Some Number-theoretic Conjectures and Their Relation to the Generation of Cryptographic Primes [1]

Ueli M. Maurer

Institute for Signal and Information Processing
Swiss Federal Institute of Technology
CH-8092 Zürich, Switzerland

Abstract. The purpose of this paper is to justify the claim that a method for generating primes presented at EUROCRYPT'89 generates primes with virtually uniform distribution. Using convincing heuristic arguments, the conditional probability distributions of the size of the largest prime factor $p_1(n)$ of a number n on the order of N is derived, given that n satisfies one of the conditions $2n+1$ is prime, $2an+1$ is prime for a given a, or the d integers $u_1, \ldots, u_d$, where $u_1 = 2a_1n+1$ and $u_t = 2a_tu_{t-1}+1$ for $2 \leq t \leq d$, are all primes for a given list of integers $a_1, \ldots, a_d$. In particular, the conditional probabilities that n is itself a prime, or is of the form "k times a prime" for $k = 2, 3, \ldots$, is treated for the above conditions. It is shown that although for all k these probabilities strongly depend on the condition placed on n, the probability distribution of the relative size $\sigma_1(n) = \log_N p_1(n)$ of the largest prime factor of n is virtually independent of any of these conditions. Some number-theoretic conjectures related to this analysis are stated. Furthermore, the probability distribution of the size of the smaller prime factor of an RSA-modulus of a certain size is investigated.

[1] To appear in the proceedings of the second IMA conference on cryptography and coding, Dec. 18-20, Cirencester, England, to be published by Oxford University Press.

1. Introduction

At EUROCRYPT'89 a new method for generating cryptographic primes was presented by this author [6] that offers several advantages over all presently known methods for generating primes. The generated numbers are provable primes rather than only pseudo-primes as for most common approaches [8, 13], certain security conditions imposed upon the primes by the cryptosystem (e.g. RSA, Diffie-Hellman, ElGamal, etc.) can easily be satisfied, and the method is faster than all previous methods for generating primes or even only pseudo-primes. Moreover it is claimed in [6] that the achieved distribution is virtually uniform over all primes in a given interval that satisfy the security conditions. This claim cannot be proved but it is based on convincing heuristic arguments leading to some number-theoretic conjectures which may also be of independent interest. The main aim of this paper is to state and justify these conjectures and to explain their relation to cryptography. In this section, we give a motivation for the questions treated in this paper.

It is often recommended [9] to select the RSA-primes of a very special form: Both primes should have equally many binary digits, and they should be of the form $2p'+1$ where p' is again a prime of the form $2p''+1$ and p'' is a prime. Since the security of the RSA system should rely on the general problem of factoring the product of two large primes, and not on a very special case of this problem, it was recommended [6] to select the RSA-modulus randomly from the set of all RSA-moduli that satisfy certain reasonably restrictive security constraints, rather than to require that the modulus is of a very special form. This recommendation is the motivation for the analysis presented here.

The algorithm of [6] is based on a theorem by Pocklington [7] that allows one to prove the primality of a number n, provided that one knows part of the prime factorization of $n-1$. More precisely, the prime factorization of some integer F, where $n-1 = 2RF$ with $F > R$, must be known. Note that it is computationally infeasible to generate a large prime by selecting odd integers n of the desired size and partially factoring $n-1$, until an n is selected which can be proved to be prime by Pocklington's theorem. A better approach to the generation of primes of a certain size, say on the order of N, which is also considered in [11], is to construct an integer $F > \sqrt{N}$ as the product of some known primes and to repeatedly select an integer R on the order of $N/2F$ until $2RF+1$ can be proved to be prime. However, if the prime factors of F are selected from a subset of all primes (e.g. the set of primes smaller than 10^6), then only a very small fraction of all primes of the desired size can be reached by this construction approach. In order to avoid that the factored part of $p-1$ consists only of small prime factors, one can use this method repeatedly to construct larger and larger primes, by letting every prime be the factored part F of the next prime (see [11]). However, this modification does not increase the diversity of primes that can be reached by this construction.

The aim of [6] is to use the above construction for generating primes, but nevertheless to obtain primes that are uniformly distributed over a given interval $[N_1, N_2]$ (e.g. $[10^{100}, 10^{101}]$) centered at $N = \sqrt{N_1 N_2}$, and at the same time to speed up the generation process. This is achieved by a recursive algorithm: instead of constructing F as the product of some known primes, the largest prime factor p_1 of $(p-1)/2$ is generated first. The size of p_1, or, more specifically, the relative size $\sigma_1 = \log_{N/2} p_1$ of p_1, where $0 < \sigma_1 \leq 1$, is randomly selected according to the conditional probability distribution (which will be considered in section 2) of the size of the largest prime factor of an integer n on the order of $N/2$, given that $2n + 1$ is prime. Then p_1 is generated on the order of $(N/2)^{\sigma_1}$ by recursive application of the prime generating procedure.

It will be shown that, with or without the condition that $2n + 1$ is prime, the probability that the largest prime factor of a random integer n is greater than its square root is approximately $\log 2 = 0.69$. In this case, i.e., when $p_1 > \sqrt{N}$, we let $F = p_1$ and generate p by repeatedly and randomly selecting R on the order of $N/2p_1$ until $2Rp_1 + 1$ is proved prime by application of Pocklington's theorem. It is only in the remaining 31% of the cases that we need to generate a second prime factor, p_2, of $(p-1)/2$. Its size is selected according to the conditional probability distribution of the size of the largest prime factor of a number r on the order of $N/2p_1$, given that $p_2 \leq p_1$ and given that $2p_1 r + 1$ is prime. In section 2, we will therefore also consider the conditional probability distribution of the size of the largest prime factor of n, given that $2an + 1$ is prime for some fixed constant $a \geq 1$.

If necessary, the third, fourth, etc., largest prime factors of $(p-1)/2$ are generated similarly, until their product $F = \prod_{i \geq 1} p_i$ is large enough to guarantee that the remaining factor $R \approx N/2F$ of $(p-1)/2$, which is chosen at random, is not greater than the smallest generated prime factor of F. This condition must be satisfied to guarantee that a randomly selected R does not contain a prime factor greater than the smallest prime in F, a circumstance that would destroy the uniform distribution established by using the appropriate distributions of the sizes of the prime factors of F.

When the relative size σ_1 of the largest prime factor p_1 of $(p-1)/2$ is very close to one, a different type of recursion than mentioned above must be used. With small, but non-negligible probability (which will be considered in section 2), $(p-1)/2$ is itself a prime or a small multiple of a prime, i.e., $p_1 = (p-1)/2$ or $p_1 = (p-1)/2k$ for a small integer k. In this case the factor $R = (p-1)/2F = k$ is fixed and can therefore not be randomly selected. Instead, we have to repeatedly generate p_1 until $2kp_1 + 1$ is prime.

To make a concrete example, assume that a prime p on the order of $N = 10^{100}$ (e.g. $p \in [10^{100}/2, 2 \cdot 10^{100}]$) has to be generated, and that the random experiment selecting the size of the largest prime factor p_1 of $(p-1)/2$ indicates that $(p-1)/2$

should be of the form $(p–1)/2 = 3p_1$ (i.e., $R = 3$). Assume further that in the process of generating p_1, the random experiment selecting the size of the largest prime factor p_{11} of $(p_1–1)/2$ indicates that $(p_1–1)/2$ should be of the form $(p_1–1)/2 = 4p_{11}$. Then, in the process of generating p_{11}, the size of the largest prime factor of $(p_{11} - 1)/2$ must be selected according to the conditional probability distribution of the size of the largest prime factor of an integer n on the order of $10^{100}/(2 \cdot 6 \cdot 8) \approx 10^{98}$, given that $2n + 1$, $8(2n + 1) + 1 = 16n + 9$ and $6(16n + 9) + 1 = 96n + 55$ are primes. In section 3, we discuss the conditional probability distribution of the size of the largest prime factor of an number n, given that the d integers $u_1 = 2a_1n + 1$, $u_2 = 2a_2u_1 + 1 = 4a_1a_2n + 2a_2 + 1) + 1$, ... , $u_d = 2a_du_{d-1} + 1$ are all primes for a given list $a_1, \ldots, a_d$ of d integers.

The described method for generating primes can be turned into a method for generating random secure RSA-moduli (see [6]). One of the security constraints provably satisfied by the generated moduli is that the iterated encryption attack, first mentioned by Simmons and Norris [12], is infeasible. The generated numbers are claimed to be virtually uniformly distributed over the set of integers that lie in a given interval, are the product of exactly two primes and satisfy the security constraints. This claim is justified in Section 3 by considering the probability distribution of the size of the smaller prime factor of a number randomly selected from the set of integers in a given interval that are the product of exactly two primes, none of which is smaller than a given limit.

2. Distribution of the size of the largest prime factor of numbers n for which $2an+1$ is prime

In the following, we consider the probability that an integer n on the order of N satisfies certain conditions. Since "on the order of N" is not a mathematically precise condition we will instead use the condition $1 \leq n \leq N$ for stating definitions, equations and conjectures, although the arguments given as well as the intended application (generation of uniformly distributed primes) are often related to the former, less precise condition. However, for large N, both conditions are virtually equivalent. It is well-known that the fraction of primes among the integers on the order of N is well approximated by $1/\log N$. This fact can also be stated as

$$\pi(N) \sim \frac{N}{\log N} , \tag{1}$$

where here and in the sequel $\log x$ denotes the natural logarithm of x, $\pi(N)$ denotes the number of primes less or equal to N and $f(N) \sim g(N)$ stands for $\lim_{N\to\infty} f(N)/g(N) = 1$.

The arguments that will be used in this section are extensions of those given by

Koblitz [5] for heuristically justifying the conjecture that the asymptotic behaviour of the number $s(N)$ of primes $p \leq N$ for which $(p-1)/2$ is also prime is $s(N) \sim C_2 N/(\log N)^2$ where $C_2 \approx 0.66016$ is a well defined constant introduced later. Note that this conjecture is strongly related to the conjecture by Hardy and Littlewood [2] that the number $r(N)$ of prime pairs less than N satisfies $r(N) \sim 2C_2 N/(\log N)^2$. Like this conjecture, our conjectures are heuristically and numerically evident, but no rigorous proofs can be given. Note that although the Hardy-Littlewood conjecture and the Koblitz conjecture can experimentally very precisely be verified, even the questions whether there exist infinitely many prime pairs, and whether there exist infinitely many primes p for which $(p-1)/2$ is also a prime, respectively, are still open problems.

We first consider the probability distribution of the size of the largest prime factor of an integer n on the order of N. Let $p_1(n)$ denote the largest prime factor of n and let $\nu(N,k)$ denote the number of positive integers less or equal to N that are of the form "k times a prime", i.e.,

$$\nu(N,k) \;=\; \#\{n : 1 \leq n \leq N,\; n = kp \text{ with } p \text{ prime}\} \tag{2}$$

where $\#S$ denotes the cardinality of the set S. Note that $\nu(N,1) = \pi(N)$ and that

$$\nu(N,k) \;=\; \pi\left(\left\lfloor \frac{N}{k} \right\rfloor\right) \sim \frac{N}{k\log(N/k)} \;=\; \frac{N}{k(\log N - \log k)}. \tag{3}$$

Let $\omega(N,\alpha)$ be the number of integers less or equal to N for which the largest prime factor is not greater than N^α, i.e.,

$$\omega(N,\alpha) \;=\; \#\{n : 1 \leq n \leq N,\; p_1(n) \leq N^\alpha\} \tag{4}$$

The number of integers n less or equal to N and satisfying $p_1(n) = n/k$ for a given (small) k is well approximated by $\nu(N,k)$. Therefore the number of integers less or equal to N whose largest prime factor is greater than N^α, $N - \omega(N,\alpha)$, is well approximated by the sum of the number of integers of the form "k times a prime" over those k for which $k < N^{1-\alpha}$ and $p_1(k) \leq N/k$. The first condition on k is to guarantee that we only count integers with a prime factor greater or equal to N^α and the second condition is to prevent that we count certain integers twice: If k has a prime factor p greater than N/k, then the corresponding integers have been taken into account already for a smaller value of k (on the order of N/p). Hence

$$\omega(N,\alpha) \;\approx\; N - \sum_{\substack{k:\, 1 \leq k < N^{1-\alpha},\\ p_1(k) \leq N/k}} \nu(N,k)\;. \tag{5}$$

For $\alpha \geq 1/2$, $p_1(k) \leq N/k$ for all k with $1 \leq k < N^{1-\alpha}$ and hence the second condition on k can be omitted. For $\alpha < 1/2$, we can replace the condition $p_1(n) \leq N/k$ by taking into account the density $\omega(k, \log_k(N/k))/k = \omega(k, \log N/\log k - 1)/k$ of numbers k having no prime factor greater than N/k, and summing over all k with $1 \le k \le N^{1-\alpha}$. Therefore we have

$$\begin{aligned}
\omega(N,\alpha)/N &\approx 1 - \frac{1}{N} \sum_{\substack{k:\, 1 \le k < N^{1-\alpha},\\ p_1(k) \le N/k}} \nu(N,k) \\
&\approx 1 - \sum_{\substack{k:\, 1 \le k < N^{1-\alpha},\\ p_1(k) \le N/k}} \frac{1}{k(\log N - \log k)} \qquad (6)\\
&\approx 1 - \int_1^{N^{1-\alpha}} \frac{1}{k(\log N - \log k)} \cdot \frac{\omega(k, \log N/\log k - 1)}{k}\, dk \qquad (7)
\end{aligned}$$

where in the last step we have approximated the sum by an integral. For the moment we make the assumption, which can actually be proved (see theorem 1), that for large N, $\omega(N,\alpha)/N$ is virtually independent of N. Thus we can replace both $\omega(N,\alpha)/N$ and $\omega(k,\alpha)/k$ by $F_1(\alpha)$, a function of α only. Using the variable substitution $\xi = \log k/\log N$, which implies $d\xi = dk/(k \log N)$, (7) can be transformed into the integral equation

$$F_1(\alpha) = 1 - \int_0^{1-\alpha} \frac{F_1((1-\xi)/\xi)}{1-\xi}\, d\xi, \qquad (8)$$

where by convention $F_1(\alpha) = 1$ for $\alpha \ge 1$. For $\alpha \ge 1/2$ we have

$$F_1(\alpha) = 1 - \int_0^{1-\alpha} \frac{d\xi}{1-\xi} = -\log\alpha \,.$$

Knuth and Trabb Pardo [4] proved the following theorem which justifies the above heuristic derivation. The purpose of introducing the heuristic arguments, even though there exists a rigorous proof for the result, is both to convey intuitive evidence of the correctness and to prepare for later use of similar arguments, which will allow us to treat other problems (which are at present beyond theoretical tractability) in a heuristically evident way.

Theorem 1: *Let $F_1(\alpha)$ be defined by (8). Then, for $0 < \alpha \le 1$,*

$$\lim_{N\to\infty} \frac{\omega(N,\alpha)}{N} = F_1(\alpha).$$

$F_1(\alpha)$ is tabulated and depicted in [4]. For instance, the probability that the largest prime factor of a randomly selected integer is greater than its square root is $\log 2 = 0.69$, and the probability that it is smaller than the fourth root is less than 1%. It can be shown (see [4]) that the functions $F_t(\alpha)$ for $t \ge 2$, denoting the fraction of integers whose t-th largest prime factor is at most N^α, are given by similar integral equations.

In the following we repeat the above derivation for integers n on the order of N that satisfy the condition that $2n+1$ is prime, or, more generally, that $2an+1$ is prime. We introduce the following definitions which will be further generalized in section 3:

$$\begin{aligned} \pi_{[a]}(N) &= \#\{n : 1 \le n \le N,\ 2an+1 \text{ is prime}\}, \\ \nu_{[a]}(N,k) &= \#\{n : 1 \le n \le N,\ n = kp \text{ with } p \text{ prime},\ 2an+1 \text{ is prime}\} \\ \text{and} \quad \omega_{[a]}(N,\alpha) &= \#\{n : 1 \le n \le N,\ p_1(n) \le N^\alpha,\ 2an+1 \text{ is prime}\}. \end{aligned} \tag{9}$$

It is a well-known fact that

$$\pi_{[a]}(N) \ \sim\ \frac{2a}{\varphi(2a)} \cdot \frac{N}{\log N} \ =\ \frac{N}{\log N} \cdot \prod_{q|2a} \frac{q}{q-1} \tag{10}$$

for every integer $a \ge 1$, where $\varphi(.)$ denotes Euler's totient function and where the product is over all primes q dividing $2a$. Here and in the following, q always denotes a prime, and products are usually taken over all primes q satisfying certain constraints. An intuitive explanation for (10) is that for every prime q dividing $2a$ the probability is 1 that $2an+1$ is not divisible by q, whereas the same probability is only $(q-1)/q$ for random integers. Thus for every prime q dividing $2a$, a correction factor $q/(q-1)$ has to be taken into account.

The condition "$2n+1$ is prime" implies that $2n+1 \not\equiv 0 \ (mod\ q)$ for all odd primes $q < 2n+1$ and thus that

$$n \ \not\equiv\ -\frac{1}{2} \quad (mod\ q) \tag{11}$$

i.e., that $n \not\equiv 1 \ (mod\ 3)$, $n \not\equiv 2 \ (mod\ 5)$, $n \not\equiv 3 \ (mod\ 7)$, etc.. Since $n \equiv 0 \ (mod\ 3)$ or $n \equiv 2 \ (mod\ 3)$ with equal probability, the probability that a randomly selected n (with $2n+1$ prime) is not divisible by 3 is $1/2$, whereas it would be $2/3$ if no condition were placed on n. Similarly, for every odd prime $q < 2n+1$ the probabilities that, with and without condition, n is not divisible by q are $(q-2)/(q-1)$ and $(q-1)/q$, respectively. Taking into account these slight deviations in probability due to the condition "$2n+1$ is prime" for all odd primes, we obtain that a random integer n on the order of N, given that "$2n+1$ is prime", is prime with probability

$$P\left[n \text{ prime} \,\middle|\, 2n{+}1 \text{ prime}\right] \ \approx\ P[n \text{ prime}] \cdot \prod_{q \ge 3} \frac{P[q \nmid n \,|\, 2n{+}1 \text{ prime}]}{P[q \nmid n]} \ \approx\ \frac{C_2}{\log N}, \tag{12}$$

where $x \nmid y$ means that y is not divisible by x and where C_2 (this notation is in accordance with [2] and [5]) is a constant defined by

$$C_2 = \prod_{q \ge 3,\, q \text{ prime}} \frac{(q-2)/(q-1)}{(q-1)/q} \ =\ \prod_{q \ge 3,\, q \text{ prime}} \left(1 - \frac{1}{(q-1)^2}\right) \ \approx\ 0.660164\ . \tag{13}$$

Let us now consider the generalized condition that "$2an+1$ is prime", which implies that $2an+1 \not\equiv 0\ (mod\ q)$ for all odd primes $q < 2an+1$. For all odd primes q dividing a this inequality is trivially satisfied. For those q not dividing a we have

$$n \ \not\equiv \ -\frac{1}{2a} \quad (mod\ q) \tag{14}$$

and hence the same argument applies as given above for the case $a = 1$. We have

$$P\left[n \text{ prime} \,\middle|\, 2an{+}1 \text{ prime}\right] \ \approx \ \frac{1}{\log N} \prod_{q\geq 3,\ q\not| a} \frac{(q-2)/(q-1)}{(q-1)/q} \ = \ \frac{B_{[a]}(1)}{\log N} \tag{15}$$

where $B_{[a]}(1)$ is defined by

$$B_{[a]}(1) \ = \prod_{q\geq 3:\ q\not| a} \frac{q(q-2)}{(q-1)^2} \ = \ C_2 \prod_{q\geq 3:\ q|a} \frac{(q-1)^2}{q(q-2)} \ . \tag{16}$$

The notation $B_{[a]}(1)$ is in accordance with two later generalizations. Using (10) and the fact that $\lim_{N\to\infty} \log(2aN)/\log N = 1$ we state the following conjecture about the density of primes p for which $2ap+1$ is also prime.

Conjecture 1: *For every $a \geq 1$,*

$$\lim_{N\to\infty} \frac{\nu_{[a]}(N,1)\log N}{\pi_{[a]}(N)} \ = \ B_{[a]}(1)\ , \qquad i.e., \qquad \nu_{[a]}(N,1) \ \sim \ A_{[a]}(1)\frac{N}{(\log N)^2}\ ,$$

where $A_{[a]}(1)$ is defined by

$$A_{[a]}(1) \ = \ \frac{2aB_{[a]}(1)}{\varphi(2a)} \ = \ 2C_2 \prod_{q\geq 3:\ q|a} \frac{q-1}{q-2}\ . \tag{17}$$

The conditional probability that a number n on the order of N is exactly k times a prime for a given k, given that $2an+1$ is prime, is

$$\begin{aligned} P\left[n = kp \text{ with } p \text{ prime} \,\middle|\, 2an{+}1 \text{ prime}\right] \ &= \\ P\left[k|n \,\middle|\, 2an{+}1 \text{ prime}\right] \cdot P\left[n/k \text{ is prime} \,\middle|\, k|n,\ 2an{+}1 \text{ prime}\right]&. \end{aligned} \tag{18}$$

For all odd primes q dividing k but not dividing a, inequality (14) implies that $P[q|n\,|\,2an{+}1 \text{ prime}] \approx 1/(q-1)$. Without the condition "$2an+1$ is prime" the same probability would be $1/q$. Therefore we have

$$P\left[k|n \,\middle|\, 2an{+}1 \text{ prime}\right] \ \approx \ \frac{1}{k}\prod_{q\geq 3:\ q|k} \frac{P[q|n\,|\,2an{+}1 \text{ prime}]}{P[q|n]} \ \approx \ \frac{1}{k}\prod_{q\geq 3:\ q\not| a,\ q|k} \frac{q}{q-1}\ . \tag{19}$$

The term $P[n/k \text{ is prime} \mid k|n,\ 2an+1 \text{ prime}]$ is equal to $P[m \text{ is prime} \mid 2akm+1 \text{ is prime}]$ where m is on the order of N/k. Using (15), (16), (18) and (19) therefore yields

$$\begin{aligned}
P\left[n = kp \text{ with } p \text{ prime} \,\middle|\, 2an+1 \text{ prime}\right] &\approx \tfrac{1}{k} \prod_{q\geq 3,\, q\not| a,\, q|k} \tfrac{q}{q-1} \cdot \tfrac{B_{[ak]}(1)}{\log(N/k)} \\
&= \frac{C_2}{k\log(N/k)} \prod_{q\geq 3:\ q\not| a, q|k} \frac{q}{q-1} \cdot \prod_{q\geq 3:\ q|ak} \frac{(q-1)^2}{q(q-2)} \\
&= \frac{B_{[a]}(k)}{k\log(N/k)}, \qquad (20)
\end{aligned}$$

where $B_{[a]}(k)$ is defined by

$$B_{[a]}(k) \;=\; B_{[a]}(1) \prod_{q\geq 3:\ q|k,\ q\not| a} \frac{q-1}{q-2}\ . \qquad (21)$$

The above derivation shows that the condition "$2an+1$ is prime" changes the probability that an integer n is prime by a factor $B_{[a]}(1)$, and more generally, the probability that n is of the form "k times a prime" by a factor $B_{[a]}(k)$. For instance, for $a = 1$ we have $B_{[1]}(1) = B_{[1]}(2) = C_2 = 0.6602$, $B_{[1]}(3) = 2C_2 = 1.3203$, $B_{[1]}(4) = C_2 = 0.6602$, $B_{[1]}(5) = 4/3 \cdot C_2 = 0.8802$, etc., and for $a = 3$ we have $B_{[3]}(1) = B_{[3]}(2) = B_{[3]}(3) = B_{[3]}(4) = 4/3 \cdot C_2 = 0.8802$, $B_{[3]}(5) = 1.1736$, $B_{[3]}(6) = 0.8802$, $B_{[3]}(7) = 1.0563$, etc.. The density of primes among the numbers n for which $2n+1$ is prime is thus only roughly $2/3$ of the density of primes. Because $P[2an+1 \text{ prime}] \approx 1/\log(2aN)$ we have

$$\nu_{[a]}(N,k) \;\approx\; \frac{B_{[a]}(k)\cdot 2a\cdot N}{k\log(N/k)\varphi(2a)\log(2aN)}$$

which leads to the following generalization of conjecture 1.

Conjecture 2: *For every $a \geq 1$ and $k \geq 1$,*

$$\lim_{N\to\infty} \frac{\nu_{[a]}(N,k)\log N}{\pi_{[a]}(N)} \;=\; \frac{B_{[a]}(k)}{k}\ , \quad \textit{i.e.,} \quad \nu_{[a]}(N,k) \;\sim\; A_{[a]}(k)\frac{N}{k(\log N)^2}\ ,$$

where $A_{[a]}(k)$ is defined by

$$A_{[a]}(k) \;=\; A_{[a]}(1) \prod_{q\geq 3:\ q|k,\ q\not| a} \frac{q-1}{q-2} \;=\; 2C_2 \prod_{q\geq 3:\ q|ak} \frac{q-1}{q-2}\ .$$

In order to estimate the influence of the condition "$2an+1$ is prime" on the distribution of the size of the largest prime factor of n, we adapt equations (5) to (8) to take this condition into account:

$$
\begin{aligned}
\omega_{[a]}(N,\alpha)/N &= 1-\tfrac{1}{N}\textstyle\sum_{\substack{k:\,1\le k<N^{1-\alpha},\\ p_1(k)\le N^\alpha}} \nu_{[a]}(N,k) \\
&\approx 1-\sum_{\substack{k:\,1\le k<N^{1-\alpha},\\ p_1(k)\le N^\alpha}} \frac{2aB_{[a]}(k)}{k\varphi(2a)\log(N/k)\log(2aN)} \\
&\approx \frac{2a}{\varphi(2a)\log(2aN)}\int_1^{N^{1-\alpha}} \frac{B_{[a]}(k)}{k(\log N-\log k)}\cdot\frac{\omega(k,\log N/\log k-1)}{k}\,dk \\
&\approx E_k[B_{[a]}(k)]\,\frac{2a}{\varphi(2a)\log(2aN)}\int_1^{N^{1-\alpha}} \frac{1}{k(\log N-\log k)}\cdot\frac{\omega(k,\log N/\log k-1)}{k}\,dk \\
&\approx E_k[B_{[a]}(k)]\,\frac{2a}{\varphi(2a)\log(2aN)}\,F_1(\alpha)\ . \qquad (22)
\end{aligned}
$$

where $E_k[f(k)]$ denotes the expectation of $f(k)$ when k is a random positive integer. We will show that the effect of the terms $B_{[a]}(k)$ in the above derivation averages out by proving the following theorem.

Theorem 2: *For every $a\ge 1$, $E_k[B_{[a]}(k)] \;=\; 1$.*

This theorem together with (22) suggests that the probability distribution of the size of the largest prime factor of a random integer n is virtually independent of the condition "$2an+1$ is prime", which leeds to the following conjecture.

Conjecture 3: *For every $a\ge 1$ and for $0<\alpha\le 1$,*

$$
\lim_{N\to\infty}\frac{\omega_{[a]}(N,\alpha)}{\pi_{[a]}(N)} \;=\; F_1(\alpha)\ , \qquad \text{i.e.,}\qquad \omega_{[a]}(N,\alpha)\;\sim\;\frac{2aF_1(\alpha)}{\varphi(2a)}\frac{N}{\log N}.
$$

Proof of Theorem 2: Let q_i denote the i-th odd prime and let X_i be the the indicator random variable for the event that $q_i|k$, i.e., $X_i=1$ if $q_i|k$ and $X_i=0$ else. Hence $P_{X_i}(0) \;=\; (q_i-1)/q_i$ and $P_{X_i}(1) \;=\; 1/q_i$. The sequence $X_1,X_2,\ldots$ is an infinite sequence of statistically independent binary random variables. Define the function $f_a:\; N\times\{0,1\}\longrightarrow Q$ by

$$
f_a(i,x) \;=\; \begin{cases} \dfrac{q_i-1}{q_i-2} & \text{if } x=1 \text{ and } q_i\not|a \\ \quad 1 & \text{else,}\end{cases}
$$

where N and Q denote the positive integers and the rational numbers, respectively. Then we have

$$
E_k[B_{[a]}(k)] \;=\; B_{[a]}(1)\cdot E_k\left[\prod_{q\ge 3:\;q|k,\;q\not|a}\frac{q-1}{q-2}\right] \;=\; B_{[a]}(1)\cdot E_k\left[\prod_{i=1}^{\infty} f_a(i,X_i)\right]
$$

$$
\begin{aligned}
&= B_{[a]}(1) \lim_{l\to\infty} \sum_{[x_1,\dots,x_l]\in\{0,1\}^l} \prod_{i=1}^{l} P_{X_i}(x_i) f_a(i,x_i) \\
&= B_{[a]}(1) \lim_{l\to\infty} \prod_{i=1}^{l} \sum_{x\in\{0,1\}} P_{X_i}(x) f_a(i,x)
\end{aligned}
$$

where the last step follows from the fact that product and summation can be interchanged. Using

$$
\sum_{x\in\{0,1\}} P_{X_i}(x) f_a(i,x) \;=\; \begin{cases} \dfrac{q_i-1}{q_i}+\dfrac{1}{q_i}=1 & \text{if } q_i|a \\[2ex] \dfrac{q_i-1}{q_i}+\dfrac{1}{q_i}\dfrac{q_i-1}{q_i-2}=\dfrac{(q_i-1)^2}{q_i(q_i-2)} & \text{if } q_i \not| a \end{cases}
$$

as well as equation (16) we obtain

$$
E_k[B_{[a]}(k)] \;=\; B_{[a]}(1) \prod_{q\ge 3:\; q\not| a} \frac{(q-1)^2}{q(q-2)} \;=\; 1. \qquad \square
$$

3. More conditions on the numbers n

In this section, we treat the same problems as in the previous section, but with more conditions placed on the integers n. For n randomly selected from the set of integers "on the order of N" that satisfy these conditions, the probability that n is of the form "k times a prime" and the probability distribution of the size of the largest prime factor of n are discussed. The conditions on n, which are motivated by the analysis of the prime generating procedure of [6], are that for given integers $a_1, a_2, \dots a_d$, the numbers

$$
\begin{aligned}
u_1 &= 2a_1 n + 1, \\
u_2 &= 2a_2 u_1 + 1 = 4a_1a_2 n + 2a_2 + 1, \\
u_3 &= 2a_3 u_2 + 1 = 8a_1a_2a_3 n + 4a_2a_3 + 2a_3 + 1, \\
&\dots \\
u_d &= 2a_d u_{d-1} + 1,
\end{aligned}
$$

all are primes. More formally, we say that n satisfies the condition $Q_{\underline{a}}(n)$ for a given list $\underline{a} = [a_1, \dots, a_d]$ of positive integers if and only if the d integers $u_1, \dots, u_d$ are primes:

$$
Q_{\underline{a}}(n) \iff u_t(\underline{a}) \text{ is prime for } 1 \le t \le d, \tag{23}
$$

where $u_t(\underline{a})$ is defined recursively by $u_0(\underline{a}) = n$ and

$$
u_t(\underline{a}) \;=\; 2a_t u_{t-1}(\underline{a}) + 1 \qquad \text{for } 1 \le t \le d. \tag{24}
$$

In the following we use the notation

$$u_t(\underline{a}) = r_t(\underline{a})\, n + s_t(\underline{a}), \quad \text{where} \quad r_t(\underline{a}) = 2^t \prod_{i=1}^{t} a_i \tag{25}$$

for $1 \le t \le d$, and where $s_t(\underline{a})$ is defined recursively by $s_1(\underline{a}) = 1$ and

$$s_t(\underline{a}) \;=\; 2a_t s_{t-1}(\underline{a}) + 1 \quad \text{for } 2 \le t \le d. \tag{26}$$

The number $u_t(\underline{a})$ is prime if and only if

$$u_t(\underline{a}) \;=\; r_t(\underline{a})\, n + s_t(\underline{a}) \;\not\equiv\; 0 \;\; (mod\ q) \tag{27}$$

for all odd primes q smaller that $u_t(\underline{a})$. If for some prime q and some t we have $r_t(\underline{a}) \equiv s_t(\underline{a}) \equiv 0$ $(mod\ q)$ or, equivalently, if $(r_t(\underline{a}), s_t(\underline{a})) \ne 1$ for some t, then there exists no integer n satisfying $Q_{\underline{a}}(n)$. A list $\underline{a} = [a_1, \dots, a_d]$ for which $Q_{\underline{a}}(n)$ is satisfied by some positive integer n is called *admissible*, otherwise it is called *non-admissible*. Examples of non-admissible lists are $[3,1]$, $[3,2,2]$ and $[5,2]$. In the sequel we only consider lists $\underline{a} = [a_1, \dots, a_d]$ for which

$$(r_t(\underline{a}), s_t(\underline{a})) \;=\; 1 \quad \text{for } 1 \le t \le d. \tag{28}$$

Let $d_q(\underline{a})$ denote the greatest integer $i \le d$ such that q does not divide a_i, i.e.,

$$d_q(\underline{a}) \;=\; \max\{i : 0 \le i \le d,\ q \not| a_i\}. \tag{29}$$

The arguments given in the following are valid for all odd primes $q < u_1(\underline{a})$. For $t > d_q(\underline{a})$, i.e., if $q|a_t$, equation (27) is trivially satisfied if (28) is satisfied. For $t \le d_q(\underline{a})$, which guarantees that $r_t(\underline{a}) \not\equiv 0$ $(mod\ q)$, the condition $Q_{\underline{a}}(n)$ implies that

$$n \;\not\equiv\; -\frac{s_t(\underline{a})}{r_t(\underline{a})} \;\; (mod\ q). \tag{30}$$

and hence that the remainder $R_q(n)$ of n modulo q satisfies

$$R_q(n) \;\notin\; S_q(\underline{a}) \;=\; \left\{ -\frac{s_t(\underline{a})}{r_t(\underline{a})}\ (mod\ q) : \; 1 \le t \le d_q(\underline{a}) \right\}. \tag{31}$$

Another reason for a list $\underline{a}$ to be non-admissible is the existence of an odd prime q such that $\#S_q(\underline{a}) = q$, i.e., such that all remainders modulo q of n are ruled out by the conditions (30). Examples of such lists $\underline{a}$ are $[1,2,2]$, $[2,1,2]$, $[2,2,2]$ and $[2,3,3,3,3]$.

Let us now consider for a given $\underline{a}$ the conditional probability that n is prime, given $Q_{\underline{a}}(n)$. In the following we only consider admissible lists $\underline{a}$. For every odd prime q we can distinguish between two cases: $0 \in S_q(\underline{a})$ and $0 \notin S_q(\underline{a})$. The probability that n is not divisible by q, given $Q_{\underline{a}}(n)$, is 1 or $(q - \#S_q(\underline{a}) - 1)/(q - \#S_q(\underline{a}))$ depending on $0 \in S_q(\underline{a})$ or $0 \notin S_q(\underline{a})$, respectively, compared to $(q-1)/q$ when no condition is placed on n. Therefore we have

$$P[n \text{ prime} \,|\, Q_{\underline{a}}(n)] \;\approx\; P[n \text{ prime}] \cdot \prod_{q\geq 3} \frac{P[q \not| n \,|\, Q_{\underline{a}}(n)]}{P[q \not| n]} \;\approx\; \frac{B_{\underline{a}}(1)}{\log N}, \tag{32}$$

where $B_{\underline{a}}(1)$ is defined (in accordance with the generalization $B_{\underline{a}}(k)$ considered below) by

$$B_{\underline{a}}(1) \;=\; \prod_{q\geq 3:\, 0\notin S_q(\underline{a})} \frac{q(q-\#S_q(\underline{a})-1)}{(q-1)(q-\#S_q(\underline{a}))} \prod_{q\geq 3:\, 0\in S_q(\underline{a})} \frac{q}{q-1}. \tag{33}$$

The natural generalizations of definitions (9) are

$$\begin{aligned} \pi_{\underline{a}}(N) &= \#\{n : 1 \leq n \leq N,\ n \text{ is prime},\ Q_{\underline{a}}(n)\}, \\ \nu_{\underline{a}}(N,k) &= \#\{n : 1 \leq n \leq N,\ n = kp \text{ with } p \text{ prime},\ Q_{\underline{a}}(n)\} \\ \text{and}\quad \omega_{\underline{a}}(N,\alpha) &= \#\{n : 1 \leq n \leq N,\ p_1(n) \leq N^{\alpha},\ Q_{\underline{a}}(n)\}, \end{aligned} \tag{34}$$

and the corresponding generalization of conjecture 1 is

Conjecture 4: *For every* $\underline{a} = [a_1, \ldots, a_d]$,

$$\lim_{N\to\infty} \frac{\nu_{\underline{a}}(N,1)\log N}{\pi_{\underline{a}}(N)} \;=\; B_{\underline{a}}(1).$$

Bateman and Horn [1] have considered the conditional probability that a random integer is prime, given that it simultaneously satisfies a set of polynomial equations. Conjectures 1 and 4 can be regarded as a special case of their conjecture, but the other conjectures are generalizations of those in [1].

Let us now consider for given $k \geq 1$ and $\underline{a} = [a_1, \ldots, a_d]$ the conditional probability that $n = kp$ with p prime, given $Q_{\underline{a}}(n)$:

$$P\Big[n = kp \text{ with } p \text{ prime} \,\Big|\, Q_{\underline{a}}(n)\Big] \;=\; P\Big[k|n \,\Big|\, Q_{\underline{a}}(n)\Big] \cdot P\Big[n/k \text{ is prime} \,\Big|\, k|n,\ Q_{\underline{a}}(n)\Big]. \tag{35}$$

Conditions (31) imply that for all odd primes q dividing k, $P[q|n \,|\, Q_{\underline{a}}(n)] = 0$ and $P[q|n \,|\, Q_{\underline{a}}(n)] \approx 1/(q-\#S_q(\underline{a}))$ when $0 \in S_q(\underline{a})$ and $0 \notin S_q(\underline{a})$, respectively. Therefore we have

$$\begin{aligned} P[k|n \,|\, Q_{\underline{a}}(n)] &\approx \frac{1}{k} \prod_{q\geq 3:\, q|k} \frac{P[q|n \,|\, Q_{\underline{a}}(n)]}{P[q|n]} \\ &\approx \frac{1}{k} \prod_{q\geq 3:\, q|k} \left\{ \begin{array}{cl} 0 & \text{if } 0 \in S_q(\underline{a}) \\ \dfrac{q}{q-\#S_q(\underline{a})} & \text{if } 0 \notin S_q(\underline{a}) \end{array} \right\}. \end{aligned} \tag{36}$$

The term $P[n/k \text{ is prime} \mid k|n, Q_{\underline{a}}(n)]$ is equal to $P[m \text{ is prime} \mid Q_{k*\underline{a}}(m)]$, where $k * \underline{a} = [ka_1, a_2, \ldots, a_d]$ and where m is on the order of N/k. Using (32), (35) and (36) therefore yields

$$P[n = kp \text{ with } p \text{ prime} \,|\, Q_{\underline{a}}(n)] \;\approx\; \frac{B_{\underline{a}}(k)}{k \log(N/k)} \;, \tag{37}$$

where

$$B_{\underline{a}}(k) \;=\; \prod_{q\geq 3:\; q|k} \left\{ \begin{array}{cc} 0 & \text{if } 0 \in S_q(\underline{a}) \\ \dfrac{q}{q - \#S_q(\underline{a})} & \text{if } 0 \notin S_q(\underline{a}) \end{array} \right\} B_{k*\underline{a}}(1) \;. \tag{38}$$

It is not difficult to verify that when $q|k$, then $\#S_q(k*\underline{a}) = 0$. Moreover, $q|k$ together with $0 \in S_q(\underline{a})$ implies that $k*\underline{a}$ is non-admissible. Similarly, when $q \not| k$, then we have $0 \in S_q(\underline{a}) \Longleftrightarrow 0 \in S_q(k*\underline{a})$ and $\#S_q(\underline{a}) = \#S_q(k*\underline{a})$. Thus

$$B_{k*\underline{a}}(1) \;=\; \prod_{q\geq 3:\; q|k} \left\{ \begin{array}{ll} 0 & \text{if } 0 \in S_q(\underline{a}) \\ 1 & \text{if } 0 \notin S_q(\underline{a}) \end{array} \right\} \prod_{q\geq 3:\; q\not|k} \left\{ \begin{array}{cc} \dfrac{q}{q-1} & \text{if } 0 \in S_q(\underline{a}) \\ \dfrac{q(q-\#S_q(\underline{a})-1)}{(q-1)(q-\#S_q(\underline{a}))} & \text{if } 0 \notin S_q(\underline{a}) \end{array} \right\} . \tag{39}$$

Combining (38) and (39) yields

$$B_{\underline{a}}(k) = \prod_{q\geq 3:\; q|k} \left\{ \begin{array}{cc} 0 & \text{if } 0 \in S_q(\underline{a}) \\ \dfrac{q}{q-\#S_q(\underline{a})} & \text{if } 0 \notin S_q(\underline{a}) \end{array} \right\} \prod_{q\geq 3:\; q\not|k} \left\{ \begin{array}{cc} \dfrac{q}{q-1} & \text{if } 0 \in S_q(\underline{a}) \\ \dfrac{q(q-\#S_q(\underline{a})-1)}{(q-1)(q-\#S_q(\underline{a}))} & \text{if } 0 \notin S_q(\underline{a}) \end{array} \right\} . \tag{40}$$

Conjectures 2 and 4 generalize to the following conjecture.

Conjecture 5: *For every* $\underline{a} = [a_1, \ldots, a_d]$,

$$\lim_{N\to\infty} \frac{\nu_{\underline{a}}(N,k) \log N}{\pi_{\underline{a}}(N)} \;=\; \frac{B_{\underline{a}}(k)}{k}.$$

It is possible to generalize the second half of conjecture 2 to obtain

$$\nu_{\underline{a}}(N,k) \;\sim\; A_{\underline{a}}(k) \frac{N}{k(\log N)^{d+1}}$$

where an expression similar to that for $A_{[a]}(k)$, but much more complicated, can be given for $A_{\underline{a}}(k)$. Since the aim of this paper is to consider only conditional probability distributions, the expressions for $A_{\underline{a}}(k)$ are not derived here.

Although the expression (40) for $B_{\underline{a}}(k)$ is rather complicated, the following theorem can be proved (see Appendix for a proof). The subsequent conjecture, which is based on theorem 3 and can be obtained by a derivation similar to that of (22), is the straight-forward generalization of conjecture 3.

Theorem 3: *For every* $\underline{a} = [a_1, \ldots, a_d]$, $E_k[B_{\underline{a}}(k)] \;=\; 1.$

Conjecture 6: *For every* $\underline{a} = [a_1, \ldots, a_d]$ *and for* $0 < \alpha \leq 1$,

$$\lim_{N\to\infty} \frac{\omega_{\underline{a}}(N,\alpha)}{\pi_{\underline{a}}(N)} \;=\; F_1(\alpha)\ .$$

4. The distribution of the size of the smaller RSA-prime

The method for generating primes is in [6] also used to devise a method for generating random RSA-moduli, i.e., for randomly selecting an integer $n = pq$ (with virtually uniform distribution) from the set of integers in a given interval that are the product of exactly two primes, are compatible with a predefined encryption exponent, and that satisfy certain security requirements. One of these requirements is that neither of the prime factors must be too small. In this section, we consider the density of integers on the order of N that are the product of exactly 2 primes, neither of which is smaller than N^γ for a given γ.

Let $\mu(n)$ denote the number of prime factors (counting multiple occurrences) of n, $\overline{p}(n)$ the smallest prime factor of n, and let

$$\begin{aligned} \rho_l(N) &\;=\; \#\{n:\ 1\leq n\leq N,\ \mu(n) = l\} && (41)\\ \text{and}\quad \tau_2(N,\gamma) &\;=\; \#\{n:\ 1\leq n\leq N,\ \mu(n) = 2,\ \overline{p}(n) \geq N^\gamma\}. && (42)\end{aligned}$$

Hence $\rho_2(N) = \tau_2(N,0)$ denotes the number of integers less or equal to N that are the product of exactly two primes. We have

$$\begin{aligned} \tau_2(N,\gamma)/N &\;\approx\; \sum_{\substack{k:\ N^\gamma \leq k < N^{1/2},\\ k \text{ prime}}} \frac{\nu(N,k)}{N} \;\approx\; \sum_{\substack{k:\ N^\gamma \leq k < N^{1/2},\\ k \text{ prime}}} \frac{1}{k\log(N/k)}\\ &\;\approx\; \int_{N^\gamma}^{N^{1/2}} \frac{1}{k\log(N/k)}\cdot\frac{1}{\log k}\,dk\\ &\;=\; \frac{1}{\log N}\int_\gamma^{1/2}\frac{d\xi}{\xi(1-\xi)} \;=\; \frac{\log(1-\gamma)-\log\gamma}{\log N} \qquad (43)\end{aligned}$$

where we have approximated the sum by an integral and taken into account the condition "k prime" by introducing the density $1/\log k$ of primes on the order of k.

The probability distribution of the size of the smaller prime of an RSA-modulus on the order of N, given that none of the prime factors is smaller than N^γ, is given by

$$\begin{aligned} P\left[\overline{p}(n) < N^\alpha \,\middle|\, \mu(n) = 2,\ \overline{p}(n) \geq N^\gamma\right] &= 1 - \frac{\tau_2(N,\alpha)}{\tau_2(N,\gamma)} \\ &\approx 1 - \frac{\log(1-\alpha) - \log\alpha}{\log(1-\gamma) - \log\gamma}. \end{aligned} \tag{44}$$

For $\gamma > 0.4$, the relative size of the smaller prime factor of the RSA-modulus is virtually uniformly distributed over the interval $[\gamma, 1/2]$. Based on (44) we state the following conjecture.

Conjecture 7: *For* $0 < \gamma \leq 1$,

$$\tau_2(N,\gamma) \sim \frac{N}{\log N}(\log(1-\gamma) - \log\gamma)\ .$$

Since $\rho_2(N) = \#\{n:\ 1 \leq n \leq N,\ \mu(n) = 2,\ \overline{p}(n) \geq 2\} = \tau_2(N, \log 2/\log N)$, a special case of conjecture 7 is the conjecture that $\rho_2(N) \sim N \log\log N/\log N$. In fact, this conjecture, and also its generalization from 2 to l prime factors, can be proved (see [3], Theorem 437).

Theorem 4: *For* $l \geq 2$, $\quad \rho_l(N) \sim \dfrac{N(\log\log N)^{l-1}}{(l-1)!\log N}$.

5. Conclusions

The claim that the procedure for generating RSA-moduli described in [6] selects virtually randomly from the set of all RSA-moduli of a given size satisfying certain security constraints has been justified by heuristic arguments. Some heuristically and numerically evident number-theoretic conjectures have been stated. Their evidence is supported by the fact that some special cases can be proved.

Appendix: Proof of Theorem 3

This proof is similar to that of Theorem 2. Equation (40) can be rewritten as

$$\begin{aligned} B_{\underline{a}}(k) &= \prod_{q\geq 3:0\in S_q(\underline{a})} \left\{ \begin{array}{cl} 0 & \text{if } q|k \\ \dfrac{q}{q-1} & \text{if } q \not| k \end{array} \right\} \prod_{q\geq 3:0\notin S_q(\underline{a})} \left\{ \begin{array}{cl} \dfrac{q}{q-\#S_q(\underline{a})} & \text{if } q|k \\ \dfrac{q(q-\#S_q(\underline{a})-1)}{(q-1)(q-\#S_q(\underline{a}))} & \text{if } q \not| k \end{array} \right\} \\ &= \prod_{i=1}^{\infty} g_{\underline{a}}(i, X_i) \end{aligned}$$

where the random variables $X_1, X_2, \ldots$ are defined in the proof of theorem 2 and where

$$g_{\underline{a}}(i,x) = \begin{cases} 0 & \text{if } 0 \in S_{q_i}(\underline{a}) \text{ and } x = 0 \\ \dfrac{q_i}{q_i - 1} & \text{if } 0 \in S_{q_i}(\underline{a}) \text{ and } x = 1 \\ \dfrac{q_i}{q_i - \#S_{q_i}(\underline{a})} & \text{if } 0 \notin S_{q_i}(\underline{a}) \text{ and } x = 0 \\ \dfrac{q_i(q_i - \#S_{q_i}(\underline{a}) - 1)}{(q_i - 1)(q_i - \#S_{q_i}(\underline{a}))} & \text{if } 0 \notin S_{q_i}(\underline{a}) \text{ and } x = 1. \end{cases}$$

Hence

$$\begin{aligned} E_k[B_{\underline{a}}(k)] &= E_k\left[\prod_{i=1}^{\infty} g_{\underline{a}}(i, X_i)\right] \\ &= \lim_{l\to\infty} \sum_{[x_1,\ldots,x_l]\in\{0,1\}^l} \prod_{i=1}^{l} P_{X_i}(x_i) g_{\underline{a}}(i, x_i) \\ &= \lim_{l\to\infty} \prod_{i=1}^{l} \sum_{x\in\{0,1\}} P_{X_i}(x) g_{\underline{a}}(i, x) \end{aligned}$$

where the last step follows from the fact that product and summation can be interchanged. It is easy to verify that

$$\sum_{x\in\{0,1\}} P_{X_i}(x) g_{\underline{a}}(i, x) = 1,$$

independently of whether $0 \in S_{q_i}(\underline{a})$ or $0 \notin S_{q_i}(\underline{a})$. This completes the proof. $\Box$

References

[1] P.T. Bateman and R.A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp., vol. 16, pp. 363-367, 1962.

[2] G.H. Hardy and J.E. Littlewood, *Some problems of 'partitio numerorum'; III: on the expression of a number as a sum of primes*, Acta Math., vol. 44, pp. 1-70, 1923.

[3] G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers*, fourth ed., Oxford University Press, 1960.

[4] D.E. Knuth and L. Trabb Pardo, *Analysis of a simple factorization algorithm*, Theoretical Computer Science, vol. 3, pp. 321-348, 1976.

[5] N. Koblitz, *Primality of the number of points on an elliptic curve over a finite field*, Pacific Journal of Mathematics, vol. 131, no. 1, pp. 157-165, 1988.

[6] U.M. Maurer, *Fast generation of secure RSA-moduli with almost maximal diversity*, to appear in Proc. EUROCRYPT'89, in Lecture Notes in Computer Science, New York, NY: Springer Verlag.

[7] H.C. Pocklington, *The determination of the prime or composite nature of large numbers by Fermat's theorem*, Proc. Cambridge Philos. Soc., vol. 18, pp. 29-30, 1914-1916.

[8] M.O. Rabin, *Probabilistic algorithm for testing primality*, Journal on Number Theory, vol. 12, pp. 128-138, l980.

[9] R.L. Rivest, *Remarks on a proposed cryptanalytic attack on the M.I.T. public key cryptosystem*, Cryptologia, vol. 2, no. 1, pp. 62-65, Jan. 1978.

[10] R.L. Rivest, A. Shamir, and L. Adleman, *A method for obtaining digital signatures and public-key cryptosystems*, Comm. ACM, vol. 21, pp.120-126, Feb. 1978.

[11] J. Shawe-Taylor, *Generating strong primes*, Electronics Letters, vol. 22, no. 16, pp. 875-877, July 1986.

[12] G. Simmons and M. Norris, *Preliminary comments on the M.I.T public key cryptosystem*, Cryptologia, vol. 1, no. 4, pp. 406-414, Oct. 1977.

[13] R. Solovay and V. Strassen, *A fast Monte-Carlo test for primality*, SIAM J. Computing, vol. 6, no. 1, pp. 84-85, March 1977.