Designs, Codes and Cryptography (2024) 92:1341–1378
https://doi.org/10.1007/s10623-023-01347-w

Average case error estimates of the strong Lucas test

Semira Einsele[1] · Kenneth Paterson[2]

Received: 13 February 2023 / Revised: 6 September 2023 / Accepted: 2 December 2023 /
Published online: 18 January 2024

Abstract

Reliable probabilistic primality tests are fundamental in public-key cryptography. In adversarial scenarios, a composite with a high probability of passing a specific primality test could be chosen. In such cases, we need worst-case error estimates of the test. However, in many scenarios, the numbers are randomly chosen and thus have a significantly smaller error probability. We are hence interested in average-case error estimates. In this paper we establish such bounds for the strong Lucas primality test, as there exist only worst-case, but no average-case error bounds. This allows us to use this test with more confidence. Let us examine an algorithm that draws odd k-bit integers uniformly and independently, runs t independent iterations of the strong Lucas test with randomly chosen parameters, and outputs the first number that passes all t consecutive rounds. We attain numerical upper bounds on the probability that a composite is returned. Moreover, we examine a slight modification of this algorithm that only considers integers that are not divisible by small primes, yielding improved bounds. In addition, we classify the numbers that contribute most to our estimate.

Keywords Strong Lucas test · Secure prime generation · Average case error estimate · Lucas–Carmichael numbers

Mathematics Subject Classification 11Y11

1 Introduction

Prime generation is a basic cryptographic operation as most modern public-key cryptosystems make use of large prime numbers, either as secret or public parameters. A way to generate large primes is to choose integers of appropriate size at random and then test them for primality until a prime is found. This encourages us to find primality testing algorithms that are polynomial in complexity. Several sophisticated general-purpose algorithms that deterministically test primality exist, but their efficiency is not sufficient for most applications. In practice, one

Communicated by L. Mérai.

✉ Semira Einsele
semira.einsele@fu-berlin.de

[1] Department of Mathematics and Computer Science, FU Berlin, Takustr. 9, Berlin, Berlin, Germany

[2] Department of Computer Science, ETH Zürich, Universitätstrasse 6, 8092 Zürich, Zürich, Switzerland

therefore makes use of *probabilistic primality* tests, which are randomized primality testing algorithms that have a small probability of letting a composite, pass as prime. Nearly all known probabilistic primality tests are based on the same principle. From the input number n, one defines an Abelian group and then tests if the group structure we expect to see if n were prime is present. If n is composite this structure is not always, but often absent. Such a test can be strenghtened by running multiple independent rounds of the test. In this paper, we will refer to both probabilistic and deterministic tests as primality tests.

There are certain scenarios where the public-key parameters, such as in the Diffie-Hellman key exchange protocol, may have been chosen by an adversary. The integer could be constructed in such a way that it has a high probability of falsely being declared prime by a specific primality test, even though it is actually composite. Hence, in this scenario, the *worst case* error probability ([9], [2]) of the primality test should be small. However, for many other applications where the integer is randomly chosen, such as prime generation, it is more important to know how the test behaves in the *average case*, as it seems that most randomly chosen composites would be accepted with a probability much smaller than the so-called worst-case numbers. More formally, let us examine an algorithm that repeatedly chooses random odd k-bit integers and runs t iterations of the primality test on each candidate. If the candidate passes all t consecutive iterations, the algorithm returns that number, otherwise, another randomly chosen odd k-bit integer is selected and tested. The algorithm ends when a number that passes all t consecutive rounds is found. The error probability that this algorithm returns a composite is called the average-case error probability.

Numerous probabilistic primality tests exist, some of which are better than others. Among these, a class utilizes so-called Lucas sequences as their foundation. Let D, P and Q be integers such that $D = P^2 - 4Q$ is non-zero and $P > 0$. Let $U_0(P, Q) = 0$, $U_1(P, Q) = 1$, $V_0(P, Q) = 2$ and $V_1(P, Q) = P$. The *Lucas sequences* $U_n(P, Q)$ and $V_n(P, Q)$ associated with the parameters P, Q are defined recursively for $n \geq 2$ by

$$U_n(P, Q) = PU_{n-1}(P, Q) - QU_{n-2}(P, Q),$$
$$V_n(P, Q) = PV_{n-1}(P, Q) - QV_{n-2}(P, Q).$$

For an integer n, let $\epsilon_D(n) = \left(\frac{D}{n}\right)$ denote the Jacobi symbol. In 1980, Baillie and Wagstaff [4] gave a thorough treatment of the use of Lucas sequences in primality testing and examined various congruences that hold for prime numbers. One of their results, stated in the following theorem, lays the foundation of this paper.

Theorem 1 *(Baillie, Wagstaff [4]) Let P and Q be integers and $D = P^2 - 4Q$. Let p be a prime number not dividing $2QD$. Write $p - \epsilon_D(p) = 2^\kappa q$, where q is odd. Then*

$$\text{either } p \mid U_q \text{ or } p \mid V_{2^i q} \text{ for some } 0 \leq i < \kappa. \tag{1}$$

From this theorem, we can derive a probabilistic primality test for an integer n with a fixed D by checking property (1) of Theorem 1 for several bases (P, Q) that are chosen uniformly at random, where $1 \leq P, Q \leq n$, $\gcd(Q, n) = 1$ and $P = D^2 - 4Q$. This test is called *the strong Lucas test.* If (1) does not hold for some base (P, Q), then n is certainly composite. We call such (P, Q) a *witness for compositeness* using the strong Lucas test, which is a short proof that n is composite. However if (1) is true for several bases, even though this does not serve as proof, it is very likely that n is a prime.

Arnault [3] demonstrated that the worst-case numbers of the strong Lucas test occur for twin-prime products, which are products of two primes with a prime gap of 2. In such cases, half of the bases (P, Q) used in the test declare the integer to be prime. For the remaining odd integers, at most $4/15$th of the bases pass the test. These results serve as the worst-case error estimate of the strong Lucas test. Luckily, excluding twin-prime products does not impose great restrictions, as they can easily be detected by running Newton's method for square roots prior to conducting the actual test.

From Arnault's result, it may be tempting to directly conclude that for non-twin-prime products, the average-case estimate for t rounds of the strong Lucas test is at most $(4/15)^t$. This reasoning is wrong, as the following discussion shows. For any k, denote M_k the set of odd k-bit integers. Let $t \geq 1$ be fixed and choose k sufficiently large such that the density of the primes in M_k is much less than $(4/15)^t$. Assume that for most composites in M_k the probability that the integer that we test for primality passes a test with randomly chosen bases is about $4/15$. Then, of course, the probability of it passing t independent tests is about $(4/15)^t$. Suppose that we have an integer n from M_k that passes t tests. Since we are assuming that primes in M_k are scarce, it will be much more likely that n is composite rather than prime, so the average-case estimate would be close to 1. Naturally, the average-case estimate is much smaller than the worst-case, so we need a different argument for obtaining average-case bounds.

When it comes to choosing dependable primality tests with both average and worst-case bounds, the Miller-Rabin test stands out as a widely used method in practice. Its widespread use can be attributed to its well-established theoretical foundations and its straightforward implementation. Rabin [17] and Monier [15], working independently and almost simultaneously, laid the groundwork for the worst-case error bounds associated with the Miller-Rabin test. Additionally, Damgård, Landrock, and Pomerance [7] established average-case error bounds, further solidifying trust in this test within the cryptographic community.

For the strong Lucas test, on the contrary, only worst-case upper bounds are known, and randomly choosing worst-case numbers is rather an unusual occurrence. We are hence concerned with finding average-error estimates for the strong Lucas test. Such results would allow us to employ this test with more confidence in practice. For this, we consider an algorithm that draws odd k-bit odd integers independently from the uniform distribution, runs t independent iterations of the strong Lucas test with randomly chosen parameters on each candidate, and outputs the first one that passes all t consecutive rounds. Let $q_{k,t}$ denote the probability that a number output by this algorithm is composite.

In this paper, we conduct a thorough analysis of this error probability and derive explicit numerical upper bounds for $q_{k,t}$. These bounds are obtained by adapting the methods used in [7] to the strong Lucas case.

Furthermore, we note that integrating trial division by small primes before conducting the strong Lucas test results in improved error estimates. In practice, the inclusion of trial division is not a restrictive assumption but rather a common practice in cryptographic software to enhance the run-time efficiency of the test. Consequently, we introduce a new error probability, denoted as $q_{k,l,t}$. This probability accounts for divisibility by the first l odd primes before running the strong Lucas test. Any integer found to be divisible by these primes is discarded, and a new random k-bit integer is selected. Once the integer successfully passes this initial stage, the algorithm proceeds to execute t rounds of the strong Lucas test, as previously described. By performing a similar analysis as before, we derive numerical upper bounds for $q_{k,l,t}$.

Furthermore, we identify the numbers that contribute most to our probability estimate in the strong Lucas test and realize that, amongst others, special types of Lucas-Carmichael numbers belong to this set. We are only able to bound the cardinality of these integers upon an additional assumption. Under this assumption, we proceed to treat them differently in our analysis, leading to improved bounds for large values of t. Bounding these numbers unconditionally remains an open question.

$$
\begin{aligned}
q_{k,1} &< \ln(k)k^2 4^{2.3-\sqrt{k}} \text{ for } k \geq 2,\\
q_{k,l,1} &< k^2 4^{1.8-\sqrt{k}} \rho_l^{2\sqrt{k-1}-2} \text{ for } k \geq 2, l \in \mathbb{N},\\
q_{k,t} &< \ln(k)^t \frac{k^{3/2}}{\sqrt{t}} 4^{2.12+t-\sqrt{tk}} \text{ for } k \geq 79,\ 3 \leq t \leq k/9 \text{ or } k \geq 88,\ t = 2,\\
q_{k,l,t} &< 4^{1.72-\sqrt{tk}} k^{3/2} 2^t \rho_l^{2\sqrt{kt}+t} \text{ for } k \geq 21,\ 2 \leq t \leq (k-1)/9,\ l \in \mathbb{N},\\
q_{k,l,t} &\leq 2^{-1.52-4t} \frac{\rho_l^{6t}}{2^t - \rho_l^t} k + \rho_l^{3t} 2^{-3.55-\frac{4k}{9}-2t} k^{15/4} + \rho_l^{5t} 2^{1.74-\frac{k}{4}-3t} k\\
&\quad \text{for } k \geq 122,\ t \geq k/9,\ l \in \mathbb{N},
\end{aligned}
$$

where $\tilde{p}_l$ is the l-th odd prime and $\rho_l = 1 + \frac{1}{\tilde{p}_{l+1}}$.

2 Preliminaries

2.1 The Miller–Rabin test

The Miller-Rabin primality test, often referred to as the strong probable prime test, is a probabilistic primality test. It is one of the most widely used primality tests and exploits the following theorem:

Theorem 2 *(Miller [14], Rabin [18]) Let p be a prime and write $p - 1 = 2^{\tilde{\kappa}} \tilde{q}$, with $\tilde{q}$ odd. Then,*

$$\text{either } a^q \equiv 1 \bmod p \text{ or } a^{2^i \tilde{q}} \equiv -1 \bmod p \text{ for } 0 \leq i < \tilde{\kappa}. \tag{2}$$

The Miller-Rabin test consists of checking property (2) of Theorem 2 for multiple bases a that are chosen uniformly at random. Finding an a for which (2) does not hold is a direct proof of the compositeness of n. On the other hand, if the property (2) holds for several bases a, then n is likely to be prime.

Composite numbers that satisfy condition (2) are called *strong pseudoprimes* with respect to the base a. Rabin [17] and Monier [15] showed in 1980 independently the following theorem, which provides an upper bound for the probability that this test gives an incorrect answer.

Theorem 3 *(The Rabin-Monier Theorem [17], [15]) Let $n \neq 9$ be an odd composite integer. Let $S(n)$ denote the number of all bases a relatively prime to n such that $0 < a < n$ makes n a strong pseudoprime. We have*

$$S(n) \leq \frac{1}{4}\varphi(n),$$

where φ is the Euler function.

With Theorem 3, we can directly conclude that the Miller-Rabin test has a worst-case error probability of $1/4$, as $\varphi(n)$ is upper bounded by n. However, when applied to most composite numbers, the test tends to exhibit a significantly smaller error probability than indicated by the worst-case behavior. The first known result that took advantage of this observation was by Damgård, Landrock, and Pomerance in [7]. They considered an algorithm that repeatedly chooses random odd k-bit integers, subjects each number to t iterations of the Miller-Rabin test with randomly chosen bases, and outputs the first number found that passes all t consecutive tests. Let $p_{k,t}$ denote the probability that the algorithm falsely outputs a composite. The authors obtained numerical upper bounds for $p_{k,t}$ for various choices of k, t and obtained an upper bound for $p_{k,t}$ for certain infinite classes of k, t. These bounds, which are formulated in the following theorem, are still the best bounds we have for this primality testing algorithm.

Theorem 4 *(Damgård, Landrock, Pomerance [7]) Let k and t be integers with $k \geq 2$.*

(i) $p_{k,1} < k^2 4^{2-\sqrt{k}}$ *for* $k \geq 2$.
(ii) $p_{k,t} < k^{3/2} \frac{2^t}{\sqrt{t}} 4^{2-\sqrt{tk}}$ *for* $k \geq 21, 3 \leq t \leq k/9$ *or* $k \geq 88, t = 2$.
(iii) $p_{k,t} < \frac{7}{20} k 2^{-5t} + \frac{1}{7} k^{15/4} 2^{-k/2-2t} + 12k2^{-k/4-3t}$ *for* $k \geq 21$ *and* $t \geq k/9$.
(iv) $p_{k,t} < \frac{1}{7} k^{15/4} 2^{-k/2-2t}$ *for* $k \geq 21$ *and* $t \geq k/4$.

For specific large values of k, the paper has even better results, for example, $p_{500,1} < 4^{-28}$. Thus, if a randomly chosen odd 500-bit number passes just one iteration of a random Miller-Rabin test, the probability of it being composite is rather small. Therefore, in most practical applications, such numbers can safely be accepted as “prime”.

2.2 Strong Lucas pseudoprimes

For the remainder of the paper, let D be fixed. The strong Lucas test is based on Theorem 1, which states that for prime numbers property (1) is always fulfilled. Unfortunately, there exist composites that satisfy property (1) for a specific base (P, Q), while the property might not hold for another base. This gives rise to the following definition:

Definition 1 Let n be an odd composite number relatively prime to $2QD$. Write $n - \epsilon_D(n) = 2^\kappa q$, with q odd. Suppose that

$$\text{either } n \mid U_q \text{ or } n \mid V_{2^i q} \text{ for some } \quad 0 \leq i < \kappa,$$

i.e., it satisfies property (1) for the choice of parameters (P, Q). Then n is called a *strong Lucas pseudoprime with respect to P and Q*. For short, we write $slpsp(P, Q)$.

Definition 2 Let n be an odd composite integer relatively prime to $2D$. We let $SL(D, n)$ denote the number of pairs (P, Q) with $0 \leq P, Q < n, \gcd(Q, n) = 1, P^2 - 4Q \equiv D \bmod n$ that make n a $slpsp(P, Q)$, as defined in Definition 1.

Arnault [3] proved the explicit formula for $SL(D, n)$ for fixed n and D.

Theorem 5 *(Arnault [3]) Let D be an integer and $n = p_1^{r_1} \cdot \ldots \cdot p_s^{r_s}$ be the prime decomposition of an integer $n \geq 2$ relatively prime to $2D$. Put*

$$\begin{cases} n - \epsilon_D(n) = 2^\kappa q \\ p_i - \epsilon_D(p_i) = 2^{k_i} q_i \text{ for } 1 \leq i \leq s \end{cases} \quad \text{with } q, q_i \text{ odd},$$

ordering the p_i's such that $k_1 \leq \ldots \leq k_s$. We have

$$SL(D, n) = \prod_{i=1}^{s} (\gcd(q, q_i) - 1) + \sum_{j=0}^{k_1 - 1} 2^{js} \prod_{i=1}^{s} \gcd(q, q_i), \tag{3}$$

where $SL(D, n)$ is as defined in Definition 2. If n is not relatively prime to $2D$, we set $SL(D, n) = 0$.

Let us introduce the following function, which serves as a variant of the Euler function φ.

Definition 3 (Arnault [3]) Let D be an integer. The following number-theoretic function is defined only on integers relatively prime to $2D$:

$$\begin{cases} \varphi_D(p^r) = p^{r-1}(p - \epsilon_D(p)) \text{ for any prime } p \nmid 2D \text{ and } r \in \mathbb{N}, \\ \varphi_D(p_1 p_2) = \varphi_D(p_1)\varphi_D(p_2) \text{ if } \gcd(p_1, p_2) = 1. \end{cases}$$

We observe that this definition closely resembles the definition of φ. Specifically, if $\epsilon_D(p) = 1$ for all p, then the two definitions would coincide. The next theorem could be seen as an analog to Theorem 3 with φ_D.

Theorem 6 *(Arnault [3]) If n is an odd composite integer relatively prime to D, then*

$$SL(D, n) \leq \frac{\varphi_D(n)}{4}.$$

In the context of the Miller-Rabin test, Theorem 3 directly implies that $S(n) < n/4$ as $\varphi(n) < n$ for every $n \in \mathbb{N}$. However, in contrast, we will see in Corollary 1 that there are infinitely many n for which $\varphi_D(n)$ remains unbounded by n. Consequently, Theorem 6 is here not of the same interest as Theorem 3. Nonetheless, we have the following useful result:

Theorem 7 *(Arnault [3]) Let D be an integer and $n \neq 9$ a composite integer relatively prime to $2D$. For every integer D, we have*

$$SL(D, n) \leq \frac{4n}{15},$$

except if n is a product of twin primes of the form $n = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1)$ with q_1 odd and such that the Jacobi symbols satisfy $\epsilon_D(2^{k_1}q_1 - 1) = -1,\ \epsilon_D(2^{k_1}q_1 + 1) = 1$. In this case we have $SL(D, n) \leq n/2$.

Hence, Theorem 7 implies that for every odd composite integer n, not a product of twin primes, at most $4/15$ of the bases declare the integer to be prime. The constraint of excluding twin-prime products does not impose a great restriction, since for $\epsilon_D(n) = -1$ with $n = p(p+2)$, the decomposition $n - \epsilon_D(n) = (p+1)^2$ can easily be detected using Newton's method for square roots before running the expensive primality test. Similarly, for $\epsilon_D(n) = 1$, Newton's method can still be applied as in this case $n - \epsilon_D(n)$ is almost a square.

However, the worst-case upper bound of $SL(D, n)/n$ of $4/15$ for n is rather an uncommon occurrence. Therefore, our primary focus will be on determining the average value of $SL(D, n)/n$ for an odd composite n, rather than concentrating solely on identifying the worst-case numbers. This approach allows us to gain an insight into the typical error behaviour of the strong Lucas test when we choose n uniformly at random.

2.3 Some Lemmas

In this section, we establish lemmas that will be used in later proofs. Let us first introduce some definitions.

Definition 4 For an odd integer n, let

$$\alpha_D(n) = \frac{SL(D, n)}{\varphi_D(n)}.$$

Thus, by Theorem 6, we have $\alpha_D(n) \leq 1/4$ for odd composite n.

Definition 5 For an $n \in \mathbb{N}$, let $\omega(n)$ denote the number of distinct prime factors of n and let $\Omega(n)$ denote the number of prime factors of n counted with multiplicity.

Thus, $\omega(n) = s$ and $\Omega(n) = \sum_{i=1}^{s} r_i$.

Now let $n - \epsilon_D(n) = 2^\kappa q$, with q odd. Also let $n = p_1^{r_1} \cdot \ldots \cdot p_s^{r_s}$ be the prime decomposition of an integer relatively prime to $2D$, ordering the $p_i's$ such that $k_1 \leq \ldots \leq k_s$ in the decomposition $p_i - \epsilon_D(p_i) = 2^{k_i} q_i$, where q_i is odd. This implies that k_1 is the largest integer such that $2^{k_1} \mid p_i - \epsilon_D(p_i)$ for all $i = 1, 2, \ldots, s$. We shall always let p denote a prime number.

Lemma 1 *(Suwa [20]) Let $n = p_1^{r_1} \cdot \ldots \cdot p_s^{r_s} > 1$ be an odd integer. Let κ and k_i with $i = 1, \ldots, s$ be as defined above. Then, $\kappa \geq k_1$. Furthermore, equality holds if and only if there is an odd number of prime factors p_i of n with odd exponent such that $k_i = k_1$.*

Lemma 2 *Let $m, s \in \mathbb{N}$. Then,*

$$\left(1 + \sum_{j=0}^{m} 2^{js}\right) \leq 2^{ms+1}.$$

The details of the proofs are omitted since this can be shown by induction on m.

Lemma 3 *If $n = p_1^{r_1} \cdot \ldots \cdot p_s^{r_s} > 1$ is relatively prime to $2D$, then,*

$$\alpha_D(n) \leq 2^{1-s} \prod_{i=1}^{s} p_i^{1-r_i} \frac{\gcd(p_i - \epsilon_D(p_i), n - \epsilon_D(n))}{p_i - \epsilon_D(p_i)}$$
$$\leq 2^{1-\Omega(n)} \prod_{i=1}^{s} \frac{\gcd(p_i - \epsilon_D(p_i), n - \epsilon_D(n))}{p_i - \epsilon_D(p_i)}.$$

Proof We see that the identity $\sum_{i=1}^{s}(r_i - 1) = \Omega(n) - s$ trivially holds. Thus,

$$2^{(1-s)} = 2^{1-\Omega(n)+\sum_{i=1}^{s}(r_i-1)} = 2^{1-\Omega(n)} \prod_{i=1}^{s} 2^{r_i-1}.$$

Using the fact that $\frac{2}{p} \leq 1$ for every prime p and $r_i \geq 1$ for all i, the second inequality follows by

$$2^{1-s} \prod_{i=1}^{s} p_i^{1-r_i} = 2^{1-\Omega(n)} \prod_{i=1}^{s} \frac{2^{r_i-1}}{p_i^{r_i-1}} \leq 2^{1-\Omega(n)} \prod_{i=1}^{s} \left(\frac{2}{p_i}\right)^{r_i-1} \leq 2^{1-\Omega(n)}.$$

For the first inequality we use Theorem 5, which implies that for n such that $\gcd(n, 2D) = 1$, we have

$$SL(D, n) \leq \left(1 + \sum_{j=0}^{k_1-1} 2^{js}\right) \prod_{i=1}^{s} \gcd(q, q_i). \tag{4}$$

Using this upper bound and the definition of $\varphi_D(n)$, we get

$$\alpha_D(n) = \frac{SL(D, n)}{\varphi_D(n)} \leq \left(1 + \sum_{j=0}^{k_1-1} 2^{js}\right) \prod_{i=1}^{s} \frac{\gcd(q_i, q)}{p_i^{r_i-1}(p_i - \epsilon_D(p_i))}$$
$$= \left(1 + \sum_{j=0}^{k_1-1} 2^{js}\right) \prod_{i=1}^{s} \frac{\gcd(p_i - \epsilon_D(p_i), q)}{p_i^{r_i-1}(p_i - \epsilon_D(p_i))}.$$

Since in the factorization $n - \epsilon_D(n) = 2^\kappa q$, the two factors 2^κ and q are coprime, we get

$$\prod_{i=1}^{s} \gcd(p_i - \epsilon_D(p_i), n - \epsilon_D(n)) = \prod_{i=1}^{s} \gcd(p_i - \epsilon_D(p_i), q) \gcd(2^{k_i}, 2^\kappa).$$

By Lemma 1, we know that $k_1 \le \kappa$, and by the way we have defined the order of $k_1, k_2, \ldots, k_s$, we get

$$\prod_{i=1}^{s} \gcd(p_i - \epsilon_D(p_i), q) \gcd(2^{k_i}, 2^\kappa) \ge 2^{sk_1} \prod_{i=1}^{s} \gcd(p_i - \epsilon_D(p_i), q).$$

Hence,

$$\prod_{i=1}^{s} \gcd(p_i - \epsilon_D(p_i), q) \le 2^{-sk_1} \prod_{i=1}^{s} \gcd(p_i - \epsilon_D(p_i), n - \epsilon_D(n)).$$

Using Lemma 2, we get

$$\begin{aligned}
\alpha_D(n) = \frac{SL(D,n)}{\varphi_D(n)} &\le \left(1 + \sum_{j=0}^{k_1-1} 2^{js}\right) \prod_{i=1}^{s} \frac{\gcd(p_i - \epsilon_D(p_i), q)}{p_i^{r_i-1}(p_i - \epsilon_D(p_i))} \\
&\le \left(1 + \sum_{j=0}^{k_1-1} 2^{js}\right) 2^{-k_1 s} \prod_{i=1}^{s} \frac{\gcd(p_i - \epsilon_D(p_i), n - \epsilon_D(n))}{p_i^{r_i-1}(p_i - \epsilon_D(p_i))} \\
&\le \left(2^{(k_1-1)s+1}\right) 2^{-k_1 s} \prod_{i=1}^{s} \frac{\gcd(p_i - \epsilon_D(p_i), n - \epsilon_D(n))}{p_i^{r_i-1}(p_i - \epsilon_D(p_i))} \\
&= 2^{1-s} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}} \frac{\gcd(p_i - \epsilon_D(p_i), n - \epsilon_D(n))}{p_i - \epsilon_D(p_i)},
\end{aligned}$$

which proves the assertion.

Lemma 4 *If $t \in \mathbb{R}$ with $t \ge 1$, then*

$$\sum_{n=\lfloor t \rfloor + 1}^{\infty} \frac{1}{n(n-1)} = \frac{1}{\lfloor t \rfloor} < \frac{2}{t}.$$

Proof

$$\sum_{n=\lfloor t \rfloor+1}^{\infty} \frac{1}{n(n-1)} = \lim_{k\to\infty} \sum_{n=\lfloor t \rfloor+1}^{k} \left(\frac{1}{n-1} - \frac{1}{n}\right) = \lim_{k\to\infty} \frac{1}{\lfloor t \rfloor} + \frac{1}{k} = \frac{1}{\lfloor t \rfloor} < \frac{2}{t},$$

where we used the partial fraction decomposition of $\frac{1}{n(n-1)}$ and the fact that $\sum_n \left(\frac{1}{n-1} - \frac{1}{n}\right)$ is a telescopic sum.

The following lemma will also be used frequently.

Lemma 5 *(Damgård, Landrock, Pomerance [7]) For all $k, t, j \in \mathbb{N}$, we have*

$$2\sqrt{tk} - \sqrt{\frac{t}{k-1}} \le 2\sqrt{t(k-1)} \le jt + \frac{k-1}{j}.$$

2.4 A simple estimate

Let us define the following two sets of integers that will be important in our analysis:

Definition 6 For $m, D \in \mathbb{N}$, let

$$C_{m,D} = \{n \in \mathbb{N} : \gcd(n, 2D) = 1, n \text{ composite and } \alpha_D(n) > 2^{-m}\}$$

and let M_k denote the set of odd k-bit integers.

By Theorem 6, we already know that $\alpha_D(n) \leq \frac{1}{4}$. Hence, we have $C_{1,D} = C_{2,D} = \emptyset$. In Theorem 15 we will classify $C_{3,D}$. For $k \geq 2$, the cardinality of M_k is $|M_k| = 2^{k-2}$. We are interested in determining the proportion of odd integers in M_k that also belong to the set $C_{m,D}$.

Theorem 8 *If m, k are positive integers with $m + 1 \leq 2\sqrt{k-1}$, then*

$$\frac{|C_{m,D} \cap M_k|}{|M_k|} < 8 \sum_{j=2}^{m} 2^{m-j-\frac{k-1}{j}}.$$

Proof Lemma 3 with $n \in C_{m,D}$ implies $\Omega(n) \leq m$. Now let $N_D(m,k,j) = \{n \in C_{m,D} \cap M_k : \Omega(n) = j\}$. Thus,

$$|C_{m,D} \cap M_k| = \sum_{j=2}^{m} |N_D(m,k,j)|.$$

Suppose $n \in N_D(m,k,j)$, where $2 \leq j \leq m$. Let p denote the largest prime factor of n. Since $2^{k-1} < n < 2^k$, we have $p > 2^{(k-1)/j}$. Let $d_D(p,n) = \frac{p-\epsilon_D(p)}{\gcd(p-\epsilon_D(p), n-\epsilon_D(n))}$. From Lemma 3 and the definition of $C_{m,D}$, we have

$$2^m > \frac{1}{\alpha_D(n)} \geq 2^{\Omega(n)-1} d_D(p,n) = 2^{j-1} d_D(p,n),$$

so that $d_D(p,n) < 2^{m+1-j}$.

Given p, d, where p is a prime with the property that $p > 2^{(k-1)/j}$ and d is such that $d \mid p - \epsilon_D(p)$ and $d < 2^{m+1-j}$, we want to get an upper bound on how many $n \in N_D(m,k,j)$ exist that have largest prime factor p with $d_D(p,n) = d$. Let

$$S_{D,k,d,p} = \{n \in M_k : p \mid n, d_D(p,n) = d, n \text{ composite}\}.$$

The size of the set $S_{D,k,d,p}$ is at most the number of solutions of the system

$$n \equiv 0 \bmod p,\ n \equiv \epsilon_D(n) \bmod \tfrac{p-\epsilon_D(p)}{d},\ p < n < 2^k,$$

i.e., at most the size of the set

$$R_{D,k,d,p} = \{n \in \mathbb{Z} : n \equiv 0 \bmod p, n \equiv \epsilon_D(n) \bmod \frac{p-\epsilon_D(p)}{d}, p < n < 2^k\},$$

which by the Chinese Remainder Theorem has less than $\frac{2^k d}{p(p-\epsilon_D(p))}$ elements.

If $S_{D,k,d,p} \neq \emptyset$, then there exists an $n \in S_{D,k,d,p}$ with $\gcd(n - \epsilon_D(n), p - \epsilon_D(p)) = (p - \epsilon_D(p))/d$. Now let us look at the parity of $(p - \epsilon_D(p))/d$. Since both p and n are odd,

$(p-\epsilon_D(p))/d = \gcd(p-\epsilon_D(p), n-\epsilon_D(n))$ must be even, thus we only need to consider those p and d that make $(p-\epsilon_D(p))/d$ even. We conclude that

$$\begin{aligned}|N_D(m,k,j)| &\leq \sum_{p>2^{(k-1)/j}} \sum_{\substack{d|p-\epsilon_D(p)\\ d<2^{m+1-j}\\ \frac{p-\epsilon_D(p)}{d}\in 2\mathbb{Z}}} \frac{2^k d}{p(p-\epsilon_D(p))}\\ &= 2^k \sum_{d<2^{m+1-j}} \sum_{\substack{p>2^{(k-1)/j}\\ d|p-\epsilon_D(p)\\ \frac{p-\epsilon_D(p)}{d}\in 2\mathbb{Z}}} \frac{d}{p(p-\epsilon_D(p))}.\end{aligned}$$

Now, for the inner sum we have

$$\begin{aligned}\sum_{\substack{p>2^{(k-1)/j}\\ d|p-\epsilon_D(p)\\ \frac{p-\epsilon_D(p)}{d}\in 2\mathbb{Z}}} \frac{d}{p(p-\epsilon_D(p))} &< \sum_{2ud>2^{\frac{k-1}{j}}-\epsilon_D(p)} \frac{d}{(2ud+\epsilon_D(p))2ud}\\ &= \frac{1}{4d} \sum_{2ud>2^{\frac{k-1}{j}}-\epsilon_D(p)} \frac{1}{(u+\frac{\epsilon_D(p)}{2d})u}\\ &\leq \frac{1}{4d} \sum_{2ud>2^{\frac{k-1}{j}}-\epsilon_D(p)} \frac{1}{u(u-\frac{1}{2d})}\\ &\leq \frac{1}{4d} \sum_{u>\frac{2^{\frac{k-1}{j}}-\epsilon_D(p)}{2d}} \frac{1}{u(u-1)}\\ &< \frac{1}{4d} \frac{2}{\frac{2^{\frac{k-1}{j}}-\epsilon_D(p)}{2d}} = \frac{1}{2^{\frac{k-1}{j}}-\epsilon_D(p)},\end{aligned}$$

where the last inequality follows from Lemma 4. Using this estimate, we get

$$|N_D(m,k,j)| \leq 2^k \sum_{d<2^{m+1-j}} \frac{1}{2^{\frac{k-1}{j}}-\epsilon_D(p)} = 2^k \frac{2^{m+1-j}-1}{2^{\frac{k-1}{j}}-\epsilon_D(p)}.$$

Using Lemma 5 with $t = 1$ and our hypothesis that $m+1 \leq 2\sqrt{k-1}$ yields $m+1 \leq j+(k-1)/j$. Thus,

$$\frac{2^{m+1-j}-1}{2^{\frac{k-1}{j}}-\epsilon_D(p)} \leq \frac{2^{m+1-j}-1}{2^{\frac{k-1}{j}}-1} \leq \frac{2^{m+1-j}}{2^{\frac{k-1}{j}}} = 2^{m-j-\frac{k-1}{j}+1}.$$

Therefore, $N_D(m,k,j)| \leq 2^{k+m-j-\frac{k-1}{j}+1}$. Combining everything and using the fact that $|M_k| = 2^{k-2}$ yields

$$\frac{|C_{m,D}\cap M_k|}{|M_k|} = \frac{\sum_{j=2}^{m}|N_D(m,k,j)|}{2^{k-2}} \leq 8\sum_{j=2}^{m} 2^{m-j-\frac{k-1}{j}}.$$

2.5 The average case error probability

To obtain average case error estimates, we adapt techniques similar to those in [7], with appropriate modifications for the strong Lucas test.

We define X as the event that an integer n declared as *probable prime* by the strong Lucas test is composite, and Z_t as the event that the uniformly at random chosen integer $n \in M_k$ passes t consecutive rounds of the strong Lucas test with uniformly chosen bases (P, Q).

We also use $\pi(x)$ to denote the prime counting function up to x and $\sum'$ to denote the sum over composite integers. Furthermore, let us define the fraction of number of pairs (P, Q) for which the strong Lucas test is positive.

Definition 7 For an odd composite n, let

$$\overline{\alpha}_D(n) = \frac{SL(D, n)}{n - \epsilon_D(n) - 1}.$$

Using the law of conditional probability, we have

$$\begin{aligned} q_{k,t} = \mathbb{P}[X \mid Z_t] = \frac{\mathbb{P}[X \cap Z_t]}{\mathbb{P}[Z_t]} &= \frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)^t}{\sum_{n \in M_k} \overline{\alpha}_D(n)^t} \\ \leq \frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)^t}{\sum_{p \in M_k} \overline{\alpha}_D(p)^t} &= \frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)^t}{\pi(2^k) - \pi(2^{k-1})}, \end{aligned} \tag{5}$$

where p is prime.

To obtain an upper bound for $q_{k,t}$, it will suffice to find an upper bound for the final sum in the numerator of inequality (5) and a lower bound for $\pi(2^k) - \pi(2^{k-1})$. We bound the latter quantity using the following result:

Proposition 1 *(Damgård, Landrock, Pomerance [7]) For an integer* $k \geq 21$, *we have*

$$\pi(2^k) - \pi(2^{k-1}) > (0.71867)\frac{2^k}{k}. \tag{6}$$

In order to proceed, we would like to upper bound the sum

$$\sum_{n \in M_k}{}' \overline{\alpha}_D(n)^t = \sum_{m=2}^{\infty} \sum_{n \in M_k \cap C_{m,D} \setminus C_{m-1,D}} \overline{\alpha}_D(n)^t.$$

However, it is not clear how to bound $\overline{\alpha}_D(n)$ directly. Theorem 8 provides a way to upper bound $|C_{m,D} \cap M_k|$. If we achieve a method to bound $\overline{\alpha}_D(n)$ using $\alpha_D(n)$, we can use the property that for $n \in C_{m,D} \setminus C_{m-1,D}$, we have $2^{-m} < \alpha_D(n) \leq 2^{-(m-1)}$. Thus, our challenge lies in bounding $\overline{\alpha}_D(n)$ using $\alpha_D(n)$. We tackle this problem by establishing two different procedures: for the general case and for a case that involves trial division by small primes before conducting the more computationally expensive strong Lucas test. The latter procedure yields improvements over the general procedure, where the results are comparable to those obtained for the Miller-Rabin test in [7].

Unlike the normal Euler function $\varphi(n)$, the function $\varphi_D(n)$ is not bounded from above by n. Let $n = \prod_{i=1}^{s} p_i^{r_i}$ and let us look at the following trivial upper bound.

$$\varphi_D(n) = \prod_{i=1}^{s} p_i^{r_i - 1}(p_i - \epsilon_D(p_i)) \leq \prod_{i=1}^{s}(p_i^{r_i} + p_i^{r_i - 1}). \tag{7}$$

The question here is whether or not this is an overestimate, that is, whether integers $n = \prod_{i=1}^{n} p_i^{r_i}$ actually exist with $\epsilon_D(p_i) = -1$ for all prime factors p_i of n. The following theorem gives the answer to this:

Theorem 9 *(Ireland [11]) Let D be a non-square integer. Then there exist infinitely many primes p for which D is a quadratic non-residue.*

Theorem 9 directly implies the following corollary:

Corollary 1 *For every non-square integer D, there are infinitely many integers of the form $n = \prod_{i=1}^{s} p_i^{r_i}$ coprime to $2D$ with $\varphi_D(n) = \prod_{i=1}^{s}(p_i^{r_i} + p_i^{r_i-1})$.*

Hence, by Corollary 1, we can construct infinitely many integers that attain the tight upper bound given in inequality (7), proving that it can, in general, not be weakened.

3 Explicit bounds for $q_{k,t}$

In this section, we establish explicit bounds for $q_{k,t}$.

3.1 A bound for $\overline{\alpha}_D(n)$

Corollary 1 demonstrates that in general, $\varphi_D(n)$ is not necessarily bounded by n. Consequently, we cannot conclude that $\overline{\alpha}_D(n) \leq \alpha_D(n)$. However, in order to continue with our analysis, we will need to establish a relationship between the two functions.

Theorem 10 *(Akbary, Friggstad [1])*

$$\frac{n}{\varphi(n)} \leq 1.07e^{\gamma} \ln(\ln(n)) \textit{ for } \ n \geq 2^{78},$$

where γ is the Euler-Mascheroni constant:

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \ln(n) \right) < 0.58.$$

Using this result, we obtain an explicit upper bound for φ_D.

Lemma 6 *For integers $k \geq 78$ and $n \in M_k$, we have*

$$\varphi_D(n) < 2n\ln(k).$$

Proof

$$\varphi_D(n) \leq n \prod_{i=1}^{s} \left(1 + \frac{1}{p_i}\right) < n \prod_{i=1}^{s} \left(1 + \frac{1}{p_i - 1}\right) = \frac{n}{\prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right)}. \tag{8}$$

We realize that $\prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right) = \frac{\varphi(n)}{n}$. Using this in (8), we obtain $\varphi_D(n) \leq n\frac{n}{\varphi(n)}$. By Theorem 10, we have for $k \geq 78$ and $n \in M_k$ that

$$\varphi_D(n) \leq n\frac{n}{\varphi(n)} < n1.07e^{\gamma} \ln(\ln(n)) < 2n\ln(\ln(2^k)) < 2n\ln(k),$$

as $1.07e^{\gamma} < 2$, which proves the claim.

Therefore, we immediately get the following estimate for $\overline{\alpha}_D$:

Corollary 2 *For $k \geq 78$ and $n \in M_k$ we have*

$$\overline{\alpha}_D(n) \leq 2\ln(k)\alpha_D(n).$$

3.2 An intermediate result

Corollary 2 and Theorem 8 allow us to proceed with our analysis.

Proposition 2 *For any integers k, M, t with $3 \leq M \leq 2\sqrt{k-1}-1$, $t \geq 1$ and $k \geq 78$, we have*

$$\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t \leq 2^{k-2+t(1-M)}\ln^t(k) + 2^{k+1+2t}\ln^t(k)\sum_{j=2}^{M}\sum_{m=j}^{M}2^{m(1-t)-j-\frac{k-1}{j}}.$$

Proof Note that our hypothesis implies $k \geq 5$. We know that $C_{1,D} \cap M_k = \emptyset$. Thus, by Corollary 2, we have

$$\begin{aligned}\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t &= \sum_{m=2}^{\infty}\sum_{n\in M_k\cap C_{m,D}\setminus C_{m-1,D}}\overline{\alpha}_D(n)^t\\ &\leq \sum_{m=2}^{\infty}\sum_{n\in M_k\cap C_{m,D}\setminus C_{m-1,D}}\big(2\ln(k)\alpha_D(n)\big)^t.\end{aligned}$$

Since $n \in C_{m,D} \setminus C_{m-1,D}$, we have that $2^{-m} < \alpha_D(n) \leq 2^{-(m-1)}$. Hence, we get

$$\begin{aligned}\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t &< \ln^t(k)\sum_{m=2}^{\infty}2^{t-(m-1)t}|M_k \cap C_{m,D} \setminus C_{m-1,D}|\\ &\leq \ln^t(k)\Big(\sum_{m=M+1}^{\infty}2^{t(2-m)}|M_k \setminus C_{M,D}| + \sum_{m=2}^{M}2^{(2-m)t}|M_k \cap C_{m,D}|\Big)\\ &= \ln^t(k)\Big(\frac{2^{2t-Mt}}{2^t-1}|M_k \setminus C_{M,D}| + \sum_{m=2}^{M}2^{(2-m)t}|M_k \cap C_{m,D}|\Big).\end{aligned}$$

For $t \geq 1$, we have that $2^t - 2 \geq 2^{t-1}$, so $\frac{2^{2t-Mt}}{2^t-1} \leq 2^{t(1-M)+1}$. Moreover, we use Theorem 8 and get

$$\begin{aligned}\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t &\leq \ln^t(k)\Big(2^{k-1+t(1-M)} + 2^{k+1+2t}\sum_{m=2}^{M}\sum_{j=2}^{m}2^{m(1-t)-j-\frac{k-1}{j}}\Big)\\ &= \ln^t(k)\Big(2^{k-1+t(1-M)} + 2^{k+1+2t}\sum_{j=2}^{M}\sum_{m=j}^{M}2^{m(1-t)-j-\frac{k-1}{j}}\Big).\end{aligned}$$

3.3 An estimate for $q_{k,1}$

We now derive the first numerical upper bound for $q_{k,t}$ when $t = 1$.

Theorem 11 *For $k \geq 2$, we have $q_{k,1} < \ln(k)k^2 4^{2.3-\sqrt{k}}$.*

Proof We use Proposition 2 with $t = 1$ and $k \geq 78$ and let M be an integer with $3 \leq M \leq 2\sqrt{k-1} - 1$ and get

$$\begin{aligned}\sum{}'_{n\in M_k}\overline{\alpha}_D(n) &\leq \ln(k)\Big(2^{k-M} + 2^{k+3}\sum_{j=2}^{M}(M+1-j)2^{-j-\frac{k-1}{j}}\Big)\\ &\leq \ln(k)\Big(2^{k-M} + 2^{k+3-2\sqrt{k-1}}\sum_{j=2}^{M}(M+1-j)\Big), \qquad (9)\end{aligned}$$

where we used Lemma 5 to bound $2^{-jt-\frac{k-1}{j}}$. We bound the sum $\sum_{j=2}^{M}(M+1-j) = M(M-1)/2$ and let $M = \lfloor 2\sqrt{k-1} - 1\rfloor$, which yields

$$\begin{aligned}\sum{}'_{n\in M_k}\overline{\alpha}_D(n) &\leq \ln(k)\Big(2^{k-M} + 2^{k+2-2\sqrt{k-1}}M(M-1)\Big)\\ &< \ln(k)\Big(2^{k+2-2\sqrt{k-1}}(1 + (4(k-1) - 6\sqrt{k-1} + 2)\Big)\\ &< \ln(k)k2^{k+4-2\sqrt{k-1}}. \qquad (10)\end{aligned}$$

We again use Lemma 5 with $t = 1$ in inequality (10) for $k \geq 100$ and get

$$\sum{}'_{n\in M_k}\overline{\alpha}_D(n) < \ln(k)k2^{4+\frac{1}{\sqrt{99}}+k-2\sqrt{k}}. \qquad (11)$$

As $\frac{2^{4+\frac{1}{\sqrt{99}}}}{0.71867} < 4^{2.3}$ we get by Proposition 1 and inequalities (11) and (5) for $k \geq 100$ that

$$q_{k,1} \leq \frac{\sum{}'_{n\in M_k}\overline{\alpha}_D(n)}{\pi(2^k) - \pi(2^{k-1})} = \frac{\ln(k)k^2 \cdot 2^{4+\frac{1}{\sqrt{99}}-2\sqrt{k}}}{0.71867} < \ln(k)k^2 4^{2.3-\sqrt{k}}.$$

But for $k \leq 101$ we have that $\ln(k)k^2 4^{2.3-\sqrt{k}} > 1$, so this upper bound is trivially true for $k \leq 101$.

3.4 An estimate for $q_{k,t}$

We now consider average case error estimates for an integer that has passed t consecutive rounds of the strong Lucas test with respect to randomly chosen bases and obtain numerical bounds for $q_{k,t}$ when $t \geq 2$.

Theorem 12 *For integers k, t with $k \geq 78$, $3 \leq t \leq k/9$ or $k \geq 88$, $t = 2$ we have*

$$q_{k,t} < \ln^t(k)\frac{k^{3/2}}{\sqrt{t}}4^{2.12+t-\sqrt{tk}}.$$

Proof Assume $k \geq 78$ and $t \geq 2$. Let us first estimate $\sum_{m=j}^{M} 2^{m(1-t)}$. We have

$$\sum_{m=j}^{M} 2^{m(1-t)} = \frac{2^{j(1-t)+t} - 2^{M(1-t)+1}}{2^t - 2} < \frac{2^{j(1-t)+t}}{2^t - 2} = \frac{2^{j(1-t)}}{1 - 2^{1-t}},$$

as $j \leq M$. Using this estimate in Proposition 2, we get that

$$\sum{}'_{n\in M_k} \overline{\alpha}_D(n)^t \leq 2^{k-1+t(1-M)} \ln^t(k) + \frac{2^{k+1+2t}}{1-2^{1-t}} \ln^t(k) \sum_{j=2}^{M} 2^{-jt-\frac{k-1}{j}}, \tag{12}$$

for any integer M with $3 \leq M \leq 2\sqrt{k-1}-1$. By Lemma 5, we have that

$$jt + \frac{k-1}{j} \geq 2\sqrt{t(k-1)} \quad \forall j, k > 0.$$

Furthermore, we choose $M = \left\lceil 2\sqrt{(k-1)/t}+1 \right\rceil$. In order to use Proposition 2, we need to make sure that $3 \leq M \leq 2\sqrt{k-1}-1$. Thus, for $3 \leq M$ to hold, we must restrict $t \leq k-1$ for $k > 1$. For $k \geq 25$, we have $M \leq 2\sqrt{k-1}-1$.

Our choice of M implies that $M-1 < 2\sqrt{(k-1)/t}+1 < 2\sqrt{k/t}+1$. Since $t \leq k-1$, we have $1 \leq \sqrt{k/t}$, which yields $M-1 < 2^{1.6}\sqrt{k/t}$. We also see that $M-1 \geq 2\sqrt{(k-1)/t}$.

Using our chosen value for M and the inequalities established above in (12), we get

$$\begin{aligned}\sum{}'_{n\in M_k} \overline{\alpha}_D(n)^t &\leq 2^{k-1+t(1-M)} \ln^t(k) + \frac{2^{k+1+2t}}{1-2^{1-t}} \ln^t(k)(M-1)2^{-2\sqrt{t(k-1)}} \\ &< 2^{k-1-2\sqrt{t(k-1)}} \ln^t(k) + \frac{2^{k+2.6+2t}}{1-2^{1-t}} \ln^t(k)\sqrt{\frac{k}{t}}2^{-2\sqrt{t(k-1)}} \\ &= 2^{k-1-2\sqrt{t(k-1)}} \ln^t(k)\left(1 + 2^{3.6}\frac{2^{2t}}{1-2^{1-t}}\sqrt{\frac{k}{t}}\right).\end{aligned}$$

The function $f(k,t) = \frac{2^{2t}}{1-2^{1-t}}\sqrt{\frac{k}{t}}$ is a monotonically increasing function for all $t \geq 2$ and for all $k \geq 1$. Thus, we get with $k \geq 78$ and $t \geq 2$

$$2^{3.6}\frac{2^{2t}}{1-2^{1-t}}\sqrt{\frac{k}{t}} \geq 2^{3.6}\frac{2^4}{1-2^{-1}}\sqrt{\frac{78}{2}} > 2438.$$

For $x > 2423$ we have $1 + x = x\left(\frac{1}{x}+1\right) < x\frac{2424}{2423}$, which yields

$$\sum{}'_{n\in M_k} \overline{\alpha}_D(n)^t < 2^{k-1-2\sqrt{t(k-1)}} \ln^t(k)\frac{2424}{2423}2^{3.6}\frac{2^{2t}}{1-2^{1-t}}\sqrt{\frac{k}{t}}. \tag{13}$$

With Lemma 5 we upper bound $2^{-2\sqrt{t(k-1)}}$. For $t = 2$ and $k \geq 88$, and using the fact that $2^{1+\sqrt{\frac{2}{k-1}}}$ is a monotonically decreasing function for all $k \geq 1$, we have

$$\frac{2^{\sqrt{\frac{t}{k-1}}}}{1-2^{1-t}} = \frac{2^{\sqrt{\frac{2}{k-1}}}}{1-2^{1-2}} = 2^{1+\sqrt{\frac{2}{k-1}}} < 2.222.$$

For $3 \leq t \leq k/9$, we have

$$\frac{2^{\sqrt{\frac{t}{k-1}}}}{1-2^{1-t}} \leq \frac{4}{3}2^{\frac{1}{3}} < 1.7.$$

In any case, we have $\frac{2\sqrt{\frac{t}{k-1}}}{1-2^{1-t}} < 2.222$. Putting these estimates in (13), we get

$$\begin{aligned}\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t &< 2^{k-2\sqrt{tk}+2t}\ln^t(k)\frac{2439}{2438}2^{2.6}\frac{2\sqrt{\frac{t}{k-1}}}{1-2^{1-t}}\sqrt{\frac{k}{t}}\\ &< 2^{k-2\sqrt{tk}+2t}\ln^t(k)\frac{2424}{2423}2^{2.6}2.222\sqrt{\frac{k}{t}}.\end{aligned}$$

for all $3 \le t \le k/9$, $k \ge 79$ and for $t = 2$, $k \ge 88$. Now using Proposition 1 and inequality (5), we get the desired result.

The bounds obtained in Theorems 11 and 12 already show that for most choices of k and t, the average-case error estimates for the strong Lucas test are small enough to be used in practice. Yet there is still room for improvement, as, for example, Theorem 11 would give us the upper bound $q_{k,1} > 1$ for all $k \le 79$. Moreover, even though the bounds in Theorem 12 are always less than 1, especially for small choices of $t = 2, 3, 4$ and small choices of k, the bounds are not as good as we would expect them to be. In the next section, we will show even better bounds by working only with integers that are divisible by small primes.

4 Improved average-case error estimates

The probability of a random integer having a small prime divisor is comparatively large, hence, it seems obvious to test the candidate for small divisors less than a given precomputed bound B. The proportion of odd candidates eliminated through trial division is $\prod_{3\le p\le B}\left(1-\frac{1}{p}\right)$. This is, by Merten's theorem, approximately equal to $\frac{1.12}{\ln(B)}$, where p ranges over prime values. Further details and the proof can be found in [10]. For example, let $B = 256$, then 80% of odd candidates are discarded before a more costly primality test is performed. It is worth noting that the OpenSSL implementation incorporates a similar idea of dividing by small primes before calling the Miller-Rabin test. This is done to quickly eliminate numbers that are obviously not prime, reducing the need for the more computationally intensive Miller-Rabin test and thereby speeding up the process of prime number generation. In practice, this additional subroutine does not typically introduce significant extra running time.

Remark 1 The function used for primality testing in OpenSSL Version 3.1.2[1] is `BN_is_prime_fasttest_ex`, which is located in the file `bn_prime.c`. Within this function, we call the function `calc_trial_divisions`, which calculates the number of trial divisions for achieving the best speed in combination with the Miller-Rabin test, based on the bit-length denoted by k. It is worth noting that in the file, the variable l represents the l-th prime. However, in the notation used in this paper, l signifies the l-th odd prime. Therefore, the numbering below differs from that in the file.

```
static int calc_trial_divisions(int k)
{
if (k <= 512)
return l = 63;
else if (k<=1024)
return l= 127;
```

[1] See https://github.com/openssl/openssl/blob/master/crypto/bn/bn_prime.c

```
else if (k<=2048)
return l = 383;
else if  (k<= 4096)
return l = 1023;
}
```

After this, the function that invokes the MR testing with pseudo-random bases is called. The number of rounds also depends on the size of k.

Two of the authors of [7], Brandt and Damgård, stated in a different paper [5] that it seems like a difficult problem to analyze the error probability of the Miller-Rabin test that includes trial division. Yet, it is clear that the error probability will be at most that of the initial algorithm as trial division can never reject a prime and thus only give us a better chance of rejecting composites.

For the strong Lucas test, in contrast, we can take advantage of this modified procedure. This permits us to establish a new bound for $\alpha_D(n)$, which eventually results in average-case error estimates that yield tighter estimates than the ones on $q_{k,t}$ established in Section 3. For this, let us consider a modified version of the primality testing algorithm. In this approach, we first perform trial division by the first l odd primes. If the integer is divisible by any of these primes, another candidate from M_k is randomly selected. Following this preliminary step, we apply t independent rounds of the strong Lucas test, as described previously.

4.1 An improved bound for $\overline{\alpha}_D(n)$

Let us define the following quantities:

Definition 8 Let $l, t, k \in \mathbb{N}$. We define $\tilde{p}_l$ to be the l-th odd prime and

$$\rho_l = 1 + \frac{1}{\tilde{p}_{l+1}}.$$

Moreover, we define $M_{k,l}$ to be the set of odd k-bit integers that are not divisible by the first odd l primes.

Given the introduction of the modified primality testing algorithm, we need to redefine our error probability.

Definition 9 We let $q_{k,l,t}$ denote the probability that a composite integer chosen uniformly at random from $M_{k,l}$ passes t consecutive rounds of the strong Lucas test with randomly chosen bases (P, Q).

The following two lemmas will be important in our analysis:

Lemma 7 *Let $n, l \in \mathbb{N}$ and let n be relatively prime to $2D$ and not divisible by all of the first l odd primes. Let $\omega(n)$ be as defined in Definition 5. Then,*

$$\varphi_D(n) \leq \rho_l^{\omega(n)} n,$$

which implies

$$\overline{\alpha}_D(n) \leq \rho_l^{\omega(n)} \alpha_D(n).$$

Proof For $n_1, n_2 \in \mathbb{N}$ with $\gcd(n_1, n_2) = 1$, we have the relation

$$\varphi_D(n_1, n_2) = \varphi_D(n_1)\varphi_D(n_2).$$

It is thus sufficient to only treat the case $n = p^r$. We have

$$\frac{\varphi_D(p^r)}{p^r} = \frac{p^{r-1}(p - \epsilon_D(p))}{p^r} = 1 - \frac{\epsilon_D(p)}{p} \leq 1 + \frac{1}{p}.$$

With $p \geq \tilde{p}_{l+1}$, we have $\frac{\varphi_D(p^r)}{p^r} \leq 1 + \frac{1}{\tilde{p}_{l+1}} = \rho_l$ and the result follows directly.

Lemma 8 *Let $n \in C_{m,D}$. Then,*

$$\omega(n) \leq m.$$

Proof The result directly follows from Lemma 3, which states

$$\alpha_D(n) \leq 2^{1-\omega(n)} \prod_{i=1}^{\omega(n)} p^{1-r_i} \frac{\gcd(p - \epsilon_D(p), n - \epsilon_D(n))}{p - \epsilon_D(p)} \leq 2^{1-\omega(n)}.$$

We can now give a bound similar to Proposition 2 for $\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t$.

Proposition 3 *For any integers k, t, M, l with $3 \leq M \leq 2\sqrt{k-1} - 1$, we have*

$$\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t \leq 2^{k-2+t} \sum_{m=M+1}^{\infty} \rho_l^{mt} 2^{-mt} + 2^{k+1+t} \sum_{m=2}^{M} \sum_{j=2}^{m} \rho_l^{mt} 2^{m(1-t)-j-\frac{k-1}{j}}.$$

Proof The theorem follows by closely following the proof of Proposition 2 while using Lemmas 7 and 8:

$$\begin{aligned}
\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t &= \sum_{m=2}^{\infty} \sum_{n \in M_k \cap C_{m,D} \setminus C_{m-1,D}} \overline{\alpha}_D(n)^t \\
&\leq \sum_{m=2}^{\infty} \sum_{n \in M_k \cap C_{m,D} \setminus C_{m-1,D}} \rho_l^{mt} 2^{-(m-1)t} \\
&\leq \sum_{m=M+1}^{\infty} \rho_l^{mt} 2^{-(m-1)t} \mid M_k \mid + \sum_{m=2}^{M} \rho_l^{mt} 2^{-(m-1)t} \mid M_k \cap C_{m,D} \mid \qquad (14) \\
&\leq 2^{k-2+t} \sum_{m=M+1}^{\infty} \rho_l^{mt} 2^{-mt} + 2^{k+1+t} \sum_{m=2}^{M} \sum_{j=2}^{m} \rho_l^{mt} 2^{m(1-t)-j-\frac{k-1}{j}}.
\end{aligned}$$

With this new bound for $\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t$ we are able to get rid of the factor $\ln(k)^t$ completely, instead, we have some power of ρ_l. If l is chosen as discussed in Remark 1, ρ_l will be very close to 1.

Let us adapt inequality (5) accordingly, which gives us

$$q_{k,l,t} = \frac{\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t}{\pi(2^k) - \pi(2^{k-1})}. \qquad (15)$$

4.2 An estimate for $q_{k,l,1}$

In this subsection we let $t = 1$. We need the following proposition to establish a new estimate:

Proposition 4 *For any integers k, l, M with $3 \le M \le 2\sqrt{k-1} - 1$, we have*

$$\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n) \le 2^{k-1-M}{\rho_l}^{M+1} + 2^{k-2\sqrt{k-1}+1}\rho_l^M M(M-1).$$

Proof Evaluating the first part of the sum in Proposition 3 with $t = 1$ yields $\sum_{m=M+1}^{\infty} \rho_l^m 2^{-m} = \frac{2^{-M}{\rho_l}^{M+1}}{2-\rho_l} \le 2^{-M}{\rho_l}^{M+1}$.

For the second part of the sum, we use Lemma 5 with $t = 1$ and the condition $m \le M$, and conclude that

$$2^{k+2}\sum_{m=2}^{M}\sum_{j=2}^{m}\rho_l^m 2^{-j-\frac{k-1}{j}} \le 2^{k-2\sqrt{k-1}}\rho_l^M \sum_{j=2}^{M}\sum_{m=j}^{M} 1 = 2^{k+1-2\sqrt{k-1}}\rho_l^M M(M-1).$$

Theorem 13 *For integers k, l with $k \ge 2$, we have*

$$q_{k,l,1} < k^2 4^{1.8-\sqrt{k}} \rho_l^{2\sqrt{k-1}-2}.$$

Proof Using inequality (15) and Proposition 4 with $M = \lfloor 2\sqrt{k-1} - 2 \rfloor$, we get

$$q_{k,l,1} = \frac{\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t}{\pi(2^k) - \pi(2^{k-1})} \le k^2 4^{1.73-\sqrt{k-1}} \rho_l^{2\sqrt{k-1}-1}. \tag{16}$$

With Lemma 5 we have for $k \ge 53$ that

$$4^{-\sqrt{k-1}} < 4^{\frac{1}{4\sqrt{13}}-\sqrt{k}} < 4^{0.07-\sqrt{k}}.$$

Therefore, we get

$$q_{k,l,1} < k^2 4^{1.8-\sqrt{k}} \rho_l^{2\sqrt{k-1}-2},$$

which proves the theorem for $k \ge 53$. The theorem is trivially true for $k \le 52$, as $k^2 4^{1.8-\sqrt{k}} \rho_l^{2\sqrt{k-1}-2} \ge k^2 4^{1.8-\sqrt{k}} \ge 1$ for $k \ge 52$ and $l \ge 1$, since $\rho_l \ge 1$.

Let us examine the bound for $q_{k,l,1}$ in Theorem 13 in more detail. When the l-th prime is sufficiently large, ρ_l is approximately equal to 1. For example, when $k = 1024$ and $l = 127$, we have $\rho_l^{2\sqrt{k-1}-1} < 1.09$. Thus,

$$q_{k,l,1} < k^2 4^{1.8-\sqrt{k}} \rho_l^{2\sqrt{k-1}-1} \approx k^2 4^{1.87-\sqrt{k}}.$$

The following corollary provides an explicit bound for k-bit integers that are not divisible by the first 127 odd primes and successfully pass a single round of the strong Lucas test. The first condition, as pointed out in Remark 1, is ensured whenever $k > 512$ before applying the Miller-Rabin test.

Corollary 3 *Let n be an odd integer that is not divisible by the first 127 odd primes. Then, for all $k \ge 2$, we have $q_{k,127,1} < k^2 4^{1.729-0.998\sqrt{k-1}}$.*

Proof By (16) we have that

$$q_{k,127,1} \le k^2 4^{-\sqrt{k-1}+1.73} \left(\frac{728}{727}\right)^{(2\sqrt{k-1}-1)}.$$

With $\left(\frac{728}{727}\right)^{(2\sqrt{k-1}-1)} \le 4^{(2\sqrt{k-1}-1)0.001}$ we get

$$q_{k,127,1} \le k^2 4^{1.73-\sqrt{k-1}+0.001(2\sqrt{k-1}-1)} \le k^2 4^{1.729-0.998\sqrt{k-1}},$$

which proves the corollary.

Table 1 Comparing the lower bound probabilities for $-\log_2(prob)$, where $prob = p_{k,1}, q_{k,1}, q_{k,l,1}$, and l was chosen with respect to k as discussed in Remark 1

k	$-\log_2 p_{k,1}$	$-\log_2 q_{k,1}$	$-\log_2 q_{k,l,1}$
100	2	-1	3
200	8	5	9
400	18	15	18
512	23	20	23
1024	40	36	40
2048	64	60	64
4096	100	96	100

In Table 1 we compare the bounds of $p_{k,1}$ for the Miller-Rabin test, as stated in Theorem 4 in [7], with the bounds of $q_{k,1}$ and $q_{k,l,1}$ for the strong Lucas test as established in Theorem 11 and Theorem 13 respectively. For the bounds of $q_{k,l,1}$, we select the values of *l* with respect to k as defined in the `calc_trial_division` function as discussed in Remark 1.

4.3 An estimate for $q_{k,l,t}$

In this section, we establish bounds for $q_{k,l,t}$ for various choices of k, l and $t \geq 2$.

Corollary 4 *Let* $\rho_l = 1 + \frac{1}{\tilde{p}_{l+1}}$. *Then*

$$2^t - \rho_l^t \geq \frac{1}{2}\rho_l^t.$$

Proof With $\tilde{p}_{l+1} \geq 2$, it follows that $\rho_l \leq \frac{4}{3} < 2$. This implies that $2^t - \rho_l^t \geq \rho_l^t\left(\frac{2}{\rho_l} - 1\right) \geq \frac{1}{2}\rho_l^t$.

We are now ready to prove the bound for $q_{k,l,t}$.

Theorem 14 *For any integers* $2 \leq t \leq (k-1)/9$, $k \geq 21$, $l \in \mathbb{N}$ *we have*

$$q_{k,l,t} \leq 4^{1.72-\sqrt{tk}} k^{3/2} 2^t \rho_l^{2\sqrt{kt}+t}.$$

Proof By Proposition 3, we know that

$$\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t \leq 2^{k-2+t} \sum_{m=M+1}^{\infty} \rho_l^{mt} 2^{-mt} + 2^{k+1+t} \sum_{j=2}^{M} \sum_{m=j}^{M} \rho_l^{mt} 2^{m(1-t)-j-\frac{k-1}{j}}. \tag{17}$$

for any integer $2 \leq M \leq 2\sqrt{k-1} - 1$. Let us first look at the left-hand side of the sum of (17). Using Corollary 4, we get that

$$2^{k-2+t} \sum_{m=M+1}^{\infty} \rho_l^{mt} 2^{-mt} = 2^{k-2+t} \frac{2^{-Mt}\rho_l^{t(M-1)}}{2^t - \rho_l^t} \leq 2^{k-1-t(M-1)} \rho_l^{t(M-2)}. \tag{18}$$

Now let us look at the right-hand side of the sum of (17). Using $\sum_{m=j}^{M} 2^{m(1-t)} < \frac{2^{j(1-t)+t}}{2^t-2}$ and $m \leq M$, we obtain

$$2^{k+1+t} \sum_{j=2}^{M} \sum_{m=j}^{M} \rho_l^{mt} 2^{m(1-t)-j-\frac{k-1}{j}} \leq \frac{2^{k+1+2t}\rho_l^{Mt}}{2^t - 2} \sum_{j=2}^{M} 2^{-jt-\frac{k-1}{j}}. \tag{19}$$

Table 2 Lower bounds for $-\log_2(q_{k,t})$ using Theorem 12

k/ t	2	4	8	16	32	64
100	6	9	10			
200	15	24	30	28		
400	30	45	60	71	64	
512	37	55	74	91	92	
1024	62	90	124	162	191	188
2048	97	141	197	264	335	390
4096	149	214	300	410	542	681

Further, we let $M = \left\lceil 2\sqrt{(k-1)/t} \right\rceil$. To have $M \ge 3$, we must restrict t to $t \le k-1$. Also, for $k \ge 9$, we have $M = \left\lceil 2\sqrt{(k-1)/t} \right\rceil \le \left\lceil 2\sqrt{(k-1)/2} \right\rceil \le 2\sqrt{k-1} - 1$. From (17), using the inequalities (18) and (19) and Lemma 5, we get

$$\begin{aligned}
\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t &\le 2^{k-1-t(M-1)} \rho_l^{t(M-2)} + \frac{2^{k+1+2t-2\sqrt{t(k-1)}}}{2^t-2} \rho_l^{Mt}(M-1) \\
&\le 2^{k-1+t-2\sqrt{t(k-1)}} \rho_l^{2\sqrt{t(k-1)}-t} \\
&\quad + \frac{2^{k+2+2t-2\sqrt{t(k-1)}}}{2^t-2} \sqrt{\frac{k}{t}} \rho_l^{2\sqrt{t(k-1)}+t} \\
&= 2^{k-1+t-2\sqrt{t(k-1)}} \rho_l^{2\sqrt{t(k-1)}+t} \left(\rho_l^{-2t} + \sqrt{\frac{k}{t}} \frac{2^{3+t}}{2^t-2} \right) \\
&< 2^{k-1+t-2\sqrt{t(k-1)}} \rho_l^{2\sqrt{t(k-1)}+t} \left(1 + \sqrt{\frac{k}{t}} \frac{2^{3+t}}{2^t-2} \right).
\end{aligned} \tag{20}$$

As $\frac{2^t}{\sqrt{t}(2^t-2)}$ is monotonically decreasing in $t \ge 2$, we have for $t \ge 2$

$$\frac{2^{3+t}}{2^t-2}\sqrt{\frac{k}{t}} < \frac{2^5}{2}\sqrt{\frac{k}{2}} = 4^{1.75}\sqrt{k}.$$

For $k \ge 1$, we have $1 + 4^{1.75}\sqrt{k} < 4^{1.812}\sqrt{k}$. For $t \le (k-1)/9$, we get by Lemma 5 that

$$2^{\sqrt{t/(k-1)}} \le 2^{\sqrt{1/9}} = 1.25992 < 1.26.$$

Thus, we get from (20)

$$\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t \le 2^{k+t} \rho_l^{2\sqrt{kt}+t} 4^{1.312-\sqrt{tk}} (1.26)\sqrt{k}.$$

By the same argument as in inequality (5), we get that

$$q_{k,l,t} \le \frac{\sum'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t}{\pi(2^k) - \pi(2^{k-1})}. \tag{21}$$

Using the bound we have obtained for $\sum'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t$ and Proposition 1 in (21), we get the desired result.

Tables 2, 3 and 4 compare the bounds for $q_{k,t}$ (Theorem 12), $q_{k,l,t}$ (Theorem 14) and $p_{k,t}$ (Theorem 4), where l was chosen with respect to k as discussed in Remark 1.

Table 3 Lower bounds for $-\log_2(q_{k,l,t})$ using Theorem 14 with l chosen with respect to k defined as in Remark 1

k/ t	2	4	8	16	32	64
100	12	22	34			
200	22	37	56	81		
400	37	59	88	126	176	
512	44	69	102	147	205	
1024	69	105	154	221	310	428
2048	105	156	227	325	459	639
4096	157	230	332	474	670	938

Table 4 Lower bounds for $-\log_2(p_{k,t})$ using Theorem 4 4

k/ t	2	4	8	16	32	64
100	12	23	36			
200	23	38	58	83		
400	38	60	89	129	179	
512	45	70	104	149	209	
1024	70	106	155	223	313	432
2048	106	157	229	327	462	642
4096	157	231	333	476	672	941

5 The worst-case numbers

The dominant contributors to our probability estimate are the numbers with the largest value for $\alpha_D(n)$, as highlighted in our analysis. According to Definition 6. These are the sets $C_{m,D}$ characterized by a small value for m. Notably, the sets $C_{1,D}$ and $C_{2,D}$ are empty due to Theorem 6, which states that $\alpha_D(n) \leq 1/4$ for all $n \in \mathbb{N}$. In our pursuit of refining estimates for $q_{k,l,t}$ for larger values of t, we opt to treat the sets $C_{3,D}$, $C_{4,D}$, and $C_{5,D}$ separately.

However, this strategy encounters a complication due to the presence of a particular subset of Lucas–Carmichael numbers within this category. Lucas–Carmichael numbers can be seen as a generalization of Carmichael numbers. Unfortunately, bounding the number of Lucas-Carmichael numbers remains an unsolved problem.

Yet, if we assume that this specific subset of composites with three prime factors has the property that $\epsilon_D(n) = \epsilon_D(p_i)$ for $i = 1, 2, 3$, we can establish an upper bound on the size of this set. It is important to emphasize that this requirement is needed for this specific group of composites. Building upon this assumption, we are able to deduce upper bounds for $q_{k,l,t}$ that exhibit a better performance for larger values of t. It is worth noting that the bounds we obtain in Subsections 5.3 and 5.4 are exclusively valid under this particular assumption. We provide reasons for the plausibility of this assumption whenever an integer passes many rounds of both the (strong) Lucas test and the (strong) probable prime test. For the remaining subsections we do not need this assumption.

The task of establishing bounds for the size of the set of arbitrary Lucas–Carmichael numbers with three prime factors remains unsolved. If such bounds could be determined, it would straightforwardly lead to the derivation of expressions for $q_{k,l,t}$ that are good for large values of t, independent of the assumption regarding $\epsilon_D(p)$ for the prime factors p of the integer.

In this section we always assume that $\epsilon_D(n) = -1$.

5.1 Classifying $C_{3,D}$

First, we classify the members of $C_{3,D}$. In this subsection, unless specified otherwise, let n always denote an integer relatively prime to $2D$ with prime decomposition $n = p_1^{r_1} \cdot \ldots \cdot p_s^{r_s}$, and write $n - \epsilon_D(n) = 2^{\kappa} q$ and $p_i - \epsilon_D(p_i) = 2^{k_i} q_i$, where q, q_i are odd, and the prime factors p_i are ordered such that $k_1 \leq \ldots \leq k_s$.

We will later make use of the following lemmas in our proofs.

Lemma 9 *(Arnault [3]) Let n be as described above. Then,*

$$\frac{SL(D,n)}{\varphi_D(n)} \leq \begin{cases} \frac{1}{2^{s-1}} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{q_i}, \\ \frac{1}{2^{s-1}} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}}, \\ \frac{1}{2^{s-1+\delta_2+\ldots+\delta_s}}, \ \textit{where } \delta_i = k_i - k_1. \end{cases} \tag{22}$$

Lemma 10 *Let n be as described above. Then,*

$$\frac{SL(D,n))}{\varphi_D(n)} \leq 2^{1-s+\sum_{i=1}^{s}(k_1-k_i)} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{q_i}.$$

Proof From inequality (4) and Lemma 2 we get that

$$SL(D,n) \leq 2^{1+(k_1-1)s} \prod_{i=1}^{s} \gcd(q,q_i).$$

We also have that

$$\varphi_D(n) = \prod_{i=1}^{s} p_i^{r_i-1}(p_i - \epsilon_D(p_i)) \geq \prod_{i=1}^{s}(p_i - \epsilon_D(p_i)) = \prod_{i=0}^{s} 2^{k_i} q_i.$$

We combine these expressions and see that

$$2^{1+(k_1-1)s} \prod_{i=1}^{s} 2^{-k_i} = 2^{1-s+\sum_{i=1}^{s}(k_1-k_i)},$$

to get the desired result.

Lemma 11 *Let n be as described above. Then,*

$$2^k q = 2^{2k_1+\delta_2} q_1 q_2 \pm 2^{k_1}(q_1 \pm 2^{\delta_2} q_2).$$

Proof

$$\begin{aligned} 2^k q =& p_1 p_2 - \epsilon_D(p_1 p_2) = (2^{k_1} q_1 + \epsilon_D(p_1))(2^{k_1+\delta_2} q_2 + \epsilon_D(p_2)) - \epsilon_D(p_1 p_2) \\ =& 2^{2k_1+\delta_2} q_1 q_2 + 2^{k_1} q_1 \epsilon_D(p_2) + 2^{k_1+\delta_2} q_2 \epsilon_D(p_1) + \epsilon_D(p_1)\epsilon_D(p_2) - \epsilon_D(p_1 p_2) \\ =& 2^{2k_1+\delta_2} q_1 q_2 + 2^{k_1}(q_1 \epsilon_D(p_2) + 2^{\delta_2} q_2 \epsilon_D(p_1)) \\ =& 2^{2k_1+\delta_2} q_1 q_2 \pm 2^{k_1}(q_1 \pm 2^{\delta_2} q_2). \end{aligned}$$

Lemma 12

$$\frac{SL(D,n)}{\varphi_D(n)} = \frac{1}{2^{k_1+k_2+\cdots+k_s}} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}} \left(\prod_{i=1}^{s} \frac{\gcd(q,q_i)-1}{q_i} + \frac{2^{sk_1}-1}{2^s-1} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{q_i} \right).$$

Proof We have

$$\varphi_D(n)=\prod_{i=1}^{s}\varphi_D(p_i^{r_i})=\prod_{i=1}^{s}p_i^{r_i-1}(2^{k_i}q_i)=2^{k_1+k_2+\cdots+k_s}\prod_{i=1}^{s}q_i\prod_{i=1}^{s}p_i^{r_i-1}.$$

Together with Theorem 5 and $\sum_{j=0}^{k_1-1}2^{js}=\frac{2^{sk_1}-1}{2^s-1}$ we get the desired result.

Lemma 13 *(Arnault [3]) Let* $n=(2^{k_1}q_1-1)(2^{k_1}q_1+1)$. *Then, for all* $q_1,k_1\in\mathbb{N}$ *with* $q_1\neq 1$ *odd, we have* $\frac{SL(D,n)}{\varphi_D(n)}>\frac{1}{3}$. *For* $q_1=1$, *we have* $\frac{SL(D,n)}{\varphi_D(n)}=\frac{1}{3}-\frac{1}{3\cdot 4^{k_1}}$.

Now we have all the ingredients to prove the main theorem of this section. For integers m,n,β, we mean by $m^\beta\ ||\ n$ that $m^\beta\mid n$ and $m^{\beta+1}\nmid n$.

Theorem 15 *Let* $n=p_1^{r_1}\dots p_s^{r_s}$ *be the prime decomposition of an integer* n *relatively prime to* $2D$. *Let* $n-\epsilon_D(n)=2^\kappa q$ *and* $p_i-\epsilon_D(p_i)=2^{k_i}q_i$, *with* q,q_i *odd, ordering the* p_i*'s such that* $k_1\le\cdots\le k_s$. $C_{3,D}$ *consists of the following numbers:*

1. $n=9,25,49.$
2. $n=p_1p_2=\begin{cases}(2^{k_1}q_1-1)(2^{k_1}q_1+1),\\(2^{k_1}q_1+\epsilon_D(p_1))(3\cdot 2^{k_1}q_1+\epsilon_D(p_2)),\\(2^{k_1}q_1+\epsilon_D(p_1))(2\cdot 2^{k_1}q_1+\epsilon_D(p_2))\ \textit{with}\ (q_1,k_1)\neq(1,1),\end{cases}$
   where each factor is prime.
3. $n=p_1p_2p_3$ *is a product of three distinct prime factors,* $p_i-\epsilon_D(p_i)\mid n-\epsilon_D(n)$ *and there is some integer* k_1 *such that* $2^{k_1}\ ||\ p_i-\epsilon_D(p_i)$ *for all* $i\in\{1,2,3\}$.

Proof 1. Let $s=1$, hence $n=p_1^{r_1}$, where $r_1\ge 2$. By the second inequality of Lemma 9 we know that $\alpha_D(n)\le\frac{1}{p_i^{r_i-1}}$. Thus, if $r_1\ge 3$, then $\alpha_D(n)\le\frac{1}{9}$ and $n\notin C_3$. If $r_1=2$, then $\alpha_D(n)\le\frac{1}{11}$ for $p_i\ge 11$. Hence, the only possibilities are $n=9,25,49$. Let us check if such an $n\in C_{3,D}$.

Let $n=9$. If $\epsilon_D(3)=1$, we have by Lemma 12 that $\alpha_D(9)=\frac{1}{6}$. If $\epsilon_D(3)=-1$ however, we get by Lemma 12 that $\alpha_D(9)=\frac{1}{4}$. In both cases $9\in C_{3,D}$.

Let $n=25$. If $\epsilon_D(5)=1$, we get by Lemma 12 that $\alpha_D(25)=\frac{3}{20}$. If $\epsilon_D(5)=-1$, we get by Lemma 12 that $\alpha_D(25)=\frac{5}{30}$. In both cases $25\in C_{3,D}$.
Let $n=49$. If $\epsilon_D(7)=1$, we get by Lemma 12 that $\alpha_D(49)=\frac{5}{42}<\frac{1}{8}$, so such a decomposition of 49 would not be in $C_{3,D}$. If $\epsilon_D(7)=-1$ however, we get by Lemma 12 that $\alpha_D(49)=\frac{1}{8}$, so in this case $49\in C_{3,D}$.

2. Now let $s=2$, hence $n=p_1^{r_1}p_2^{r_2}$. If $p_1=3$, then $r_1\le 2$ and $r_2\le 1$, otherwise by the second inequality of Lemma 9 $\alpha_D(n)\le\frac{1}{18}$. If $p_1,p_2\ge 5$ it follows from the second inequality of Lemma 9 that $r_i=1$, because otherwise $\alpha_D(n)\le\frac{1}{2}\cdot\frac{1}{5}=\frac{1}{10}$. Thus, either $n=p_1p_2$ with $p_1,p_2>3$ or $n=3^2p_2$. We shall first treat the case $n=p_1p_2$ with $p_1,p_2>3$.
   Now let $n=p_1p_2$ with $p_1-\epsilon_D(p_1)=2^{k_1}q_1$ and $p_2-\epsilon_D(p_2)=2^{k_2}q_2$. If $k_2\ge k_1+2$, we have $\alpha_D(n)\le\frac{1}{8}$ by the third inequality of Lemma 9. Hence, either $k_2=k_1$ or $k_2=k_1+1$.
   By the first inequality of Lemma 9 either both $\frac{\gcd(q,q_1)}{q_1}=\frac{\gcd(q,q_2)}{q_2}=1$ or $\frac{\gcd(q,q_i)}{q_i}=\frac{1}{3}$ for exactly one i and $\frac{\gcd(q,q_j)}{q_j}=1$ for the other $j\neq i$, as otherwise $\alpha_D(n)\le\frac{1}{18}$.

If $k_2 = k_1 + 1$, it must hold that $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$, otherwise by Lemma 10 $\alpha_D(n) \leq \frac{1}{12}$.
Thus, we are left to check the following three cases: The first one is $k_1 = k_2$ with $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$, the second one is $k_1 = k_2$ with $\frac{\gcd(q,q_1)}{q_1} = 1$ and $\frac{\gcd(q,q_2)}{q_2} = \frac{1}{3}$, and the third one is $k_2 = k_1 + 1$ with $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$.
Let us look at the case where $k_1 = k_2$ with $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$. This is equivalent to saying $q_1, q_2 \mid q$. By Lemma 11 both q_1, q_2 divide $2^\kappa q = 2^{2k_1} q_1 q_2 \pm 2^{k_1}(q_1 \pm q_2)$. This is only possible if $q_1 = q_2$, and thus $p_1 - \epsilon_D(p_1) = 2^{k_1} q_1$ and $p_2 - \epsilon_D(p_2) = 2^{k_1} q_1$. In order for p_1 and p_2 to be distinct primes, we must have that $\epsilon_D(p_1) \neq \epsilon_D(p_2)$. Without loss of generality we let $\epsilon_D(p_1) = 1$ and $\epsilon_D(p_2) = -1$. Therefore,

$$n = (2^{k_1} q_1 - 1)(2^{k_1} q_1 + 1).$$

Let us check if such an n is indeed in $C_{3,D}$. By Lemma 13, we know that $\frac{SL(D,n)}{\varphi_D(n)} > \frac{1}{3}$ for all odd $q_1 \neq 1$, and for $q_1 = 1$ we have $\frac{SL(D,n)}{\varphi_D(n)} = \frac{1}{3} - \frac{1}{3 \cdot 4^{k_1}}$. Since this is monotonically increasing in k_1, we have $\alpha_D(n) = \frac{1}{3} - \frac{1}{3 \cdot 4^{k_1}} \geq \frac{1}{4}$. Thus, $n \in C_{3,D}$.
Now let us look at the case where $k_1 = k_2$ with $\frac{\gcd(q,q_1)}{q_1} = 1$ and $\frac{\gcd(q,q_2)}{q_2} = \frac{1}{3}$. Thus, q_1 and $\frac{1}{3} q_2$ both divide q and by Lemma 11 also $2^\kappa q = 2^{2k_1} q_1 q_2 \pm 2^{k_1}(q_1 \pm q_2)$. Hence $q_1 \mid q_2$ and $\frac{1}{3} q_2 \mid q_1$. This implies that there exists $a \in \mathbb{N}$ such that $q_1 \cdot a = q_2$, and $b \in \mathbb{N}$ such that $\frac{1}{3} q_2 b = q_1$. Solving the two equations yields $a = 3$ and $b = 1$, thus $q_2 = 3q_2$. Therefore, $p_1 - \epsilon_D(p_1) = 2^{k_1} q_1$ and $p_2 - \epsilon_D(p_2) = 2^{k_1} 3 q_1$. Thus,

$$n = (2^{k_1} q_1 + \epsilon_D(p_1))(2^{k_1} 3 q_1 + \epsilon_D(p_2)).$$

Let us check if an n is indeed $C_{3,D}$. By Lemma 12, we have $\alpha_D(n) = \frac{1}{4^{k_1}} \left(\left(\frac{q_1 - 1}{3 q_1} \right)^2 + \frac{4^{k_1} - 1}{9} \right)$. If $q_1 = 1$, we have $\alpha_D(n) = \frac{4^{k_1} - 1}{9 \cdot 4^{k_1}} < \frac{1}{8}$, so $n \notin C_3$. If $q_1 \neq 1$, we have $\alpha_D(n) = \frac{1}{4^{k_1}} \left(\frac{1}{3} \left(\frac{q_1 - 1}{q_1} \right)^2 + \frac{4^{k_1} - 1}{9} \right) \geq \frac{1}{3 \cdot 4^{k_1}} \cdot \frac{4^{k_1 + 1} - 1}{12} > \frac{1}{8}$. We used the fact that both $\frac{q_1 - 1}{q_1}$ and $\frac{4^{k_1 + 1} - 1}{4^{k_1}}$ are monotonically increasing functions in q_1 and k_1 respectively. Thus, $n \in C_3$.
Now let us look at the case $k_2 = k_1 + 1$ with $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$. By Lemma 11, both q_1, q_2 divide $2^k q = 2^{2k_1 + 1} q_1 q_2 \pm 2^{k_1}(q_1 \pm 2 q_2)$. Thus, $q_1 \mid 2 q_2$ and $q_2 \mid q_1$. Since q_1 is odd, we must have that $q_1 \mid q_2$, which is only possible when $q_1 = q_2$. Hence, $p_1 - \epsilon_D(p_1) = 2^{k_1} q_1$ and $p_2 - \epsilon_D(p_2) = 2^{k_1 + 1} q_1 = 2(2^{k_1} q_1) = 2(p_1 - \epsilon_D(p_1))$. Therefore,

$$n = p_1 p_2 = (2^{k_1} q_1 + \epsilon_D(p_1))(2 \cdot 2^{k_1} q_1 + \epsilon_D(p_2)).$$

Let us check if such an n is in C_3. By Lemma 12, we have that $\alpha_D(n) = \left(\frac{q_1 - 1}{q_1} \right)^2 \frac{1}{2 \cdot 4^{k_1}} + \frac{4^{k_1} - 1}{6 \cdot 4^{k_1}}$. If $q_1 = 1$ we obtain $\alpha_D(n) = \frac{4^{k_1} - 1}{6 \cdot 4^{k_1}}$. This is only greater than $\frac{1}{8}$ for $k_1 > 1$. For $k_1 = 1$, we obtain $\alpha_D(n) = \frac{1}{8}$, the only possibility is $n = (2 + \epsilon_D(p_1))(4 + \epsilon_D(p_2)) = 3 \cdot 5$. If $q_1 \neq 1$, using the fact that $(q_1 - 1)/q_1$ is monotonically increasing in q_1, we obtain

$$\alpha_D(n) = \left(\frac{q_1 - 1}{q_1} \right)^2 \frac{1}{2 \cdot 4^{k_1}} + \frac{4^{k_1} - 1}{6 \cdot 4^{k_1}} \geq \frac{4}{18 \cdot 4^{k_1}} + \frac{4^{k_1} - 1}{6 \cdot 4^{k_1}} = \frac{4}{18 \cdot 4^{k_1}} + \frac{1}{8} > \frac{1}{8}.$$

Now let us treat the case when $n = 3^2 p_2$. Since $3 - \epsilon_D(3) = 2^{k_1} q_1$, but $\epsilon_D(3) = \pm 1$, we have that $3 - \epsilon_D(3) \in \{2, 4\}$, which implies that $q_1 = 1$ and $k_1 \in \{1, 2\}$.
By the third inequality of Lemma 9, we have for $k_2 \geq k_1 + 2$ that $n \notin C_3$, thus either $k_1 = k_2$ or $k_2 = k_1 + 1$. Now let $k_2 = k_1$. Again it must hold that either $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$ or $\frac{\gcd(q,q_1)}{q_1} = 1$ and $\frac{\gcd(q,q_2)}{q_2} = 3$, since $q_1 = 1$. We have

$$\begin{aligned} 2^\kappa q =& n - \epsilon_D(n) = 3^2 p_2 - \epsilon_D(3^2 p_2) && (23)\\ =&(2^{k_1} + \epsilon_D(3))^2 (2^{k_1+\delta_2} q_2 + \epsilon_D(p_2)) - \epsilon_D(p_2)\\ =&(2^{2k_1} + 2^{k_1+1}\epsilon_D(3) + 1)(2^{k_1+\delta_2} q_2 + \epsilon_D(p_2)) - \epsilon_D(p_2)\\ =&q_2(2^{3k_1+\delta_2} + 2^{k_1+\delta_2} + \epsilon_D(3) 2^{2k_1+1+\delta_2}) + \epsilon_D(p_2)(2^{2k_1} + \epsilon_D(3) 2^{k_1+1}). && (24) \end{aligned}$$

Now let us look at the case where $\frac{\gcd(q,q_2)}{q_2} = 1$, meaning $q_2 \mid q$. Inequality (23) implies that $q_2 \mid 2^{2k_1} + \epsilon_D(3) 2^{k_1+1}$. If $k_1 = 1$ we must have that $\epsilon_D(3) = 1$, otherwise $2^{k_1} + \epsilon_D(3) \neq 3$. Hence $q_2 \mid 8$. If $k_1 = 2$, we must have that $\epsilon_D(3) = -1$, otherwise $2^{k_1} + \epsilon_D(3) \neq 3$. Hence, $q_2 \mid 8$. Since q_2 must be odd, the only possibility is $q_2 = 1$. This analysis holds for both $k_2 = k_1$ and $k_2 = k_1 + 1$. Therefore, we get

$$p_2 = 2^{k_2} q_2 + \epsilon_D(p_2) = \begin{cases} 2^{k_1} q_2 \pm 1 = 1, 3, & \text{for } k_2 = k_1 = 1, q_2 = 1,\\ 2^{k_1} q_2 \pm 1 = 3, 5, & \text{for } k_2 = k_1 = 2, q_2 = 1,\\ 2^{k_1+1} q_2 \pm 1 = 3, 5 & \text{for } k_2 = k_1 + 1, k_1 = 1, q_2 = 1,\\ 2^{k_1+1} q_2 \pm 1 = 7, 9 & \text{for } k_2 = k_1 + 1, k_1 = 2, q_2 = 1. \end{cases}$$

Since p_2 is a prime different from 3, we discard all other cases and are left with $p_2 \in \{5, 7\}$.
Now let us look at the case where $\frac{\gcd(q,q_2)}{q_2} = \frac{1}{3}$, meaning $\frac{1}{3} q_2 \mid q$. Here it must hold that $k_1 = k_2$. By the same reasoning as above we have $\frac{1}{3} q_2 \mid 2^{2k_1} + \epsilon_D(3) 2^{k_1+1}$, which implies $q_2 \mid 3(2^{2k_1} + \epsilon_D(3) 2^{k_1+1})$. For $k_1 = 1$ we have $\epsilon_D(3) = 1$ and hence $q_2 \mid 24$. If $k_1 = 2$ it holds that $\epsilon_D(3) = -1$ and hence $q_2 \mid 24$. Again since q_2 must be odd, the only possibility now is $q_2 = 3$. Thus, we get

$$p_2 = 2^{k_2} q_2 + \epsilon_D(p_2) = \begin{cases} 2^{k_1} q_2 \pm 1 = 5, 7, & \text{for } k_1 = 1, q_2 = 3,\\ 2^{k_1} q_2 \pm 1 = 11, 13, & \text{for } k_1 = 2, q_2 = 3. \end{cases}$$

Again we discard the cases where p_2 is not prime or divisible by 3 and are left with $p_2 = 5, 7, 11, 13$.
We see that for $n = 3^2 p_2$ with $p_2 \geq 5$ prime and $n \in C_{3,D}$ the only possibilities are $n = 45, 63, 99, 117$. Now let us check if such an $n \in C_{3,D}$.
Let $n = 45$. By the arguments above, there are only three possible decompositions that would make $45 \in C_{3,D}$. The first being $\epsilon_D(5) = 1$, $\epsilon_D(3) = -1$ with $k_1 = k_2 = 2$ and $q_1 = q_2 = 1$, $q = 11$. By Lemma 12, this yields $\alpha_D(n) = \frac{5}{48} < \frac{1}{8}$. The second decomposition is $\epsilon_D(5) = \epsilon_D(3) = 1$ with $k_1 = 1, k_2 = 2$ and $q_1 = q_2 = 1, q = 11$. Again by Lemma 12, we get $\alpha_D(n) = \frac{1}{24}$. The third decomposition is $\epsilon_D(5) = -1$, $\epsilon_D(3) = 1$ with $k_1 = k_2 = 1$, and $q_1 = 1, q_2 = 3, q = 23$. This gives us $\alpha_D(n) = \frac{1}{36}$. In any case $45 \notin C_{3,D}$.
Let $n = 63$. By the arguments above there are only two possible decompositions that would make $63 \in C_{3,D}$. The first one being $\epsilon_D(7) = -1, \epsilon_D(3) = 1$, with $k_1 = 2, k_2 = 3$ and $q_1 = q_2 = 1$, $q = 1$. By Lemma 12, this yields $\alpha_D(63) = \frac{5}{96} < \frac{1}{8}$. The second decomposition is $\epsilon_D(5) = \epsilon_D(3) = 1$ with $k_1 = k_2 = 1$ and $q_1 = 1, q_2 = 3, q = 31$. Again by Lemma 12, we get $\alpha_D(63) = \frac{1}{36}$. In any case $63 \notin C_{3,D}$.

By the above arguments the values for s, k_1, k_2, q_1, q_2 and $\gcd(q, q_i)$ for $i = 1, 2$, so that $n = 99$ and $n = 117$ could be in $C_{3,D}$ are the same and by Lemma 12, define $\alpha_D(n)$, we get that both $\alpha_D(99) = \alpha_D(117) = \frac{5}{144}$, so both $99, 117 \notin C_{3,D}$.

3. Now let $s = 3$ with $n = p_1^{r_1} p_2^{r_2} p_3^{r_3}$. By the second inequality of Lemma 9, it must hold that $r_i = 1$ for all $i = 1, 2, 3$, otherwise $\alpha_D(n) \leq \frac{1}{12}$. Therefore, $n = p_1 p_2 p_3$ with $p_i \neq p_j$ for every $i \neq j$. By the first inequality of Lemma 9, we have that $\frac{\gcd(q,q_i)}{q_i} = 1$ for all $i = 1, 2, 3$, otherwise $\alpha_D(n) \leq \frac{1}{12}$. Thus, $q_i \mid q$ for every $i = 1, 2, 3$. By the third inequality of (9), we must have that $k_1 = k_2 = k_3$, as else $\alpha_D(n) \leq \frac{1}{8}$.
Therefore, we have $k_1 = k_2 = k_3$ with $q_i \mid q$ for all $i \in \{1, 2, 3\}$ thus also $q_i \mid 2^\kappa q$. Since $r_i = 1$ is odd for all i and also the number of $k_i = \kappa$ is odd, this implies that $2^{k_i} q_i \mid 2^\kappa q$, which is the same as saying that $p_i - \epsilon_D(p_i) \mid n - \epsilon_D(n)$.
Let us check if such an n is indeed in C_3. Using Lemma 12 and the fact that $k_1 = k_2 = k_3$, $q_i \mid q$ and $r_i = 1$ for $i = 1, 2, 3$ we get
$$\alpha_D(n) = \frac{1}{2^{3k_1}} \left(\prod_{i=1}^{3} \frac{q_i - 1}{q_i} + \frac{2^{3k_1} - 1}{7} \right) = \frac{1}{2^{3k_1}} \prod_{i=1}^{3} \frac{q_i - 1}{q_i} + \frac{1}{7} \cdot \frac{2^{3k_1} - 1}{2^{3k_1}}.$$
Since $\frac{2^{3k_1}-1}{3k_1}$ is monotonically increasing in k_1, we get $\frac{2^{3k_1}-1}{2^{3k_1}} \geq \frac{2^3-1}{2^3} = \frac{7}{8}$. Thus
$$\alpha_D(n) = \frac{1}{2^{3k_1}} \prod_{i=1}^{3} \frac{q_i - 1}{q_i} + \frac{1}{7} \cdot \frac{2^{3k_1} - 1}{2^{3k_1}} \geq \frac{1}{2^{3k_1}} \prod_{i=1}^{3} \frac{q_i - 1}{q_i} + \frac{1}{8} > \frac{1}{8}.$$
With this, we indeed have that such an $n \in C_3$.
4. Now let $s \geq 4$. By the second inequality of Lemma 9, we immediately have that $\alpha_D(n) \leq \frac{1}{8}$, thus, $n \notin C_3$.

5.2 Twin-prime products

Let $\pi_2(x) =\mid \{p \leq x : \Omega(p+2) = 1\} \mid$ denote the twin-prime counting function, which counts the number of twin-prime tuples up to x. The following theorem bounds the number of twin-primes for $x > e^{42}$.

Theorem 16 *(Riesel, Vaughan [19]) For $x > e^{42}$, we have*
$$\pi_2(x) < \frac{16\alpha x}{(7.5 + \ln(x))\ln(x)},$$
where α is called the Twin Prime Constant,
$$\alpha = \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) = \prod_{p>2} \frac{p(p-2)}{(p-1)^2} \approx 0.6602\ldots$$

By Theorem 15, we know that $n = (2^{k_1} q_1 - 1)(2^{k_1} q_1 + 1)$, where both factors are prime, belong to $C_{3,D}$. This is nothing but a subset of the set of products of twin-primes and Theorem 16 gives us a way to upper bound the size of this set for integers in M_k.

Lemma 14 *For $k \geq 122$ there exists less than $6\frac{2^{k/2}}{k^2}$ k-bit integers that are twin-prime products.*

Proof As $n = p(p+2)$ is a k-bit integer, p must be a $k/2$-bit integer. Thus, we only have to consider the number of twin-primes up to $2^{k/2}$. With Theorem 16 we obtain

$$\pi_2(2^{k/2}) < \frac{16\alpha 2^{k/2}}{(7.5 + \ln(2^{k/2}))\ln(2^{k/2})} < \frac{16\alpha}{4\ln^2(2)}\frac{2^{k/2}}{k^2} < 6\frac{2^{k/2}}{k^2},$$

which holds for $2^{k/2} > e^{42}$, so that $k \geq 122$.

5.3 Lucas–Carmichael numbers with three prime factors

We will now examine the numbers falling into the third category as outlined in Theorem 15. It is important to note that these numbers are not arbitrary integers; they have already been classified. Let us consider another congruence based on the Lucas sequence and applicable to prime numbers.

Theorem 17 *(Baillie, Wagstaff [4]) Let D be a fixed integer. If n is an odd prime, and if* $\gcd(n, Q) = 1$, *then*

$$U_{n-\epsilon_D(n)}(P, Q) \equiv 0 \bmod n. \tag{25}$$

In a similar manner, we can formulate a probabilistic primality test using this property, referred to as the *(weak) Lucas test*.

If n is composite, but congruence (25) of Theorem 17 still holds, then we call n a *Lucas pseudoprime with parameters P and Q*, or *lpsp*(P, Q).

Unfortunately, there are integers that satisfy this congruence for all suitable bases. This leads us to the introduction of the following definition.

Definition 10 Let D be a fixed integer. If n is an odd composite integer, such that for all $P, Q \in \mathbb{N}$ with $\gcd(P, Q) = 1$, $P^2 - 4Q = D$ and $\gcd(n, QD) = 1$ property (25) still holds, we call n a *Lucas–Carmichael number*.

These numbers draw parallels with Carmichael numbers, which are composites satisfying Fermat's Little Theorem, also called the probable prime test, the weak version of the Miller-Rabin test, for all appropriate bases. Specifically, a Carmichael number is an odd composite integer n that satisfies $a^{n-1} \equiv 1 \bmod n$ for all a such that $\gcd(a, n) = 1$. Carmichael [6] demonstrated that if n is a Carmichael number, it can be expressed as the product of $k \geq 3$ distinct primes $n = \prod_{i=1}^{k} p_i$, with the property that $p_i - 1 \mid n - 1$ for all $i = 1, 2, \ldots, k$.

If n is a Lucas–Carmichael number with respect to either $D = 1$ or D a perfect square, then it can be shown that n is a Carmichael number. In this regard, Lucas–Carmichael numbers extend the concept of Carmichael numbers. In 1977, Williams [21] established the following theorem, further strengthening the connection between these notions.

Theorem 18 *(Williams [21]) Let D be fixed. If n is a Lucas–Carmichael number, then it is a product of k distinct primes $p_1, p_2, \ldots, p_k$ and*

$$p_i - \epsilon_D(p_i) \mid n - \epsilon_D(n) \qquad \text{for all } i = 1, 2, \ldots, k.$$

Thus, as per Theorem 18, it becomes evident that the numbers of the third form in Theorem 15 precisely correspond to Lucas–Carmichael numbers with three prime factors, with the additional property that there exists a natural number k_1 satisfying $2^{k_1} \mid\mid p_i - \epsilon_D(p)$ for all prime factors p of n. Bounding the number of Lucas–Carmichael numbers below a given integer remains an open question in number theory. Our task, however, is simplified, as we only need to constrain a subset of such numbers. Supposing that we have the additional property $\epsilon_D(p) = -1$ for all primes p dividing n, we can demonstrate the following result:

Theorem 19 *Let $M(x)$ denote the number of Lucas–Carmichael numbers n up to x with exactly three prime factors and Jacobi-symbol $\epsilon_D(n) = -1$. Furthermore, let us suppose that $\epsilon_D(p) = -1$ for all primes p dividing n. Then, for all $x \geq 1$ we have*

$$M(x) \leq (0.1668)x^{5/9}\ln(x)^{11/4}.$$

Proof A Lucas–Carmichael number n with exactly three prime factors can be written as $n = pqr$ with $2 < p < q < r$ primes and has the property that $p - \epsilon_D(p) \mid pqr - \epsilon_D(pqr)$, $q - \epsilon_D(q) \mid pqr - \epsilon_D(pqr)$, $r - \epsilon_D(r) \mid pqr - \epsilon_D(pqr)$. Let $g = \gcd(p - \epsilon_D(p), q - \epsilon_D(q), r - \epsilon_D(r))$ and let a, b, c be such that

$$\begin{aligned} p - \epsilon_D(p) &= ga, \\ q - \epsilon_D(q) &= gb, \\ r - \epsilon_D(r) &= gc. \end{aligned}$$

Thus, $a \leq b \leq c$. We want to show that $\gcd(a, b, c) = 1$. Suppose that $\gcd(a, b, c) = d > 1$. Then we can write $a = a'd$, $b = b'd$ and $c = c'd$ with $\gcd(a', b', c') = 1$ for some integers $a', b', c' > 1$. Thus, $p - \epsilon_D(p) = a'dg$, $q - \epsilon_D(q) = b'dg$ and $r - \epsilon_D(r) = c'dg$. It follows that that $\gcd(p - \epsilon_D(p), q - \epsilon_D(q), r - \epsilon_D(r)) = gd > g$, which is a contradiction. Thus, $\gcd(a, b, c) = 1$. Moreover, we know that

$$\begin{aligned} n = pqr &= (ga + \epsilon_D(p))(gb + \epsilon_D(q))(gc + \epsilon_D(r)) \\ &= g^3abc + g^2ac\epsilon_D(q) + g^2ab\epsilon_D(r) + g^2bc\epsilon_D(p) + ga\epsilon_D(q)\epsilon_D(r) \\ &\quad + gb\epsilon_D(p)\epsilon_D(r) + gc\epsilon_D(p)\epsilon_D(q) + \epsilon_D(p)\epsilon_D(q)\epsilon_D(r). \end{aligned}$$

Thus,

$$\begin{aligned} n - \epsilon_D(n) = g(&g^2abc + gac\epsilon_D(q) + gab\epsilon_D(r) + gbc\epsilon_D(p) \\ &+ a\epsilon_D(q)\epsilon_D(r) + b\epsilon_D(p)\epsilon_D(r) + c\epsilon_D(p)\epsilon_D(q)). \end{aligned}$$

Since $ga = p - \epsilon_D(p) \mid n - \epsilon_D(n)$, it holds that

$$\begin{aligned} a \mid g^2abc &+ gac\epsilon_D(q) + gab\epsilon_D(r) + gbc\epsilon_D(p) + a\epsilon_D(q)\epsilon_D(r) + b\epsilon_D(p)\epsilon_D(r) \\ &+ c\epsilon_D(p)\epsilon_D(q). \end{aligned}$$

With this it follows that

$$a \mid gbc\epsilon_D(p) + b\epsilon_D(p)\epsilon_D(r) + c\epsilon_D(p)\epsilon_D(q) = \epsilon_D(p)(gbc + b\epsilon_D(r) + c\epsilon_D(q)).$$

This directly yields

$$a \mid gbc + b\epsilon_D(r) + c\epsilon_D(q). \tag{26}$$

Using a symmetric argument, we have that

$$b \mid gac + a\epsilon_D(r) + c\epsilon_D(q), \tag{27}$$

$$c \mid gbc + a\epsilon_D(q) + b\epsilon_D(r). \tag{28}$$

Relation (26) implies that $\gcd(a, b) \mid c$, but we already know that $\gcd(a, b, c) = 1$, thus it follows that $\gcd(a, b) = 1$. Relations (27) and (28) imply by a similar argument that $\gcd(a, c) = \gcd(b, c) = 1$.

Therefore, the relations in (26), (27) and (28) imply that if a, b, c are given, then g is determined $\mod abc$.
By our assumption we have $\epsilon_D(n) = \epsilon_D(p) = -1$ for all primes $p \mid n$. Thus,

$$n = pqr = (ag-1)(bg-1)(cg-1) = g^3abc - g^2(ab+ac+bc) + g(a+b+c) - 1,$$

with $q = bg - 1 \geq p + 2 = ag + 1$, and $r = bg - 1 \geq p + 2 = ag + 1$.

We now count the number M of quadruples a, b, c *and* g that satisfy both of the above conditions. We note that $M(x) \leq M$. We write $M = M_1 + M_2 + M_3$, where in M_1 we count the number of quadruples with $g > abc$, in M_2 we count the number of quadruples with $G < g \leq abc$ and in M_3 we count the number of quadruples with $g \leq G$ and $g \leq abc$. Here G is a parameter that we will choose later with the goal of optimizing the bounds of M_2 and M_3.

Bounding M_1: We know that $g^3abc - g^2(ab+ac+bc) \leq x$. Since $g > abc$, we know that $\frac{g}{ab} > c$, $\frac{g}{ac} > b$ and $\frac{g}{bc} > a$. Thus,

$$g^3abc - g^2(\frac{g}{c} + \frac{g}{a} + \frac{g}{b}) = g^3(abc - (\frac{1}{c} + \frac{1}{a} + \frac{1}{b})) \leq g^3abc - g^2(ab+ac+bc) \leq x.$$

But with $a \geq 1$, and $c > b > a$, we get that $\frac{1}{c} + \frac{1}{a} + \frac{1}{b} < 1 + \frac{1}{2} + \frac{1}{3} < 2$. Thus,

$$g^3(abc - 2) < g^3(abc - (\frac{1}{c} + \frac{1}{a} + \frac{1}{b})) < x.$$

Since $a \geq 1$ and $c > b > a$, thus $b \geq 2$ and $c \geq 3$, it holds that $\frac{1}{2}abc \leq abc - 2$ as $\frac{1}{2} \leq 1 - \frac{2}{abc} \leq 1 - \frac{2}{1 \cdot 2 \cdot 3} = \frac{2}{3}$. Thus, it holds that $\frac{1}{2}abc \leq abc - 2$, and with this we have

$$\frac{1}{2}g^3abc \leq x.$$

For a, b, c given, the number of g with $\frac{1}{2}g^3abc \leq x$, g in a specific residue class $\mod abc$, and $g > abc$ is at most $\left\lfloor \frac{(2x/(abc))^{1/3}}{abc} \right\rfloor \leq \frac{(2x)^{1/3}}{(abc)^{4/3}}$.

$$M_1 \leq \sum_{a<b<c} \frac{(2x)^{1/3}}{(abc)^{4/3}} < \frac{1}{6}(2x)^{1/3}\zeta\left(\frac{4}{3}\right)^3, \tag{29}$$

where ζ denotes the Riemann zeta function.

Bounding M_2: We know that $g^3abc - g^2(ab+bc+ac) \leq x$. Since $abc \geq ab, bc, ac$, we have that

$$g^2abc(g-3) = g^3abc - 3g^2abc \leq g^3abc - g^2(ab+bc+ac) \leq x.$$

Since $g \geq 3$, it follows that $g^2abc \leq x$, and thus $abc \leq \frac{x}{g^2} \leq \frac{x}{G^2}$. Furthermore, the area of $1/x$ from 1 to n is $\int_1^n \frac{1}{x}\,dx = \ln(n)$. Since $1/x$ is convex, we can lower bound it using rectangles as follows $\ln(n) = \int_1^n \frac{1}{x}\,dx > \sum_{k=1}^n \frac{1}{k}$. Moreover, we have that $\ln(1/a) < a$ for $a > 0$, 37. Thus,

$$\begin{aligned}
M_2 &\le \sum_{1\le a\le\left(\frac{x}{G^2}\right)^{1/3}} \ \sum_{a\le b\le\left(\frac{x}{aG^2}\right)^{1/2}} \ \sum_{b<c\le\frac{x}{abG^2}} 1 \\
&< \sum_{1\le a\le\left(\frac{x}{G^2}\right)^{1/3}} \ \sum_{a\le b\le\left(\frac{x}{aG^2}\right)^{1/2}} \frac{x}{abG^2} \\
&< \frac{x}{G^2} \sum_{1\le a\le\left(\frac{x}{G^2}\right)^{1/3}} \frac{1}{a} \sum_{1\le b\le\left(\frac{x}{aG^2}\right)^{1/2}} \frac{1}{b} \\
&< \frac{x}{G^2} \sum_{1\le a\le\left(\frac{x}{G^2}\right)^{1/3}} \frac{1}{a} \ln\left(\left(\frac{x}{aG^2}\right)^{1/2}\right) \qquad (30)\\
&= \frac{x}{2G^2} \sum_{1\le a\le\left(\frac{x}{G^2}\right)^{1/3}} \frac{1}{a}\left(\ln\left(\frac{1}{a}\right)+\ln\left(\frac{x}{G^2}\right)\right) \\
&< \frac{x}{2G^2}\left(1+\ln\left(\frac{x}{G^2}\right)\right) \sum_{1\le a\le\left(\frac{x}{G^2}\right)^{1/3}} \frac{1}{a} \\
&< \frac{x}{2G^2}\left(1+\ln\left(\left(\frac{x}{G^2}\right)\right)\right) \ln\left(\frac{x}{G^2}\right)^{1/3} \\
&< \frac{x}{6G^2}(\ln(x))^2,
\end{aligned}$$

for $G > e^{1/2}$.

Bounding M_3: From the relations (26), (27), and (28) we have that for given g, a, b, c, there is an integer h with

$$c = \frac{gab - a - b}{h} = \frac{(ga-1)b - a}{h}, \qquad (31)$$

so that

$$h \mid (ga-1)b - a.$$

We also have that $hc = (ga-1)b - a \le gab - a$, which implies $h \le ga\frac{b-a}{c}$. With $a < b < c$, it immediately follows that $h \le ga$. Note that

$$gac - a - c = (ga-1)c - a = \frac{(ga-1)^2 b - (ga-1)a}{h} - a,$$

so that (27) implies $b \mid (ga-1)a - ha$. Since $(a, b) = 1$, we have

$$b \mid ga - 1 - h. \qquad (32)$$

Also, note that

$$gac - a - c = (ga-1)c - b = (ga-1)\frac{(ga-1)b - a}{h} - b,$$

so that (26) implies $a \mid (gb-1)b - hb$, and since $(a, b) = 1$, we have

$$a \mid gb - 1 - h. \qquad (33)$$

Let j be such that

$$b = \frac{ga - 1 - h}{j}, \qquad (34)$$

so that $a < b$ and $h \leq ga$ imply $j \leq 2g$. We have

$$gb - 1 - h = g\frac{ga - 1 - h}{j} - 1 - h,$$

so that (33) implies that $a \mid -g - gh - j - jh$, which also means that $a \mid g + gh + j + jh$, that is

$$(g + j)(1 + h) \equiv 0 \mod a. \tag{35}$$

Suppose we are given g, a, j. Let $d = \gcd(a, j(g + j))$. Note that (34) and (35) imply

$$1 + h \equiv ga \mod j, \qquad 1 + h \equiv 0 \mod \frac{a}{\gcd(a, g + j)}.$$

Thus,

$$1 + h \equiv ga \mod \frac{ja}{d}. \tag{36}$$

The number of positive integers $h \leq ga$ that satisfy (36) is at most

$$\left\lfloor \frac{ga - 1}{ja/d} \right\rfloor \leq \left\lfloor \frac{gd}{j} \right\rfloor \leq \frac{gd}{j}, \tag{37}$$

since $j \leq 2g$ implies $gd/j \geq d/2 \geq 1/2$. Further, if g, a, j, h are given, then b, c are also specified via (31) and (34). Moreover, with $q = bg - 1 \geq p + 2 = ag + 1$, and $r = bg - 1 \geq p + 2 = ag + 1$ we get from $n = pqr = (ag - 1)(bg - 1)(cg - 1)$ that $(ag)^3 < n = (ag - 1)(bg - 1)(cg - 1)$ and thus $(ag)^3 < x$. Thus, by (37)

$$\begin{aligned} M_3 &\leq \sum_{g \leq G} \sum_{j \leq 2g} \sum_{a \leq x^{1/3}/g} \frac{g \gcd(a, j(j + g))}{j} \\ &\leq \sum_{g \leq G} \sum_{j \leq 2g} \sum_{d \mid j(j+g)} \frac{gd}{j} \sum^{n}_{\substack{a \leq x^{1/3}/g \\ d \mid a}} i \\ &\leq x^{1/3} \sum_{g \leq G} \sum_{j \leq 2g} \sum_{d \mid j(j+g)} \frac{1}{j}. \end{aligned} \tag{38}$$

Next note that

$$\sum_{d \mid j(j+g)} 1 = \tau(j(j + g)) \leq \tau(j)\tau(j + g),$$

where $\tau(m)$ denotes the number of divisors of m. Thus, (38) yields

$$\begin{aligned} M_3 &\leq x^{1/3} \sum_{g \leq G} \sum_{j \leq 2g} \frac{\tau(j)\tau(j + g)}{j} \\ &= x^{1/3} \sum_{j \leq 2G} \frac{\tau(j)}{j} \sum_{j/2 \leq g \leq G} \tau(j + g) \\ &\leq x^{1/3} \Bigg(\sum_{j \leq 2G} \frac{\tau(j)}{j} \Bigg) \Bigg(\sum_{m \leq 3G} \tau(m) \Bigg). \end{aligned} \tag{39}$$

From Lemma 2.6 in [12] and its proof, we have that

$$\sum_{m\leq 3G}\tau(m)\leq 3G(1+\ln(3G)),\quad \sum_{j\leq 2G}\frac{\tau(j)}{j}\leq\frac{1}{2}(2+\ln(2G))^2.$$

With (39), we get

$$M_3\leq\frac{3}{2}x^{1/3}G(1+\ln(3G))(2+\ln(2G))^2. \tag{40}$$

We let $G=\frac{x^{2/9}}{\ln(x)^{1/4}}$ and assume $x>4\cdot 10^{10}$. We use the fact that $\zeta(\frac{4}{3})<4$ and that $\frac{4^3\cdot 2^{1/3}}{6}<\frac{1}{10^4}x^{2/9}\ln(x)^{11/4}$ for $x>4\cdot 10^{10}$, and by (29) we get

$$M_1<\frac{1}{10^5}x^{5/9}\ln(x)^{11/4}.$$

With our chosen value for G and (30) we have

$$M_2<\frac{1}{6}x^{5/9}\ln(x)^{11/4}.$$

Since $\frac{\partial}{\partial x}\left(\frac{1}{4}\ln(\ln(x))\right)<0$ for $x>1$, we have that $\frac{1}{4}\ln(\ln(x))$ is a strictly monotone increasing function for $x>1$ and thus, if $x>4\cdot 10^{10}$, we have $\frac{1}{4}\ln(\ln(x))>\frac{1}{4}\ln(\ln(4\cdot 10^{10}))$. So for our value of G we get

$$1+\ln(3G)=1+\ln(3)+\frac{2}{9}\ln(x)-\frac{1}{4}\ln(\ln(x))<\frac{3}{10}\ln(x), \tag{41}$$

for $x>4\cdot 10^{10}$. For the second inequality, we get by the same argument for our value of G that

$$2+\ln(2G)=2+\ln(2)+\frac{2}{9}\ln(x)-\frac{1}{4}\ln(\ln(x))<\frac{3}{10}\ln(x), \tag{42}$$

for $x>4\cdot 10^{10}$. Using inequalities (41) and (42) in (40) we get

$$M_3\leq\frac{3}{2}(\frac{3}{10})^3\ln(x)^{3-1/4}x^{1/3+2/9}=\frac{1}{25000}\ln(x)^{11/4}x^{5/9}.$$

Hence, we get

$$M(x)\leq M_1+M_2+M_3<(0.1668)x^{5/9}\ln(x)^{11/4},$$

for $x>4\cdot 10^{10}$. From the table in [8], where all Lucas–Carmichael numbers with $p+1\mid n+1$ for all $p\mid n$ are given, we see that the theorem is true for all $x\leq 4\cdot 10^{10}$, which concludes our proof.

To assume that $\epsilon_D(p)=-1$ holds true for all primes p dividing n, when $\epsilon_D(n)=-1$, becomes plausible when the Miller-Rabin test is combined with the strong Lucas test, as shown by the following discussions and lemmas:

Lemma 15 *(Williams [21]) Let integers D and n exist such that $\epsilon_D(n)=-1$. Suppose that n simultaneously qualifies a Lucas–Carmichael and Carmichael number. In such a case, n must have an odd number of prime factors, and further, $\epsilon_D(p)=-1$ must hold for all primes $p\mid n$.*

Lemma 16 *(Muller [16]) A necessary condition for a Carmichael number n to be a lpsp(P,Q) with $P\in\mathbb{Z}_n^*$ and $\epsilon_D(n)=-1$ is that n has an odd number of prime factors and $\epsilon_D(p)=-1$ for all primes $p\mid n$.*

Lemma 17 *(Muller [16]) Let $n = \prod_{i=1}^{r} p_i$, $p_i \neq 2, 3, 5$. Suppose that for some i there exists a parameter a_i with $\frac{p_i-1}{2} \mid ord_{p_i}(a_i)$ such that $a_i^{n-1} \equiv 1 \mod n$ for all bases a with $\gcd(a, n) = 1$. A necessary condition for n to be $lpsp(P, Q)$ for $P \in \mathbb{Z}_n^*$ and $\epsilon_D(n) = -1$ is that $\epsilon_D(p) = -1$ for all primes $p \mid n$.*

Hence, Lemmas 15 and 16 suggest that when an integer n is simultaneously a Lucas pseudoprime and a Carmichael number, we must have that $\epsilon_D(p) = \epsilon_D(n)$ for all primes $p \mid n$. Lemma 17 suggests that when n is a Lucas pseudoprime and satisfies Fermat's Little Theorem for certain specific bases, we must have $\epsilon_D(p) = \epsilon_D(n)$ for all primes $p \mid n$. Thus, there are reasons to believe that $\epsilon_D(p) = \epsilon_D(n)$ for all primes $p \mid n$, whenever n successfully passed numerous rounds of both the (strong) Lucas test and the (strong) probable prime test.

Moreover, we have the following result:

Lemma 18 *(Leng [13]) If n has three distinct prime factors p_1, p_2 and p_3, where $p_i = 2^{k_1} q_i + \epsilon_D(p_i)$, and if there exists a j_1 such that $2^{j_1} \mid\mid p_i - 1$, then $\epsilon_D(n) = \epsilon_D(p_i)$ for $i = 1, 2, 3$.*

The question of whether we can draw the conclusion that $\epsilon_D(p_i) = \epsilon_D(n)$ for $i = 1, 2, 3$, from the additional requirement that there exists a $k_1 \in \mathbb{N}$ such that $2^{k_1} \mid\mid p_i - \epsilon_D(p_i)$ whenever n is a Lucas–Carmichael number with three prime factors remains open.

Furthermore, we leave it open whether a specific condition can be incorporated into the primality test, such that the relation $\epsilon_D(p_i) = \epsilon_D(n)$ for $i = 1, 2, 3$ holds for all integers of the third form in C_3. The resolution of either of these questions would enable the utilization of Theorem 19 to establish unconditional bounds for large values of t.

5.4 Bounds for large *t*

Since we are unable to use Theorem 19 to impose bounds on all Lucas–Carmichael numbers of the third form within C_3, we cannot bound $|C_{3,D} \cap M_k|$, which is pivotal for the progression of our analysis. Nonetheless, if we can effectively demonstrate the equality $\epsilon_D(p_i) = -1$ for $i = 1, 2, 3$ for these specific numbers, we could proceed to use the results outlined in this section that hold unconditionally.

Lemma 19 *Let $k \geq 122$, and assume that whenever n is a Lucas–Carmichael number of the third form of C_3, we have $\epsilon_D(n) = -1$ and $\epsilon_D(p) = -1$ for the primes $p \mid n$. Then, $|C_{3,D} \cap M_k| \leq (0.061)2^{5k/9}k^{11/4}$.*

Proof We consider the elements of $C_{3,D}$ listed in Theorem 15 that are also in M_k. If $n = p(p+2)$, then by Corollary 14 we get for $k \geq 122$ that

$$\pi_2(2^{k/2}) \leq 88\frac{2^{k/2}}{k^2} \leq 10^{-10}2^{5k/9}k^{11/4}.$$

If $n = (m-1)(2m-1) \leq 2^k$, then $2(m-1)^2 \leq 2^k$. Using that m is even, we have at most $\sqrt{2^k/8} + 1/2$ such integers $n \leq 2^k$ of the form $n = (m-1)(2m-1)$ that are in $C_{3,D} \cap M_k$. If $n = (m-1)(3m-1) \leq 2^k$, then $3(m-1)^2 \leq 2^k$, we get at most $\sqrt{2^k/12} + 1/2$ such integers $n \leq 2^k$ of the form $n = (m-1)(3m-1)$ that are in $C_{3,D} \cap M_k$. For $k \geq 122$, we get

$$\frac{2^{k/2}}{\sqrt{8}} + 1/2 \leq 10^{-8}2^{5k/9}k^{11/4}, \quad \frac{2^{k/2}}{\sqrt{12}} + 1/2 \leq 10^{-8}2^{5k/9}k^{11/4}.$$

Now for the elements of the third form of Theorem 15, we have by Theorem 19 that

$$M(2^k) \leq (0.1668)2^{5k/9}\ln(2^k)^{11/4} < (0.0609)2^{5k/9}k^{11/4}.$$

Bounding the cardinality of $C_{3,D} \cap M_k$ using the above estimates yields

$$|C_{3,D} \cap M_k| < \big(-10^{-10} + 10^{-8} + 10^{-8} + 0.0609\big)2^{5k/9}k^{11/4} < (0.061)2^{5k/9}k^{11/4}.$$

which proves the corollary for $k \geq 122$.

We simply bound $|M_k \cap C_{m,D}|$ for $m = 4, 5$ using Theorem 8.

Lemma 20 *Let $k \geq 122$, then*

$$\begin{aligned} \mid M_k \cap C_{4,D} \mid amp; &\leq (2.39)2^{k-\frac{k}{4}}, \\ \mid M_k \cap C_{5,D} \mid amp; &\leq (2.37)2^{k-\frac{k}{5}}. \end{aligned}$$

Proof Since $m \leq 5$, we have that $m + 1 \leq 2\sqrt{k-1}$ for $k \geq 10$, thus we can use Theorem 8, which says that $\mid C_{m,D} \cap M_k \mid \leq 2^{k+1}\sum_{j=2}^{m} 2^{m-j-\frac{k-1}{j}}$. We want to find a $c_m \in \mathbb{R}$ such that $\mid M_k \cap C_{m,D} \mid \leq c_m 2^{k-\frac{k}{m}}$ for each $m = 4, 5$.

$$\begin{aligned} |M_k \cap C_{4,D}| \leq 2^{k+1}\sum_{j=2}^{4} 2^{4-j-\frac{k-1}{j}} &= 2^{k+1}(2^{2-\frac{k-1}{2}} + 2^{1-\frac{k-1}{3}} + 2^{-\frac{k-1}{4}}) \leq c_4 2^{k-\frac{k}{4}} \\ \Rightarrow 2^{\frac{7}{2}-\frac{k}{4}} + 2^{\frac{7}{3}-\frac{k}{12}} + 2^{\frac{5}{4}} &\leq c_4. \end{aligned}$$

With $k \geq 122$ we get $2^{\frac{7}{2}-\frac{k}{4}} + 2^{\frac{7}{3}-\frac{k}{12}} + 2^{\frac{5}{4}} \leq 2^{\frac{7}{2}-\frac{122}{4}} + 2^{\frac{7}{3}-\frac{122}{12}} + 2^{\frac{5}{4}} \leq c_4$. Therefore, we can set $c_4 = 2.39$. The argument for $|M_k \cap C_{5,D}|$ is identical.

Theorem 20 *Let $k \geq 122$, and assume that whenever n is a Lucas–Carmichael number of the third form of C_3, we have $\epsilon_D(p) = \epsilon_D(n)$ for the primes $p \mid n$. Then, we have*

$$q_{k,l,t} \leq k\big(2^{-1.52-4t}\frac{\rho_l^{6t}}{2^t - \rho_l^t} + \rho_l^{3t}2^{-3.55-\frac{4k}{9}-2t}k^{\frac{11}{4}} + \rho_l^{4t}2^{1.74-\frac{k}{4}-3t} + \rho_l^{5t}2^{1.73-\frac{k}{5}-4t}\big).$$

Proof We let $M = 5$ in inequality (14) and get

$$\sum{}'_{n\in M_{k,l}}\overline{\alpha}_D(n)^t \leq 2^{k-2+t}\sum_{m=6}^{\infty}\rho_l^{mt}2^{-mt} + \sum_{m=3}^{5}\rho_l^{mt}2^{-(m-1)t} \mid M_k \cap C_{m,D} \mid . \tag{43}$$

Evaluating the first sum yields

$$2^{k-2+t}\sum_{m=6}^{\infty}\rho_l^{mt}2^{-mt} = 2^{k-2-4t}\frac{\rho_l^{6t}}{2^t - \rho_l^t}.$$

We then use the bounds for $\mid M_k \cap C_{m,D} \mid$ from Lemmas 19 and 20. Furthermore, we use inequality (15) with inequality (43) and (1) for the respective bounds, which gives us the desired result.

The following corollary is useful for $t \geq k/9$.

Table 5 Lower bounds for $-\log_2(q_{k,l,\lceil k/9\rceil})$ and $-\log_2(p_{k,\lceil k/9\rceil})$ where l was chosen as discussed in Remark 1. For the probabilities, we use Corollary 5 for the strong Lucas test and Theorem 4 for the Miller-Rabin test

k	$-\log_2(q_{k,l,\lceil k/9\rceil})$	$-\log_2(p_{k,\lceil k/9\rceil})$
100	33	34
200	71	74
400	146	150
512	188	192
600	223	226
1024	385	389

Corollary 5 *Let $t \geq k/9$ and $k \geq 122$. We have*

$$q_{k,l,t} \leq k\left(2^{-1.52-4t}\frac{\rho_l^{6t}}{2^t-\rho_l^t} + \rho_l^{3t}2^{-3.55-\frac{4k}{9}-2t}k^{\frac{11}{4}} + \rho_l^{5t}2^{1.74-\frac{k}{4}-3t}\right).$$

Proof We bound the last term in Theorem 20 for $t \geq k/9$ and $k \geq 122$ by

$$\rho_l^{5t}2^{1.73-\frac{k}{5}-4t} \leq \rho_l^{5t}2^{-5.72-\frac{k}{4}-3t}.$$

We then bound the sum of the last two terms in Theorem 20 as follows

$$\begin{aligned}\rho_l^{4t}2^{1.74-\frac{k}{4}-3t} + \rho_l^{5t}2^{1.73-\frac{k}{5}-4t} &< \rho_l^{4t}2^{1.74-\frac{k}{4}-3t} + \rho_l^{5t}2^{-5.72-\frac{k}{4}-3t}\\ &< \rho_l^{5t}2^{1.74-\frac{k}{4}-3t},\end{aligned}$$

which concludes our proof.

The lower bounds in Table 5 for $q_{k,l,t}$ for the strong Lucas test were computed from Corollary 5, where l is chosen as discussed in Remark 1, and the bounds for the Miller-Rabin test for $p_{k,t}$ from Theorem 4.

6 Conclusion

In this paper, we established the framework needed to find average case error bounds for the strong Lucas test. No such bounds existed previously. We were able to show that the strong Lucas test is, in fact, reliable enough for almost all practical purposes. We examined an algorithm that chooses k-bit integers at random from the uniform distribution, runs t independent iterations of the strong Lucas test on this integer, and outputs the first number that passes all t tests. Let $q_{k,t}$ be the probability that this algorithm outputs a composite integer. The bounds we obtained are $q_{k,1} \leq \ln(k)k^2 4^{2.3-\sqrt{k}}$ for $k \geq 2$ and $q_{k,t} < \ln^t(k)\frac{k^{3/2}}{\sqrt{t}}4^{2.12+t-\sqrt{tk}}$ for $k \geq 21$ and $3 \leq t \geq (k-1)/9$ or $k \geq 88$ and $t = 2$. Since it is computationally less expensive to rule out candidates that are divisible by small primes by trial division than by using the strong Lucas test, trial division is often implemented before the actual primality test. We made use of this property and imposed the additional requirement of checking for divisibility by the first l odd primes before running the strong Lucas test. This yielded improved bounds. Let $q_{k,l,t}$ be the probability that this updated algorithm returns a composite number. Let $\tilde{p}_l$ denote the l-th odd prime and let $\rho_l = 1 + \frac{1}{\tilde{p}_{l+1}}$. We showed that $q_{k,l,1} < k^2 4^{1.8-\sqrt{k}}\rho_l^{2\sqrt{k-1}-2}$ for all $l \in \mathbb{N}$ and $k \geq 1$ and $q_{k,l,t} \leq 4^{1.72-\sqrt{tk}}k^{3/2}2^t\rho_l^{2\sqrt{kt}+t}$ for all $k \geq 21$ and $2 \leq t \leq (k-1)/9$ with $l \in \mathbb{N}$. These bounds are comparable to the bounds in [7] for the Miller-Rabin test. Furthermore, we classified the numbers that add the most to our probability estimate and realized that specific Lucas–Carmichael numbers with three prime factors belong to this set. Unfortunately, we could only bound those numbers under an extra constraint. We provided bounds for large values of t if this assumption were to hold.

During the scope of this work, we found average case error bounds for the strong Lucas test. Yet, many open questions that look promising for future projects remain. For example, future work could be to get such average case estimates where we average over both D and n. Moreover, one could also find bounds for the error in case we look for a prime by incremental search from a random starting point. Furthermore, one could analyze if it is possible to get improved estimates for the Miller-Rabin test using the modified algorithm that includes division by small primes. The most interesting future work, however, is to get average case error bounds for the Baillie-PSW test, which is a probabilistic primality test that combines one round of the Miller-Rabin test with one round of the strong Lucas test. No one has found counterexamples for composites passing this test; therefore, this primality test seems very promising.

Acknowledgements Semira Einsele acknowledges the financial support by the 6G research cluster funded by the Federal Ministry of Education and Research (BMBF) in the programme of “Souverän. Digital. Vernetzt.” Joint project 6G-RIC, project identification number: 16KISK020K. Special thanks go to Mia Filić for all the fruitful discussions.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. Akbary A., Friggstad Z., Juricevic R.: Explicit upper bounds for $f(n) = \prod_{p \le p_{\omega(n)}} \frac{p}{p-1}$. Contrib. Discrete Math. **2**(2) (2007).
2. Albrecht M.R., Massimo J., Paterson K.G., Somorovsky J.: Prime and prejudice: primality testing under adversarial conditions. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, pp. 281–298 (2018).
3. Arnault F.: The Rabin-Monier theorem for Lucas pseudoprimes. Math. Comput. **66**(218), 869–881 (1997).
4. Baillie R., Wagstaff S.S.: Lucas pseudoprimes. Math. Comput. **35**(152), 1391–1417 (1980).
5. Brandt J., Damgård I.: On generation of probable primes by incremental search. In: Advances in Cryptology-CRYPTO’92: 12th Annual International Cryptology Conference Santa Barbara, California, USA August 16–20, 1992 Proceedings 12, pp. 358–370 (1993). Springer.
6. Carmichael R.D.: Note on a new number theory function. Bull. Am. Math. Soc. **16**(5), 232–238 (1910).
7. Damgård I., Landrock P., Pomerance C.: Average case error estimates for the strong probable prime test. Math. Comput. **61**(203), 177–194 (1993).
8. Donovan L.P.P.a.J.: Table of n for $a(n)$ for $n = 1 \dots 10000$. https://oeis.org/A006972/b006972.txt. Accessed 12 June 2023.
9. Galbraith S., Massimo J., Paterson K.G.: Safety in numbers: on the need for robust diffie-hellman parameter validation. In: Public-Key Cryptography–PKC 2019: 22nd IACR International Conference on Practice and Theory of Public-Key Cryptography, Beijing, China, April 14-17, 2019, Proceedings, Part II, pp. 379–407 (2019). Springer.

10. Hardy G.H., Wright E.M.: An Introduction to the Theory of Numbers. Oxford University Press, Oxford (1979).
11. Ireland K.: A Classical Introduction to Modern Number Theory. Graduate Texts in Mathematics, vol. 84. Springer, New York (1990).
12. Kim S.H., Pomerance C.: The probability that a random probable prime is composite. Math. Comput. **53**(188), 721–741 (1989).
13. Leng A.: Independence of the miller-rabin and lucas probable prime tests (2017).
14. Miller G.L.: Riemann's hypothesis and tests for primality. In: Proceedings of the Seventh Annual ACM Symposium on Theory of Computing, pp. 234–239 (1975).
15. Monier L.: Evaluation and comparison of two efficient probabilistic primality testing algorithms. Theor. Comput. Sci. **12**(1), 97–108 (1980).
16. Müller S.: On the combined fermat/lucas probable prime test. In: Cryptography and Coding: 7th IMA International Conference Cirencester, UK, December 20–22, 1999 Proceedings 7, pp. 222–235 (1999). Springer.
17. Rabin M.O.: Probabilistic algorithm for testing primality. J. Number Theory **12**(1), 128–138 (1980).
18. Rabin M.O.: Probabolistic algorithm for primality testing. J. Number Theory **12**, 128–138 (1980).
19. Riesel H., Vaughan R.C.: On sums of primes. Arkiv för Matematik **21**(1), 45–74 (1983).
20. Suwa N.: Some remarks on Lucas pseudoprimes. Math. J. Okayama Univ. **54**(1), 1–32 (2012).
21. Williams H.C.: On numbers analogous to the Carmichael numbers. Can. Math. Bull. **20**(1), 133–143 (1977).