕ-Function, *Math. Ann.* **374**(2019) 253-271.
[21] P. Pollack, C. Pomerance, Some Problems of Erdős on the Sum-of-Divisors Function, *Trans. Amer. Math. Soc. Ser. B* **3**(2016) 1-26.
[22] P. Pollack, C. Pomerance, E. Treviño, Sets of Monotonicity for Euler's Totient Function, *Ramanujan J.* **30**(2013) 379-398.
[23] D. H. J. Polymath, Variants of Selberg Sieve, and Bounded Intervals Containing Many Primes, *Res. Math. Sci.* **1**(2014) 1-83.
[24] A. Schinzel, Sur l'équation $\phi(x+k) = \phi(x)$, *Acta Arith.* **4**(1958) 181-184.
[25] A. J. Towns, Solving $\phi(x) = n$, where $\phi(x)$ is Euler's Totient Function - Testing Carmichael's Conjecture; *Number Theory Web*, available at `http://www.numbertheory.org/php/carmichael.html#carmichael`
[26] User coffeemath, Solutions to $\phi(\phi(a)) = \phi(\phi(b))$, available at `https://math.stackexchange.com/questions/3309655/solutions-to-varphi-varphia-varphi-varphib`
[27] T. Yamada, On Equations $\sigma(n) = \sigma(n+k)$ and $\phi(n) = \phi(n+k)$, *J. Combin. Number Theory* **9**(2017) 15-21.
[28] Y. Zhang, Bounded Gaps between Primes, *Ann. Math.* **179**(2014) 1121-1174.

California State University Northridge, 18111 Nordhoff St, Northridge, CA 91330

Santa Monica College, 1900 Pico Blvd, Santa Monica, CA 90405
E-mail address: `707107@gmail.com`