Electronic Colloquium on Computational Complexity - Reports Series 1997 - available via:
ECCC
TR97-027
FTP: ftp.eccc.uni-trier.de:/pub/eccc/
WWW: http://www.eccc.uni-trier.de/eccc/
Email: ftpmail@ftp.eccc.uni-trier.de with subject 'help eccc'

On the Security of Server Aided RSA Protocols

J. Merkle and R. Werchner
Fachbereich Mathematik
Universität Frankfurt
60054 Frankfurt
Germany
e-mail: merkle@cs.uni-frankfurt.de

April 29, 1997

Abstract

In this paper we investigate the security of the server aided RSA protocols RSA-S1 and RSA-S1M proposed by Matsumoto, Kato and Imai resp. Matsumoto, Imai, Laih and Yen. We prove lower bounds for the complexity of attacks on these protocols and show that the bounds are sharp by describing attacks that almost match our lower bounds. To the best of our knowledge these are the first lower bounds for efficient server aided RSA protocols.

1 Introduction

In this paper, we investigate the security of server-aided secret computations of RSA signatures. Consider the following scenario: Let n be an RSA modulus and d, e be a pair of private/public exponents. A smart card stores n and d and needs to sign a message x by computing $y = x^d \bmod n$. This takes $O(\log n)$ multiplications modulo n, which is a heavy task for the card. The solution proposed by Matsumoto, Kato and Imai in [MKI89] is server-aided secret computation (SASC). In their protocol RSA-S1 the main part of the computation is done by a more powerful server.

The RSA-S1 Protocol

0. (Preprocessing step) The card chooses an integer vector $\mathbf{d} = (d_1, \ldots, d_m) \in \mathbf{Z}^m$ and a vector $\mathbf{f} = (f_1, \ldots, f_m) \in \{0,1\}^m$ with Hamming weight k so that $d = \sum_{i=1}^m f_i d_i \bmod \phi(n)$.
1. The card sends x, n and $\mathbf{d}$ to the server.
2. The server returns $z_i = x^{d_i} \bmod n$ for $i = 1, \ldots, m$.
3. The card computes the signature y as $y = \prod_{i=1}^m z_i^{f_i} \bmod n$.

Consequently Matsumoto, Imai, Laih and Yen proposed a two-phase version of this protocol, called RSA-S1M ([MILY93]).

There exist two kinds of attacks against such protocols: classical searching ones are called *passive attacks*; specific ones where the server returns false values to get some information from the card are called *active attacks*. In [LL95] Lim and Lee showed, that active attacks can be avoided efficiently. In this paper we focus on the computational complexity of *generic attacks* against RSA-S1 and RSA-S1M — that is, passive attacks, which do not exploit any special properties of the encoding of the elements of $\mathbf{Z}_n$. A similar class of algorithms has been analysed by Shoup ([Sho]). He proved lower bounds for the complexity of generic algorithms for the discrete log and related problems.

We prove average case lower bounds of $\Omega(N^{1/2})$ and $\Omega(N^{2/3})$ for generic attacks against RSA-S1 resp. RSA-S1M. Here N is the number of possible cointosses of the card. The bound for attacks against RSA-S1 is asymptotically sharp. In addition we prove a lower bound of $\Omega(N^{3/4})$ for meet-in-the-middle attacks against RSA-S1M, which have been the most successful attacks so far. This bound is also asymptotically sharp.

Kawamura and Shimbo ([KS93]) and Quisquater and de Soete ([QS91]) proposed rather different SASC protocols. There the communication between the server and the card is independent of the secret key and consequently they are as secure against passive attacks as RSA. Unfortunately these protocols are not very efficient. Our results are the first security proofs for efficient SASC protocols.

The structure of the paper is as follows: In section 2 we define SASC protocols and the model of computation. In sections 3 and 4 we review the considered protocols and investigate their security. Finally in section 5 we discuss our results giving some concrete examples and shortly discuss active attacks.

2 Model of Computation

2.1 Server Aided Secret Computation

Let n be a product of two large primes, $d \in \mathbf{Z}^*_{\phi(n)}$ and $x \in \mathbf{Z}^*_n$. Furthermore let the security parameters of the protocol be fixed. In order to compute $x^d \bmod n$ using the help of a server the client carries out the following protocol with the server. In a preprocessing step the client generates some secret (I_s) and some public infomation (I_p) using a fixed randomized algorithm. When the client wants to sign a message he contacts a server. For a fixed number of times he sends him some information and receives an answer. This information depends on the security parameters, the public information generated in the preprocessing and the previous responses of the server. At the end the client computes the signature y using the information responded by the server and his private information generated in the preprocessing. If both parties follow the protocol the result is correct, i.e. $y = x^d \bmod n$.

2.2 Generic Attacks

In [Nec94] Nechaev investigated the complexity of the discrete log problem. Given a finite abelian group G and a fixed generator g of G this is the problem to compute for any $a \in G$ an integer x with $g^x = a$. Nechaev considered algorithms that only perform two kinds of operations on elements of G, multiplication and equality testing.

In [Sho] Shoup proved lower bounds for the complexity of the discrete log in $\mathbf{Z}_n$ and the Diffie Hellmann problem. He considers algorithms that do not have any computational restriction but get their input (g, g^x) encoded by a randomly chosen function σ. Furthermore they have access to an oracle, which on input $\sigma(g^a), \sigma(g^b)$ outputs $\sigma(g^{a+b})$ or on input $\sigma(g^a)$ outputs $\sigma(g^{-a})$. Shoup calls this kind of algorithms generic. The class of these algorithms can be seen as those that do not depend on the representation of the group G.

These models are closely related and both authors prove a lower bound $\Omega(\sqrt{p})$ for the discrete log problem, where p is the greatest prime divisor of $ord(G)$. We adapt the model of generic algorithms to prove lower bounds for the server aided RSA protocols.

Let n, x, y, d be as above and r, s be distinct primes dividing $\phi(n)$. Further let S be a set of bit strings of cardinality at least n and σ be an injective map from $\mathbf{Z}_n$ to S, the encoding function. A *generic attack* against the server aided RSA protocol is an algorithm which for randomly chosen x, d, σ takes as input $\sigma(x)$ and $\sigma(y)$. It is required to output a value d' with $x^{d'} = x^d \bmod n$. During the computation the algorithm consults an oracle, which on input (a, b) returns $\sigma(y^a x^b)$. This oracle is more powerful than that in the Shoup model. We allow the value b to depend on the public information I_p of the client, but claim that $\gcd(a_1 - a_2, rs) = 1$ holds for any pair of (distinct) oracle queries $(a_1, b_1), (a_2, b_2)$. This is motivated by the real scenario, where the attacker knows I_p but not r and s. For fixed x and I_p let δ be the number of oracle queries of the algorithm and ρ its probability of success over randomly chosen I_s and a randomly chosen encoding function σ. Then the complexity of the algorithm is defined by the expectation of δ/ρ, over randomly chosen x and I_p.

In practice an adversary can't take advantage of equations $x^\alpha = y^\beta \bmod n$, unless β divides α. The most obvious way to avoid this problem, is to look for equations with $\beta = 1$, like in the meet-in-the-middle attacks of Pfitzmann and Waidner ([PW93]) or Lim and Lee ([LL95]). We formalize this type of attack.

We call a generic attack a *meet-in-the-middle attack* iff in all oracle queries the value a is binary, i.e. all queries have the form $(0, b)$ or $(1, b)$. These kind of attacks actually cover the so called meet-in-the-middle attacks of Pfitzmann and Waidner ([PW93]) and Lim and Lee ([LL95]), which have been the most successful attacks against RSA-S1 resp. RSA-S1M so far.

All known passive attacks against RSA-S1 and RSA-S1M are generic attacks. Generic algorithms may depend on n but not on the public key e. Otherwise the algorithm could give d rightaway without computation. It seems impossible for a real attacker to use I_p and e both, because in the considered protocols they are related by a quadratic equation modulo $\phi(n)$. Since we focus on attacks that exploit the additional information given by the SASC protocols, we consequently disregard the public key.

Most of our results generalize to the case, where the value a of the queries depends on the public infomation I_p as well. Details are given in sections 3 and 4.

3 The 1 Round Protocol

Consider the RSA-S1 protocoll, described in section 1. There a weak device (called the client) generates a signature $y = x^d \bmod n$ of a message $x \in \mathbf{Z}_n^*$, using the help of a powerful device (called the server). For the sake of clarity we claim that the Hamming weight of the secret vectors is fixed, but our results hold for more general variants as well (see the remark at the end of this section). We suppose that $\phi(n)$ has a large prime factor r so that r^2 doesn't divide $\phi(n)$. This condition always holds for a secure RSA modulus.

In the RSA-S1 protocol the secret information I_s is the vector $\mathbf{f}$ and the public information I_p is the message x and the vector $\mathbf{d}$.

In the preprocessing step $(f_1, \ldots, f_m)$ is chosen uniformly from the set of all 0-1-vectors with Hamming weight k. We denote this set by X_k. Given the f_i and d, the d_i are generated by a random process as follows. Let j be the largest index i with $f_i = 1$. Then, all d_i with $i \neq j$ are drawn independently according to the uniform distribution on $\{0, \ldots, c \cdot \phi(n)\}$ for a constant integer $c > 1$ and d_j is computed as $d_j = d - \sum_{i=1}^{j-1} d_i f_i \bmod \phi(n)$. The integer c should be chosen not too small to prevent the knapsack attacks discussed in section 5.

For the generation of the f_i and d_i different scenarios are possible. They can either be generated once for each card and stored in the ROM of the card. Or the generation is done by the card while communicating with the server. In the latter case the cards needs to store all f_i but only a constant number of $\log n$ bit numbers. The pros and cons of both methods are discussed in [LL95].

3.1 The best known Attack on RSA-S1

A trivial attack is to enumerate all $\binom{m}{k}$ canditates $\mathbf{f} \in X_k$, compute $c = \sum_{i=1}^{m} f_i\, d_i$ and check if $x^c = y \bmod n$.

A more sophisticated approach, called the meet-in-the-middle attack, was proposed by Pfitzmann and Waidner [PW93]. We present a variation of their attack which was proposed by Oorschot and Wiener in [vOW96] and is slightly more efficient.

From the definition of the protocol we have $x^d = \prod_{f_i=1} z_i \bmod n$. Let m, k be even. With probability $\rho := \binom{m/2}{k/2}^2 / \binom{m}{k}$ it holds that $(f_1, \ldots, f_{\frac{m}{2}})$ has Hamming weight $k/2$. For all possible $(f_1, \ldots, f_{\frac{m}{2}})$ with Hamming weight $k/2$ the attack computes

$$\prod_{\substack{f_i=1\\ i\le m/2}} z_i \bmod n$$

and sorts them. Subsequently for all $(f_{\frac{m}{2}+1}, \ldots, f_m)$ with Hamming weight $k/2$ it computes

$$y \Big(\prod_{\substack{f_i=1\\ i> m/2}} z_i \Big)^{-1} \bmod n$$

and sorts them as well. It is easy to see that if $(f_1, \ldots, f_{\frac{m}{2}})$ has Hamming weight $k/2$ then there is a collision which reveals d.

Since the complexity of generic algorithms does not count the effort for sorting and checking the equalities the complexity of this attack is $2\binom{m/2}{k/2}/\rho = 2\binom{m}{k}/\binom{m/2}{k/2}$.

3.2 The Security of RSA-S1

The following theorem shows that the complexity of this attack is nearly optimal. Let $\gamma(n) := \phi(\phi(n))/\phi(n)$ and $N := \binom{m}{k}$.

Theorem. 3.1 *Let be n, r, k, m as above so that $(N^2+1)/r < \gamma(n)/\sqrt{2}$. Then any generic attack against RSA-S1 has at least complexity $\gamma(n)N^{1/2}$*

Proof. For a random variable X let $E(X)$ be the expectation of X. Consider the set $Z := \{\mathbf{f} \in X_k \mid \gcd(\sum_{i=1}^{m} f_i\, d_i, \phi(n)) = 1\}$. It is easy to see that $E(|Z|) = \gamma(n)N$ where the expectation is taken over a randomly chosen $\mathbf{d}$.

For randomly chosen $\mathbf{d}$ the probability that there are two vectors $\mathbf{f}$ and $\mathbf{f'}$ with $\sum_{i=1}^{m} f_i d_i = \sum_{i=1}^{m} f'_i d_i \bmod r$ (collision) is at most N^2/r. Furthermore for randomly chosen x the probability that r doesn't divides the order of x is $1/r$. Depending on x and $\mathbf{d}$ let Ψ be $|Z|$ iff there is no collision and r divides the order of x and 0 else. Since $|Z| < N$ we can estimate $E(\Psi) \geq E(|Z|) - N(N^2+1)/r > \gamma(n)N/\sqrt{2}$.

Let n, x and $\mathbf{d}$ be fixed and $\mathcal{A}$ be a generic attack that makes δ oracle queries and has probability (over a randomly chosen $\mathbf{f} \in \mathbf{Z}$) of success ρ. We show that

$$\delta/\rho > \sqrt{2/N}\Psi \tag{1}$$

For $\Psi = 0$ this is trivial.

Now let $\Psi = |Z| > 0$. Then there is no collision and r divides the order of x. The probability that $\mathcal{A}$ outputs a d' with $d = d' \bmod r$ is at least ρ. For each pair of oracle queries (a_i, b_i), (a_j, b_j) with $a_i \neq a_j \bmod r$ the oracle returns the same value only if $b_i - b_j = (a_j - a_i)d \bmod r$. Since there is no collision this holds with probability (over a randomly chosen $\mathbf{f} \in Z$) at most $1/|Z|$. Therefore the probability that the oracle returns the same value for any pair (a_i, b_i), (a_j, b_j) with $a_i \neq a_j \bmod r$ is at most $\binom{\delta}{2}/|Z|$. On the other hand, since the encoding function is random, the probability that for all pairs (a_i, b_i), (a_j, b_j) with $a_i \neq a_j \bmod r$ the oracle answers are distinct and $\mathcal{A}$ outputs a d' with $d = d' \bmod r$ is at most $1/|Z|$. Thus we get $\delta^2 > 2\rho|Z|$ which implies (1).

Taking expectations on both sides yields the claim.

■

Remark. Theorem 3.1 holds for the non-binary RSA-S1 as well, where the f_i are ℓ-bit integers. In this case the client performs $k + \ell - 1$ multiplications, a generalisation of the described meet-in-the-middle attack has complexity $2\binom{m\ell}{k}\binom{m\ell/2}{k/2}^{-1}$ and we get a lower bound of $\gamma(n)\binom{m\ell}{k}^{1/2}$. In addition Theorem 3.1 holds if $\mathbf{f}$ are chosen from the set of vectors with Hamming weight *at most* k. There the number of possible choices of $\mathbf{f}$ is $N = \sum_{i=1}^{k}\binom{m\ell}{i}$. Furthermore Theorem 3.1 remains valid if the values a_i of the oracle queries may depend on the public information $\mathbf{d}$.

4 The 2-Round Protocol

To prevent the meet-in-the-middle attack Matsumoto, Imai, Laih and Yen [MILY93] proposed a 2-round server-aided RSA computation protocol called RSA-S1M. We consider a variant where the Hamming weight of the secret vectors is fixed. This restriction is essentiell for our results.

4.1 The RSA-S1M Protocol

Let n be a product of two large primes and r, r' be large primes so that rr' divides $\phi(n)$ but r^2, r'^2 do not divide $\phi(n)$. The client wants to sign a message x with his secret key d.

0. (Preprocessing) The client chooses an integer vector $\mathbf{d} \in \mathbf{Z}^m$ and two vectors $\mathbf{f}, \mathbf{g} \in \{0,1\}^m$ with Hamming weight k so that $d = f \cdot g \bmod \phi(n)$ where f, g are defined as $f = \sum_{i=1}^{m} f_i d_i$ and $g = \sum_{i=1}^{m} g_i \bar{d}_j$ with $\bar{d}_j = d_j(j + 3m)$. Furthermore the client randomly picks an $s \in \mathbf{Z}_n$ and computes $t = s^{-g} \bmod n$.

1. The client sends x, n and $\mathbf{d}$ to the server.

2. The server returns $z_i = x^{d_i} \bmod n$ for $i = 1, \ldots, m$.

3. The client computes and sends to the server $z = s \cdot \prod_{f_i=1} z_i = s \cdot x^f \bmod n$.

4. The server returns $v_j = z^{\bar{d}_j} \bmod n$ for $j = 1, \ldots, m$.

5. The client computes the signature y as $y = t \cdot \prod_{g_j=1} v_j = t \cdot z^g \bmod n$.

In this protocol the secret information I_s are the vectors $\mathbf{f}, \mathbf{g}$ and the public information I_p is the message x and the vector $\mathbf{d}$.

Again let X_k denote the set of 0-1 vectors with Hamming weight k and set $d(\mathbf{f}, \mathbf{g}) := \sum_{i,j=1}^{m} f_i\, g_j\, d_i\, \bar{d}_j \bmod \phi(n)$. Since z is a random number it does not reveal any information about d to the server.

The vectors $\mathbf{f}, \mathbf{g}$ are uniformly drawn from X_k. Here we only consider the case where $\mathbf{f} \neq \mathbf{g}$. So there are i' and i'' with $f_{i'} = 1$, $g_{i'} = 0$, $f_{i''} = 0$, and $g_{i''} = 1$.

d is uniformly drawn from $\mathbf{Z}^*_{\phi(n)}$. All d_i except $d_{i'}$ are drawn independently and uniformly from $\mathbf{Z}_n$. $d_{i'}$ is chosen so that $\sum f_i\, d_i$ is invertible modulo $\phi(n)$. $d_{i''}$ is chosen so that $(\sum f_i\, d_i)\left(\sum g_j\, \bar{d}_j\right) = d \bmod n$. Finally x is drawn uniformly from $\mathbf{Z}_n$.

Our protocol differs from the original RSA-S1M ([MILY93]). We insist on a fixed Hamming weight k of $\mathbf{f}$, whereas in [MILY93] $\mathbf{f}$ is chosen with a Hamming weight up to k and in the second round we let the server use the $\bar{d}_j = d_j(j+3m)$ as exponents instead the d_j. This modifications have technical reasons and do not substantly affect the efficiency of the protocol. In order to achieve a security of 2^{64} we have to insist on an RSA modulus of at least 750 Bit.

4.2 The best known attack on RSA-S1M

In [LL95] Lim and Lee showed that the ideas of [PW93] are applicable to RSA-S1M as well. They gave a meet-in-the-middle attack with complexity $O\left(N^{3/4}\right)$, where N was the number of possible pairs $(\mathbf{f}, \mathbf{g})$. We give a variation of this attack which is slighly more efficient and uses ideas of [vOW96].

Let m, k be even. From the definition of the protocol we have

$$x^d = \prod_{f_i=1} x^{gd_i} \bmod n.$$

With probability $\rho := \binom{m/2}{k/2}^2 / \binom{m}{k}$ it holds that $(f_1, \ldots, f_{\frac{m}{2}})$ has Hamming weight $k/2$. The attack guesses $\mathbf{f}$, thereby determines f and for all possible tupels $(g_1, \ldots, g_{\frac{m}{2}})$ with Hamming weight $k/2$ it computes the values

$$\prod_{\substack{g_j=1 \\ j \le m/2}} v_j \bmod n$$

and sorts them. Subsequently for all $(g_{\frac{m}{2}+1}, \ldots, g_m)$ with Hamming weight $k/2$ it computes the values

$$y \Big(\prod_{\substack{g_j=1 \\ j > m/2}} v_j \Big)^{-1} \bmod n$$

and sorts them as well. It is easy to see that if $(g_1, \ldots, g_{\frac{m}{2}})$ has Hamming weight $k/2$ then there is a collusion which reveals d.

This attack is covered by the decision tree model and has complexity $2\binom{m}{k}\binom{m/2}{k/2}/\rho = 2\binom{m}{k}^2/\binom{m/2}{k/2}$.

4.3 The Complexity of Generic Attacks

Let $N := \binom{m}{k}^2$ be the number of possible choices of $(\mathbf{f}, \mathbf{g})$. Of Course, the bound $\Omega(\sqrt{N})$ for the complexity of any generic attack holds for the RSA-S1M protocol as well. The proof is analogous to the proof of Theorem 3.1. But RSA-S1M was designed to achieve a better efficiency than that of RSA-S1. In fact we are able to prove a lower bound of $\Omega(N^{2/3})$.

Let $n, r, s, x, \mathbf{d}$ be fixed and $Z := \{(\mathbf{f}, \mathbf{g}) \in X_k^2 \mid \gcd(d(\mathbf{f}, \mathbf{g}), \phi(n)) = 1\}$. Consider a generic attack $\mathcal{A}$ that makes the oracle queries $(a_1, b_1), \ldots, (a_\delta, b_\delta)$ and has probability of success ρ, over randomly chosen $(\mathbf{f}, \mathbf{g}) \in Z$ and random encoding function σ.

Definition We say that two oracle queries (a_i, b_i), (a_j, b_j) are *related via* $(\mathbf{f}, \mathbf{g}) \in Z$ iff $a_i \neq a_j$ and $y^{a_i} x^{b_i} = y^{a_j} x^{b_j} \bmod n$ holds with $y = x^{d(\mathbf{f},\mathbf{g})} \bmod n$. The latter condition means that the oracle answers of the queries are identical if $d = d(\mathbf{f}, \mathbf{g})$ and implies $b_i - b_j = (a_j - a_i) d(\mathbf{f}, \mathbf{g}) \bmod \operatorname{ord}(x)$. We say that (a_i, b_i), (a_j, b_j) are *related* if they are related via a pair $(\mathbf{f}, \mathbf{g}) \in Z$.

We define a graph $G = (V, E)$ as follows: For every oracle query (a_i, b_i) set a vertex $u_i \in V$. For $i \neq j$ set an edge $(u_i, u_j) \in E$ iff (a_i, b_i) and (a_j, b_j) are related. The following Lemma reveals the connection between the size of E and the probability of success of the attack.

Lemma. 4.1 *With probability at least* $1 - (4N^2 + r + s)/rs$ *(over randomly chosen* $\mathbf{f}, \mathbf{g}$ *and* σ*) it holds that* $\rho |Z| \leq |E| + 1$.

Proof. Assume that rs divides the order of x. This holds with probability at least $1 - 1/r - 1/s$. Then with probability at least ρ the algorithm outputs a d' with $d = d' \bmod rs$. Further assume that there is no *collision* $d(\mathbf{f}, \mathbf{g}) = d(\mathbf{f}', \mathbf{g}') \bmod rs$ with $(\mathbf{f}, \mathbf{g}) \neq (\mathbf{f}', \mathbf{g}') \in Z$. This holds with probability at least $1 - 4N^2/rs$. Then two oracle queries (a_i, b_i), (a_j, b_j) are related via at most one pair $(\mathbf{f}, \mathbf{g}) \in X_k^2$. Therefore the probability (over randomly chosen $(\mathbf{f}, \mathbf{g}) \in Z$) that there are any related oracle queries is at most $|E|/|Z|$. On the other hand, since the encoding function σ is random, the probability (over randomly chosen $\mathbf{f}, \mathbf{g} \in Z$) that there are no related oracle queries and $\mathcal{A}$ outputs a d' with $d = d' \bmod rs$ is at most $1/|Z|$.

■

Exploiting the nonexistence of certain cycles in G, we get the following result:

Theorem. 4.2 *Let* $N > 2^{92}$ *and* $k > 10$. *Then with probability at least* $1 - 4N^{20/3}/rs$ *(over a randomly chosen vector* $\mathbf{d}$*) it holds that* $|V| > 2^{-2.7} |E| N^{-1/3}$.

The proof is given in the full paper.

Using $|E| \leq N$ and $\delta = |V|$ we are now able to proof a lower bound for the complexity of a generic attack. Let $\tau := (4N^{20/3} + 4N^2 + r + s)/rs$.

Theorem. 4.3 *Let be* n, r, s, m, k *as above so that* $N \geq 2^8$, $k > 10$ *and* $\tau < \gamma(n)^2/20$. *Then every generic attacker breaking RSA-S1M has at least complexity* $2^{-3} N^{2/3} \gamma(n)^2$.

Proof. Using standard arguments, we can estimate $E(|Z|) \geq \binom{m}{k}(\binom{m}{k} - 1)\gamma(n)^2$, where the expectation is taken over a randomly chosen vector $\mathbf{d}$. Since $N \geq 2^8$ this is at least $\frac{19}{20} N \gamma(n)^2$. Depending on x and $\mathbf{d}$ let Ψ be $|Z|$ if Lemma 4.1 and Theorem 4.2 hold, and 0 else. Using $|Z| < N$ we can estimate $E(\Psi) \geq E(|Z|) - \tau N$. Since $\tau < \frac{1}{20}\gamma(n)^2$ and $E(|Z|) > \frac{19}{20} N \gamma(n)^2$ we find that $E(\Psi) > 2^{0.2} N \gamma(n)^2$.

On the other hand by Lemma 4.1 and Theorem 4.2 we have $\delta > 2^{-2.7}(\rho\Psi - 1) N^{-1/3}$. Thus for $\rho\Psi \geq 30$ we get

$$\delta/\rho \geq 2^{-2.75} \Psi N^{-1/3} \tag{2}$$

and since $N \geq 2^8$ for $\rho\Psi > 30$ equation (2) is trivial. Taking expectations on both sides yields the claim.

■

4.4 The Complexity of Meet-in-the-midle Attacks

We show that best known attack against RSA-S1M is optimal for a meet-in-the-middle-attack.

Let $n, r, s, x, \mathbf{d}$ be fixed and let $\mathcal{A}$ be a meet-in-the-middle attack that makes the oracle queries $(a_1, b_1), \ldots, (a_\delta, b_\delta)$ (i.e. $a_i \in \{0,1\}$ for $i = 1, \ldots, \delta$) and has probability of success ρ. Let the graph G be defined as above. Then Lemma 4.1 still holds.

Due to the particular form of the oracle queries, G is bipartide. We get an even better bound for its number of edges. The proof is given in the full paper.

Theorem. 4.4 *Let $N > 2^{92}$. Then with probability at least $1 - 8N^6/rs$ (over a randomly chosen vector $\mathbf{d}$) it holds that $|V| > 2^{-4.75}|E|N^{-1/4}$.*

We are now able to prove the lower bound for the complexity of birthday attacks. Let $\tau := (8N^6 + 2N^2 + r + s)/rs$.

Theorem. 4.5 *Let be n, r, s, m, k as above so that $N > 2^{92}$ and $\tau < \frac{1}{20}\gamma(n)^2$. Then every meet-in-the-middle attack against RSA-S1M has at least complexity $\gamma(n)^2 2^{-5} N^{3/4}$.*

Proof. Depending on x and $\mathbf{d}$ let Ψ be $|Z|$ if Lemma 4.1 and Theorem 4.4 hold, and 0 else. Analogously to the proof of Theorem 4.3 we find $E(\Psi) > 2^{-0.2}N\gamma(n)^2$.

On the other hand by Lemma 4.1 and Theorem 4.4 we get $\delta > 2^{-4.75}(\rho\Psi - 1)N^{-1/4}$. For $\rho\Psi \geq 30$ we get

$$\delta/\rho \geq 2^{-4.8}\Psi N^{-1/4} \tag{3}$$

and since $N \geq 2^{92}$ for $\rho\Psi > 30$ equation (3) is trivial. Taking expectations on both sides yields the claim.

■

Remark. Even if the values a_i of the oracle queries may depend on the public information $\mathbf{d}$, Theorem 4.5 still remains valid.

5 Conclusions

5.1 Discussion of our Results

We give some concrete examples of the sharpness of our results for several choices of the parameters that yield a security of 2^{64}. Since the binary RSA-S1 is not very efficient, we consider the non-binary version. There the f_i are l-bit integers and Theorem 3.1 holds as well. In the case of RSA-S1M , for technichal reasons, we suppose that $\phi(n)$ has prime factors r, s fulfilling $rs \gg \binom{m}{k}^{12}$. This holds for a secure RSA modulus of at least 750 bit. But we don't believe that RSA-S1M is less secure for 512 bit moduli.

We compare the upper bound c_1 (given by the described attacks) for the security and the lower bounds c_2 and c_3 (given by our results) for the complexity of a generic resp. a birthday attack against RSA-S1 and RSA-S1M. In the case of RSA-S1 we have $c_2 = c_3$. We omit the terms $\gamma(n)$ and $\gamma(n)^2$. They don't seem to play any role in practice because an attacker does not have much knowledge about $\phi(n)$. The client has to perform $k + l - 2$ resp. $2k + 1$ multiplications.

RSA-S1					RSA-S1M				
m	k	l	c_1	c_2	m	k	c_1	c_2	c_3
80	40	2	$2^{65.4}$	$2^{63.0}$	54	20	$2^{70.9}$	$2^{61.3}$	$2^{66.3}$
92	36	2	$2^{66.1}$	$2^{63.7}$	60	16	$2^{69.2}$	$2^{59.8}$	$2^{64.6}$

These examples show that the bounds of Theorem 3.1 and Theorem 4.5 are quite sharp. The factors $\gamma(n)$ resp. $\gamma(n)^2$ don't seem to play any role in practice because an attacker does not have much knowledge about $\phi(n)$.

5.2 Active Attacks

Various active attacks against the protocols RSA-S1 and RSA-S1M have been proposed in the past (for example see [And92],[Kaw95],[LL95]). In an active attack the server returns false values and tries to extract information out of the results presented by the client. As noted in [LL95] the active attacks can be partially prevented by checking $y = x^d$. This can be done efficiently if e is small by computing y^e. A forged y_i with $f_i = 1$ will be detected. But under certain circumstances it may be dangerous to choose e small. Lim and Lee ([LL95]) proposed a method to check the equality $y = x^d$ using only 6 multiplication, irrespective of the size of e.

However there are still multi-round active attacks possible. Lim and Lee suggested to change the secret vectors $\mathbf{f}, \mathbf{g}$ and $\mathbf{d}$ randomly after a small number of runs. But in this case it is crucial not to choose the vector $\mathbf{d}$ from $\mathbf{Z}^m_{\phi(n)}$ but as integer vectors with components from an intervall $[1, \ldots, c\phi(n)]_\mathbf{Z}$ with a constant integer c being large enough. Otherwise the following knapsack attack is possible. Let $d = \sum_{i=1}^m f_i^{[l]} d_i^{[l]} \bmod \phi(n)$ for $l = 1, 2, \ldots$. Since the vectors $\mathbf{d}^{[l]}$ are chosen randomly in $\mathbf{Z}^m_{\phi(n)}$, it is not unlikely that $\sum_{i=1}^m f_i^{[l]} d_i^{[l]} = \sum_{i=1}^m f_i^{[l']} d_i^{[l']}$ holds in $\mathbf{Z}$ for small l, l'. In this case using a $2m$ dimensional lattice it is possible to obtain the secret vectors by lattice reduction.

If the vectors are changed after a small number of runs and $\mathbf{d}$ is chosen from a large intervall (e.g. $c = 8$) then all known active attacks are prevented.

A Proof of Theorem 4.2

We give a short sketch, how Theorem 4.2 is proven. First we show that for each 4-cycle in G with high probability we get an equation in vectors $\mathbf{f}^{[1]}, \mathbf{g}^{[1]}, \ldots, \mathbf{f}^{[4]}, \mathbf{g}^{[4]} \in X_k$. This is done by the Lemmata A.1, A.5 and A.6. Using this result, we prove in Lemma A.7 that with high probability G does not contain certain 4-cycles. Finally we exploit this property to prove Theorem 4.2.

Lemma. A.1 *Let $k > 10$, $\alpha_1, \ldots, \alpha_4 \in \mathbf{Z}_r \setminus \{0\}$ with $\sum_{l=1}^4 \alpha_l = 0 \bmod r$ and let $\mathbf{f}^{[1]}, \mathbf{g}^{[1]}, \ldots, \mathbf{f}^{[4]}, \mathbf{g}^{[4]} \in X_k$ and $A := \sum_{l=1}^4 \alpha_l\, \mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} \bmod r$. Then for at least one pair i, j the equation*

$$jA_{ij} + iA_{ji} = 0 \tag{4}$$

does not hold.

Proof. We asume that equation (4) holds for all i, j. Let $f(i)$ be the number of l with $f_i^l = 1$ and $g(i)$ the number of l with $g_i^l = 1$. The following facts are easy to check:

Fact. A.2 *$\sum_l f_i^l g_i^l \alpha_l = 0$ holds for all i.*

Fact. A.3 *For $j_1 \neq j_2$ with $A_{ij_1} \neq 0$ it holds that $A_{ij_1} \neq A_{ij_2}$ or $A_{j_1 i} \neq A_{j_2 i}$*

Fact. A.4 *If $f(i) = 3$ then $A_{ji} \neq 0$ holds for at least $k/2$ many j.*

Now we consider any i with $A_i \neq \mathbf{0}$.

The case $f(i) = 1$. W.l.o.g. $f_i^1 = 1$. Then by fact A.2 we get $g_i^1 = 0$ and thus for all j we have $A_{ji} = \sum_{l=2}^{4} \epsilon_l^j \alpha_l$ with $\epsilon_l^j \in \{0, 1\}\}$. Since $k > 8$ we find $j_1 \neq j_2$ with $A_{j_1 i} = A_{j_2 i} \neq 0$ and $A_{ij_1} = A_{ij_2} = \alpha_1$ which contradicts fact A.3.

The case $g(i) = 1$ is analogous.

The case $f(i) = 2$ and $g(i) = 2$. W.l.o.g. $f_i^1 = f_i^2 = 1$ and ,by fact A.2, $g_i^1 = g_i^2$. If $g_i^1 = g_i^2 = 1$ we get $\alpha_1 + \alpha_2 = 0$ by fact A.2. Fix any j with $A_{ij} \neq 0$. Since $A_{ij}, A_{ji} \in \{\alpha_1, \alpha_2\}$ we find $\alpha_1 = \alpha_2 = 0$, a contradiction. On the other hand if $g_i^3 = g_i^4 = 1$ there are two cases: If $\alpha_1 + \alpha_2 \neq 0$ it holds that at least $k+1 > 9$ many A_{ij} are nonzero. Since $A_{ij} \in \{\alpha_1, \alpha_2, \alpha_1 + \alpha_2\}$ and $A_{ji} \in \{\alpha_3, \alpha_4, \alpha_3 + \alpha_4\}$ we find $j_1 \neq j_2$ with $A_{ij_1} = A_{ij_2} \neq 0$ and $A_{j_1 i} = A_{j_2 i} \neq 0$ which contradicts equation 4. On the other hand, if $\alpha_2 = -\alpha_1$, we get $\alpha_3 = -\alpha_4$. Since $A_i \neq \mathbf{0}$ we find $j_1 \neq j_2$ with $A_{ij_1} = \alpha_1$ and $A_{ij_2} = \alpha_2$. We get the equations

$$\begin{aligned} iA_{j_1 i} + j_1 \alpha_1 &= 0 \\ iA_{j_2 i} - j_2 \alpha_1 &= 0 \end{aligned}$$

with $A_{j_1 i}, A_{j_2 i} = \in \{\alpha_3, -\alpha_3\}$. This yields $\alpha_1 = 0$.

The case $f(i) = 3$ and $g(i) = 2$. W.l.o.g. $f_i^1 = f_i^2 = f_i^3 = 1$ and $g_i^1 \geq g_i^2 \geq g_i^3$. Then by fact A.2 we get $g_i^1 = g_i^2 = 1$, $g_i^3 = g_i^4 = 0$, $\alpha_2 = -\alpha_1$ and $\alpha_4 = -\alpha_3$. and thus $A_{ji} \in \{0, \alpha_1, \alpha_2\}$. Let $A_{j_1 i} = \alpha_1$ and $A_{j_2 i} = \alpha_2$.

If $A_{ij_1} = \alpha_1 + \alpha_3$ then we get $(j_1 + i)\alpha_1 + i\alpha_3 = 0$. On the other hand $g_{j_1}^2 = 0$ and thus by fact A.2 we get $0 = \alpha_1 + \alpha_4 = \alpha_1 - \alpha_3$. This yields $\alpha_1, \ldots, \alpha_4 = 0$.

If $A_{ij_1} = \alpha_2$ then we get a contradiction by the equation $j_1 \alpha_1 + i\alpha_2 = 0$.

Therefore we can conclude that $A_{ij_1} \in \{\alpha_2 + \alpha_3, \alpha_3\}$ and analogously we see that $A_{ij_2} \in \{\alpha_1 + \alpha_3, \alpha_3\}$. By fact A.4 at least k many A_{ij} are nonzero. Since $k > 8$ we find $j \neq j'$ with $A_{ij} = A_{ij'} \neq 0$ and $A_{ji} = A_{j'i}$ which contradicts fact A.3.

The case $f(i) = 2$ and $g(i) = 3$ is analogous.

The case $f(i) = 3$ and $g(i) = 3$. W.l.o.g. $f_i^1 = f_i^2 = f_i^3 = 1$ and $g_i^1 = g_i^2 = 1$. Then by fact A.2 we get $g_i^3 = 0$, $\alpha_1 = -\alpha_2$ and $\alpha_3 = -\alpha_4$.

If $\alpha_1 + \alpha_3 = 0$ then by fact A.4 at least $k/2$ many A_{ij} are nonzero. Since $k > 6$ we find $j \notin \{i/2, 2i\}$ with $A_{ij} \neq 0$. We get an equation $jA_{ij} + iA_{ji} = 0$ with $A_{ij} \in \{\alpha_1, -\alpha_1, -2\alpha_1\}$ and $A_{ji} \in \{\alpha_1, -\alpha_1, 2\alpha_1\}$ and thus $\alpha_1 = 0$.

If $\alpha_2 + \alpha_3 = 0$ we get a contradiction analogously.

If $\alpha_1 + \alpha_3 \neq 0$ and $\alpha_2 + \alpha_3 \neq 0$ then at least $k > 10$ many A_{ij} are nonzero. We distinguish 3 cases:

- If $|\{j \mid A_{ij} = \alpha_3\}| > 4$ then, since by fact A.2 $A_{ij} = \alpha_3$ implies $A_{ji} \in \{\alpha_1, -\alpha_1, -\alpha_3, \alpha_1 - \alpha_3, -\alpha_1 - \alpha_3\}$, we find $j_1 \neq j_2$ with $A_{ij_1} = A_{ij_2} = \alpha_3$ so that $A_{j_1 i} = A_{j_2 i}$ or $A_{j_1 i} = -A_{j_2 i}$. The equations $j_1 \alpha_3 + iA_{j_1 i} = 0$ and $j_2 \alpha_3 + iA_{j_2 i} = 0$ yield $\alpha_1 = 0$.
- If $|\{j \mid A_{ij} = \alpha_1 + \alpha_3\}| > 3$ then, since by fact A.2 $A_{ij} = \alpha_1 + \alpha_3$ implies $A_{ji} \in \{-\alpha_1, -\alpha_3, -\alpha_1 - \alpha_3\}$, we find $j_1 \neq j_2$ with $A_{ij_1} = A_{ij_2} = \alpha_1 + \alpha_3$ and $A_{j_1 i} = A_{j_2 i}$, which contradicts fact A.3.
- If $|\{j \mid A_{ij} = -\alpha_1 + \alpha_3\}| > 3$ we get a contradiction analogously.

■

Lemma. A.5 *Let* $\mathbf{f}^{[1]}, \mathbf{g}^{[1]}, \ldots, \mathbf{f}^{[4]}, \mathbf{g}^{[4]} \in X_k$ *and* $\alpha_1, \ldots, \alpha_4 \in \mathbf{Z}^*_{rs}$ *with* $\sum_l \alpha_l = 0$ *so that*

$$\sum_{l=1}^{4} \alpha_l \mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} \neq 0$$

holds. Then with probability at most $4/rs$ *(over randomly chosen* $\mathbf{d}$*) it holds that*

$$\sum_{l=1}^{4} \alpha_l \sum_{i,j=1}^{m} f_i^{[l]} \, g_j^{[l]} \, d_i \, \bar{d}_j = 0 \bmod rs. \tag{5}$$

Proof. Equality in (5) is equivalent to equality mod$r$ and mod$s$. We show that equality mod$r$ holds only with probability $2/r$. Analogously one can see that the probability of equality mod$s$ is $2/s$.

Let equation (5) hold mod$r$.Let

$$\bar{c}_{i,j} = \sum_{l=1}^{4} \alpha_l f_i^{[l]} \, g_j^{[l]} \bmod r.$$

and, for $1 \leq i < j \leq m$,

$$\begin{aligned} c_{i,j} &= \bar{c}_{i,j}\,(j+3m) + \bar{c}_{j,i}(i+3m) \bmod r \\ c_{i,i} &= \bar{c}_{i,i}(i+3m) \bmod r. \end{aligned}$$

If $\sum_{l=1}^{4} \alpha_l \, \mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} \neq 0$ then there is a $\bar{c}_{i,j} \neq 0$. By Lemma A.1 there is a pair i, j with $c_{i',j'} \neq 0$ as well.

We show that equation (5) holds mod$r$ only with probability at most $2/r$.

We distinguish two cases. First assume that $i' = j'$. In this case (5) is equivalent to

$$c_{i',i'}(i+3m) \, d_{i'}^2 + \gamma_1 \, d_{i'} + \gamma_2 = 0 \bmod r\,, \tag{6}$$

where γ_1 and γ_2 depend on the $c_{i,j}$ and the d_i $(i \neq i')$. Since $\mathbf{Z}_r$ is a field, equality (6) holds for no more than two out of the r possible values of d_i.

In the second case we have $i' \neq j'$ and all $c_{i,i} = 0$. Fixing all d_i except $d_{i'}$ and $d_{j'}$ (5) becomes

$$c_{i',j'} \, d_{i'} \, d_{j'} + \gamma_1 \, d_{i'} + \gamma_2 \, d_{j'} + \gamma_3 = 0 \bmod r \tag{7}$$

with certain constants $\gamma_1, \gamma_2, \gamma_3$. Interpreting the left hand side as a linear function in $d_{i'}$ the coefficient of $d_{i'}$ is $c_{i',j'} \, d_{j'} + \gamma_1$. This coefficient is nonzero with probability $1 - 1/r$. But in this case the linear function computes zero on a random $d_{i'}$ also with probability $1 - 1/r$. Thus the probability of (7) holding mod$r$ is at most $2/r$.

■

Now let $\mathbf{d}$ be so that there is no collision $d(\mathbf{f}, \mathbf{g}) = d(\mathbf{f}', \mathbf{g}')$ with $(\mathbf{f}, \mathbf{g}) \neq (\mathbf{f}', \mathbf{g}')$. Consider an edge (v_i, v_j) in the resulting graph G. It holds that $a_i \neq a_j$ and there exists a pair $(\mathbf{f}, \mathbf{g}) \in X_k^2$ fulfilling $b_i - b_j = (a_j - a_i) d(\mathbf{f}, \mathbf{g}) \bmod rs$. Since there is no collision this pair is unique. We label the edge by this pair. If a pair $\mathbf{f}, \mathbf{g}$ occurs more than once as a label of an edge in E we remove all but one of this edges.

The main property of G we will exploit is the non-existence of certain 4-cycles. We will then prove a variant of the well known general result that graphs with v nodes not containing cycles of length $L \leq 2\kappa$, have $O(v^{1+1/\kappa})$ edges ([Bol78]). with $\kappa = 2$. The results of non-existence of certain cycles are obtained using the following Lemmata.

Lemma. A.6 *Let* $(v_1, \ldots, v_L, v_{L+1} = v_0)$ *be a cycle of length* L *in* G. *For* $l = 1, \ldots, L$ *let* $(\mathbf{f}^{[l]}, \mathbf{g}^{[l]})$ *be the label of* (v_l, v_{l+1}). *Then*

$$\sum_{l=1}^{L} \alpha_l d(\mathbf{f}, \mathbf{g}) = 0 \bmod rs$$

holds with $\alpha_1, \ldots, \alpha_4 \in \mathbf{Z}_{rs}^*$ *and* $\sum \alpha_l = 0$.

Proof. Let $(v_1, \ldots, v_L, v_{L+1} = v_1)$ be an edge disjunct cycle of length L in G (i.e. $v_{L+1} = v_L$) so that edge (v_l, v_{l+1}) is labeled by $(\mathbf{f}^{[l]}, \mathbf{g}^{[l]})$. Then we have for $l = 1, \ldots, L$

$$(a_{l+1} - a_l)d(\mathbf{f}, \mathbf{g}) = b_l - b_{l+1} \bmod rs.$$

Since by definition $gcd(a_{l+1} - a_l, rs) = 1$, the claim follows by summation over l. ∎

The Lemmata A.6 and A.5 show that with high probability each 4-cycle in this graph G yields an equation in the tensor product of the labels of the edges in the cycle.

Definition. For an edge (x, y) labeled by $\mathbf{f}, \mathbf{g}$ we call $\mathbf{f}$ the F-colour and $\mathbf{g}$ the G-colour of (x, y) and write $\mathbf{f} = F(x, y)$, $\mathbf{g} = G(x, y)$. For any $\mathbf{f}, \mathbf{g} \in X_k$ there are at most $\binom{m}{k}$ edges of F-colour $\mathbf{f}$ and $\binom{m}{k}$ edges of G-colour $\mathbf{g}$. A path in G is *F-monochromic* iff all its edges are of the same F-colour. A path in G is *G-monochromic* iff all its edges are of the same G-colour. A path in G is *colourful* iff it is neither F-monochromic nor G-monochromic .

Definition. For $\tilde{V} \subseteq V$ we say that $\tilde{V}$ is *F-monochromic* if for every vertex $x \in \tilde{V}$ all edges incident to x are of the same F-colour. For any colour $\mathbf{f}$ and $x \in \tilde{V}V_1$ we say x *is of the F-colour* $\mathbf{f}$ if the edges incident to x are of the F-colour $\mathbf{f}$ and define $\tilde{V}^F(\mathbf{f})$ as the set of vertices in $\tilde{V}$ that are of the F-colour $\mathbf{f}$. Analogously we define $\tilde{V}$ to be *G-monochromic* and $\tilde{V}^G(\mathbf{g})$ for any colour $\mathbf{g}$.

Definition. We say a vertex x is *F-dominated* if at least $3/4$ of the edges incident to x are of the same F-colour. Analogously we define x to be *G-dominated.* We say x is *colourful* if it is neither F-dominated nor G-dominated.

Lemma. A.7 *With probability at least* $1 - 4\delta^4 N^4/rs$ G *does not contain colourful 4-cycles.*

Proof. By Lemma A.5 for each 4-tupel $\alpha_1, \ldots, \alpha_4$ fulfilling $\sum_l \alpha_l = 0$ with probability $1 - 4/rs$ (over randomly chosen $\mathbf{d}$) equation (5) only holds if

$$\sum_{l=1}^{4} \alpha_l \mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} = 0. \tag{8}$$

Since for oracle queries $(a_1, b_1), \ldots, (a_4, b_4)$, the values a_l are independent of $\mathbf{d}$, with probability at least $1 - 4\delta^4 N^4/rs$ for all 4-tupels of oracle queries $(a_1, b_1), \ldots, (a_4, b_4)$, $\alpha_l := a_l - a_{l+1 \bmod 4}$ and all $\mathbf{f}^{[1]}, \mathbf{g}^{[1]}, \ldots, \mathbf{f}^{[4]}, \mathbf{g}^{[4]} \in X_k$ equation (5) implies (8). By Lemma A.6 follows that each 4-cycle in G yields an equation (8) with $\mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} \neq \mathbf{f}^{[l']} \otimes \mathbf{g}^{[l']}$ for $l \neq l'$. We show that such an equation can only hold if all $\mathbf{f}^{[l]}$ are equal or if all $\mathbf{g}^{[l]}$ are equal.

Assume that equation (8) holds for $\mathbf{f}^{[1]}, \mathbf{g}^{[1]}, \ldots, \mathbf{f}^{[4]}, \mathbf{g}^{[4]} \in X_k$ and $\mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} \neq \mathbf{f}^{[l']} \otimes \mathbf{g}^{[l']}$ for $l \neq l'$ If not all $\mathbf{f}^{[l]}$ are equal and not all $\mathbf{g}^{[l]}$ are equal, then w.l.o.g. we can assume that $\mathbf{f}^{[1]} \neq \mathbf{f}^{[2]}$ and $\mathbf{g}^{[1]} \neq \mathbf{g}^{[2]}$. Set $A := \alpha_1 \mathbf{f}^{[1]} \otimes \mathbf{g}^{[1]} + \alpha_2 \mathbf{f}^{[2]} \otimes \mathbf{g}^{[2]}$. There always exist permutations σ_1, σ_2 so that the vectors $(f^{[1]}_{\sigma_1(1)}, \ldots, f^{[1]}_{\sigma_1(m)})$, $(f^{[2]}_{\sigma_1(1)}, \ldots, f^{[2]}_{\sigma_1(m)})$, $(g^{[1]}_{\sigma_2(1)}, \ldots, g^{[1]}_{\sigma_2(m)})$, and $(g^{[2]}_{\sigma_2(1)}, \ldots, g^{[2]}_{\sigma_2(m)})$ have the form $(0, \ldots, 0, 1, \ldots, 1, 0, \ldots, 0)$. One can easily see that one can determine $\mathbf{f}^{[1]}, \mathbf{f}^{[2]}, \mathbf{g}^{[1]}, \mathbf{g}^{[2]}$ from $(A_{\sigma_1(i), \sigma_2(j)})_{ij}$, which contradicts the precondition that all $\mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]}$ are different.

■

Definition. For a vertex x let $N(x)$ denote the set of vertices y adjacent to x. For an $y \in N(x)$ let $d_{F(x,y)}(x)$ denote the number of edges incident to x with F-colour $F(x,y)$. Analogously we define $d_{G(x,y)}(x)$.

Definition. For $G = (V, E)$ and $U_1, U_2 \subset V$ we set $E(U_1, U_2) := \{(x,y) \in E \mid x \in U_1 \wedge y \in U_2\}$.

Lemma. A.8 *Let $x \in V$ be colourful and for $y \in N(x)$ let $A_x(y)$ be the set of edges $(x,z) \neq (x,y)$ with $F(x,z) = F(x,y)$ or $G(x,z) = G(x,y)$. Then there is at most one $y \in N(x)$ with $|A_x(y)| > 7/8\, d(x)$.*

The proof follows immediately from $|A_x(y) \cap A_x(y')| \leq 3/4\, d(x)$.

Proof of Theorem 4.1. Assume that $\delta = |V| < 2^{-2.7}|E|N^{-1/3}$. Then, by Lemma A.7, with probability at least $1 - 4N^{20/3}/rs$ there is no collision and no colourful 4-cycle in G. Assume that this is the case. We divide E into the following subsets

$$\begin{aligned}
L_1 &:= \{(x,y) \in E \mid x \text{ is colourful}\} \\
L_2 &:= \{(x,y) \in E \mid x \text{ is F-dominated, } y \text{ is F-dominated}\} \\
L_3 &:= \{(x,y) \in E \mid x \text{ is G-dominated, } y \text{ is G-dominated}\} \\
L_4 &:= \{(x,y) \in E \mid x \text{ is F-dominated, } y \text{ is G-dominated}\}
\end{aligned}$$

If $|\mathbf{L_1}| > \frac{1}{4}|\mathbf{E}|$ then we set $G := (V, L_1)$, $V_c := \{x \in V \mid \text{x is colourful}\}$ and for every $x \in V_c$ we set

$$M(x) := \{(y, y') \in N(x)^2 \mid y \neq y',\, F(x,y) \neq F(x,y'),\, G(x,y) \neq G(x,y')\}.$$

Since there are no colourful 4-cycles in G, for $x \neq x' \in V_c$ the sets $M(x)$ and $M(x')$ are disjoint and we get $\sum_{x \in V_c} |M(x)| \leq |V|^2$. On the other hand we have $|M(x)| = 1/2 \sum_{y \in N(x)} \big(d(x) - A_x(y)\big)$ which, by Lemma A.8, is at least $2^{-4}\big(d(x)^2 - d(x)\big)$. Thus by the Cauchy-Schwarz inequality and $\sum_{x \in V_c} d(x) \geq |L_1|$ we get $2^{-4}|L_1|\big(|L_1| - n\big) \leq |V|^3$ and, since $E > 2^{28}$, $|V| > 2^{-2.7}|E|^{2/3} > 2^{-2.7}|E|N^{-1/3}$.

If $|\mathbf{L_2}| > \frac{1}{3}|\mathbf{E}|$ then G contains a subgraph $\tilde{G} = (\tilde{V}, \tilde{E})$ so that $\tilde{V}$ is F-monochromic and $|\tilde{E}| > \frac{1}{6}|E|$. We get $|\tilde{V}| = \sum_{\mathbf{f}} |\tilde{V}^F(\mathbf{f})| \leq \binom{m}{k}^{-1/2} \sum_{\mathbf{f}} |\tilde{V}^F(\mathbf{f})|^2 = |\tilde{E}|\binom{m}{k}^{-1/2}$ and thus $|V| > 2^{-2.7}|E|N^{-1/3}$.

The case $|\mathbf{L_3}| > \frac{1}{3}|\mathbf{E}|$ is analogous.

If $|\mathbf{L_4}| > \frac{1}{10}|\mathbf{E}|$ then G contains a bipartide subgraph $\tilde{G} = (\tilde{V_1}, \tilde{V_2}, \tilde{E})$ so that $\tilde{V_1}$ is F-monochromic , $\tilde{V_2}$ is F-monochromic and $|\tilde{E}| > \frac{1}{20}|E|$. Since $\tilde{G}$ does not contain colourful 4-cycles, it does not contain any 4-cycle at all. It is well known, that graphs with with v vertices and girth greater than 4 have at most $\frac{1}{2}v^{3/2}$ edges (see [LW92]). Thus we get $|V| > 2^{-2.7}|E|^{2/3} > 2^{-2}|E|N^{-1/3}$.

■

B Proof of Theorem 4.4

The case $|E^0| \leq 2^{28}$ is trivial. Before we prove the Theorem for $|E^0| > 2^{28}$, we need some definitions and technical lemmata. We assume that each edge has only one label. This holds with probability $1 - \binom{m}{k}^4/rs$.

Lemma. B.1 *Let* $\mathbf{f}^{[1]}, \mathbf{g}^{[1]}, \ldots, \mathbf{f}^{[8]}, \mathbf{g}^{[8]} \in X_k$ *with*

$$\sum_{l=1}^{8}(-1)^l \mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} \neq 0.$$

Let $d_1, \ldots, d_m$ *be independently randomly chosen from* $\mathbf{Z}_{\phi(n)}$. *For primes* $r \neq s$ *dividing* $\phi(n)$ *it holds with probability at most* $4/rs$ *that*

$$\sum_{l=0}^{7}(-1)^l \sum_{i,j=1}^{m} f_i^{[l]}\, g_j^{[l]}\, d_i\, \bar{d}_j = 0 \bmod rs. \tag{9}$$

Proof. The proof is almost identical to the proof of Lemma A.5. The only difference is, how we show that there is a $a_{i',j'} \neq 0$. This is obvious for $\bar{\imath} = \bar{\jmath}$. Otherwise, we have $a_{i',j'} = 0$ for $i' = \min(\bar{\imath}, \bar{\jmath})$ and $j' = \max(\bar{\imath}, \bar{\jmath})$ which is equivalent to

$$\bar{a}_{\bar{\imath},\bar{\jmath}}\,(\bar{\jmath} + 3m) + \bar{a}_{\bar{\jmath},\bar{\imath}}(\bar{\imath} + 3m) = 0 \bmod r.$$

Since $r \gg 32m$ this equation holds in $\mathbf{Z}$ as well and we get

$$\frac{\bar{\imath} + 3m}{\bar{\jmath} + 3m} = \frac{\bar{a}_{\bar{\imath},\bar{\jmath}}}{-\bar{a}_{\bar{\jmath},\bar{\imath}}}.$$

Since the right hand side of the last equation is either equal to 1 or no closer to 1 than 3/4 or 4/3, this equation cannot hold. This shows that $a_{i',j'} \neq 0$.

■

Lemma. B.2 *With probability at least* $(1 - 4\binom{m}{k}^{12}/rs)$ *(over a randomly chosen* $\mathbf{d}$*) there are no colourful 6-cycles in* G^0.

Proof. For every edge disjunct colourful 6 cycle, we get an equation

$$\sum_{l=1}^{3}\sum_{i,j=1}^{m} f_i^{[l]} g_j^{[l]} d_i \bar{d}_j = \sum_{l=4}^{6}\sum_{i,j=1}^{m} f_i^{[l]} g_j^{[l]} d_i \bar{d}_j \bmod rs.$$

Thus by Lemma B.1 with probability at least $(1 - 4N^6/rs)$ we get an equation

$$\sum_{l=1}^{3}\mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} = \sum_{l=4}^{6}\mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]}$$

for every edge disjoint colourful 6 cycle. If the cycle is colourful then it holds that $f^{[1]}, f^{[2]}, f^{[3]}$ are not all equal and that $g^{[1]}, g^{[2]}, g^{[3]}$ are not all equal. We set $A := \mathbf{f}^{[1]} \otimes \mathbf{g}^{[1]} + \mathbf{f}^{[2]} \otimes \mathbf{g}^{[2]} + \mathbf{f}^{[3]} \otimes \mathbf{g}^{[3]}$.

If $\mathbf{f}^{[2]} = \mathbf{f}^{[3]}$ we can determine $\mathbf{f}^{[1]}$ from $(f_1)_j = 1 \iff \sum_{i=1}^m A_{ij} \in \{k, 3k\}$ and $\mathbf{g}^{[1]}$ as A^t_j for any j with $\sum_{i=1}^m A_{ij} = k$. This uniqueness of $\mathbf{f}^{[1]} \otimes \mathbf{g}^{[1]}$ contradicts the edge disjunctness of the cycle. Analogously we can conclude that $\mathbf{f}^{[1]} \neq \mathbf{f}^{[3]}$, $\mathbf{f}^{[1]} \neq \mathbf{f}^{[2]}$, $\mathbf{g}^{[2]} \neq \mathbf{g}^{[3]}$, $\mathbf{g}^{[1]} \neq \mathbf{g}^{[3]}$, $\mathbf{g}^{[1]} \neq \mathbf{g}^{[2]}$.

Now we assume that the $\mathbf{f}^{[l]}$, $(l = 1, 2, 3)$, are all distinct and the $\mathbf{g}^{[l]}$, $(l = 1, 2, 3)$, are all distinct. First we show that $\mathbf{g}^{[1]}, \mathbf{g}^{[2]}, \mathbf{g}^{[3]}$ are uniquely determined up to permutation by A and therefore $\{\mathbf{g}^{[1]}, \mathbf{g}^{[2]}, \mathbf{g}^{[3]}\} = \{\mathbf{g}^{[4]}, \mathbf{g}^{[5]}, \mathbf{g}^{[6]}\}$.

If there are 3 (as vectors) distinct columns j_1, j_2, j_3 of A with $\sum_{i=1}^m A_{ij_l} = 2k$ for $l = 1, 2, 3$ then for all $a, b \in \{1, 2, 3\}$ it holds that $(f_a)_{j_b}$ is 1 iff $a \neq b$ and is 0 iff $a = b$, and $\mathbf{g}^{[1]}, \mathbf{g}^{[2]}, \mathbf{g}^{[3]}$ are uniquely determined up to permutation by
$\mathbf{g}^{[1]} = A^t_{j_3} - A^t_{j_1} + A^t_{j_2}$, $\mathbf{g}^{[2]} = A^t_{j_1} - A^t_{j_2} + A^t_{j_3}$ and $\mathbf{g}^{[3]} = A^t_{j_2} - A^t_{j_3} + A^t_{j_1}$.

If there no such 3 columns in A we have w.l.o.g.

$$\forall j: \ (f_1)_j = (f_2)_j = 1 \ \Rightarrow \ (f_3)_j = 1. \tag{10}$$

Now we can conclude that since $\mathbf{f}^{[1]} \neq \mathbf{f}^{[3]}$, $\mathbf{f}^{[2]} \neq \mathbf{f}^{[3]}$ and (10) holds, there are j_1, j_2 with $(f_i)_{j_1} = 1$ iff $i = 1$ and $(f_i)_{j_2} = 1$ iff $i = 2$. Thus there are at least 2 (as vectors) distinct columns j of A satisfying $\sum_{i=1}^m A_{ij} = k$. If there are 3 of them then they are equal to $\mathbf{g}^{[1]}, \mathbf{g}^{[2]}, \mathbf{g}^{[2]}$ and if there are only 2 of them they are equal to $\mathbf{g}^{[1]}, \mathbf{g}^{[2]}$ and we can easily determine $\mathbf{g}^{[3]}$ from A. Again $\mathbf{g}^{[1]}, \mathbf{g}^{[2]}, \mathbf{g}^{[3]}$ are uniquely determined up to permutation by A.

Analogously one can see that $\mathbf{f}^{[1]}, \mathbf{f}^{[2]}, \mathbf{f}^{[3]}$ are uniquely determined up to permutation by A and therefore $\{\mathbf{f}^{[1]}, \mathbf{f}^{[2]}, \mathbf{f}^{[3]}\} = \{\mathbf{f}^{[4]}, \mathbf{f}^{[5]}, \mathbf{f}^{[6]}\}$. We get $\mathbf{f}^{[1]} \otimes \mathbf{g}^{[1]} + \mathbf{f}^{[2]} \otimes \mathbf{g}^{[2]} + \mathbf{f}^{[3]} \otimes \mathbf{g}^{[3]} = \mathbf{f}^{[1]} \otimes \mathbf{g}_{i_1} + \mathbf{f}^{[2]} \otimes \mathbf{g}_{i_2} + \mathbf{f}^{[3]} \otimes \mathbf{g}_{i_3}$ with $\{i_1, i_2, i_3\} = \{1, 2, 3\}$. Now considering a j satisfying $(f_1)_j = 0$ and $(f_2)_j = 1$ it is easy to see that $i_k = k$ for $k = 1, 2, 3$.

■

Lemma. B.3 *With probabilty at least* $(1 - 4\binom{m}{k}^{12}/rs)$ *(over a randomly chosen* $\mathbf{d}$*) for every subgraph* $\tilde{G} = (\tilde{V}_1, \tilde{V}_2, \tilde{E})$ *of* G *with* F*-monochromic* $\tilde{V}_1$ *the following fact holds:*

Fact. B.4 *For any* $x_1, x_2 \in V_1'$ *the number of 4-paths* (x_1, a, b, c, x_2) *with* $F(x_1, a) \neq F(a, b)$ *and* $F(b, c) \neq F(c, x_2)$ *is bounded by*

- $2M$ *if* x_1 *and* x_2 *are of different F-colours.*
- $2Md(x_2)$ *if* x_1 *and* x_2 *are of the same F-colour,*

where $M := \max_{\mathbf{f}}(|\tilde{V}_1^F(\mathbf{f})|)$.

Proof. Let there be 2 such 4-paths from x_1 to x_2. Since $\tilde{V}_1$ is F-monochromic for every such 8 cycle we get an equation

$$\sum_{l=1}^{4}(-1)^l \sum_{i,j=1}^{m} f_i^{[l]} g_j^{[l]} d_i \bar{d}_j = \sum_{l=5}^{8}(-1)^l \sum_{i,j=1}^{m} f_i^{[l]} g_j^{[l]} d_i \bar{d}_j \bmod rs$$

with $f^{[1]} = f^{[8]}, f^{[4]} = f^{[5]}, f^{[2]} = f^{[3]}, f^{[6]} = f^{[7]}$. Thus by Lemma B.1 with probability at least $(1 - 4\binom{m}{k}^{12}/rs)$ we get the equation

$$\sum_{l=1}^{4}(-1)^l \mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]} \ = \ \sum_{l=5}^{8}(-1)^l \mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]}.$$

We set $A := \sum_{l=1}^{4}(-1)^l \mathbf{f}^{[l]} \otimes \mathbf{g}^{[l]}$. In A there are at most 3 (as vectors) distinct columns j that contain 1's and -1's:

$$\begin{array}{ll} a) & f_j^{[1]} = f_j^{[4]} = 1 \quad f_j^{[2]} = 1 \\ b) & f_j^{[1]} = f_j^{[4]} = 1 \quad f_j^{[2]} = 0 \\ c) & f_j^{[1]} = f_j^{[4]} = 0 \quad f_j^{[2]} = 1. \end{array}$$

$\mathbf{f}^{[2]}$ is determined by an assignment of the cases $a) - c)$ to the columns. For each column j that contains 1's and -1's it is uniquely determined whether it is of type $c)$ or $a) - b)$. Thus there are only 2 possibillities for $\mathbf{f}^{[2]}$. Now fix $\mathbf{f}^{[2]}$.

1. Let x_1 and x_2 be of different F-colours. We show that $\mathbf{g}^{[1]}$ and $\mathbf{g}^{[4]}$ are uniquely determined by A and $\mathbf{f}^{[2]}$:

If for all j it holds that $f_j^{[2]} = 1 \Leftrightarrow f_j^{[1]} \neq f_j^{[4]}$ then there are j_1, j_2, j_3 so that for all $a, b \in \{1, 2, 3\}$ it holds that $f_{j_b}^{[a]}$ is 1 iff $a \neq b$ and is 0 iff $a = b$ and $\mathbf{g}^{[1]}, \mathbf{g}^{[4]}$ are uniquely determined by $2\mathbf{g}^{[1]} = A_{j_3}^t - A_{j_1}^t + A_{j_2}^t$ and $2\mathbf{g}^{[4]} = A_{j_1}^t - A_{j_3}^t + A_{j_2}^t$.

If there is a j satisfying $f_j^{[2]} = 0$ and $f_j^{[1]} \neq f_j^{[4]}$ then A_j^t equals $\mathbf{g}^{[1]}$ or $\mathbf{g}^{[4]}$ and one can easily determine the other value from A.

2. Let x_1 and x_2 be of the same F-colour. $\mathbf{g}^{[1]}$ is uniquely determined by $\mathbf{g}^{[4]}$ and A.

For fixed $\mathbf{f}^{[1]} \otimes \mathbf{g}^{[1]}$ and $\mathbf{f}^{[2]}$ there are at most $|\tilde{V}_1^F(\mathbf{f}^{[2]})|$ many possible values $\mathbf{g}^{[2]}$ and $\mathbf{g}^{[3]}$ is uniquely determined by $\mathbf{f}^{[2]} \otimes \mathbf{g}^{[2]}$ and $\mathbf{f}^{[4]} \otimes \mathbf{g}^{[4]}$.

■

Lemma. B.5 *Let $G = (V_1, V_2, E)$ be a graph with $|V_1|, |V_2| \leq v$. For any vertex x let $U(x)$ denote the set of vertices $y \in N(x)$ that are F-dominated. Then G contains a subgraph $\tilde{G} = (\tilde{V}_1, \tilde{V}_2, \tilde{E})$ with $|\tilde{E}| > |E| - 4v \log_2 v$ so that for all $x \in \tilde{V}_1$ it holds that*

$$\max_{y \in U(x)} (d(y) - d_{F(x,y)}(y)) \leq 1/2 \sum_{y \in U(x)} d(y) - d_{F(x,y)}(y) \tag{11}$$

Proof. We consider the function $D(G) := \sum_{x \in V_1} \sum_{y \in U(x)} d(y) - d_{F(x,y)}(y)$. For every $x \in V_1$ for that (11) doesn't hold we remove the edge (x, y_x) with $y_x \in U(x)$ and $d(y_x) - d_{F(x,y_x)}(y_x) = \max_{y \in U(x)}(d(y) - d_{F(x,y)}(y))$. By that procedure we decrease $D(G)$ at least by a factor of 2. Since $D(G) \leq \sum_{x \in V_1} \sum_{y \in N(x)} d(y)$, which is at most $|E|^2$, we can perform this procedure at most $\log_2(|E|)$ many times.

■

$E(U_1, U_2) := \{(x, y) \in E \mid x \in U_1 \wedge y \in U_2\}$.

Proof of Theorem 4.4 Assume that $\delta = |V| < 2^{-4.75}|E|N^{-1/4}$. Then, by Lemma B.2 and B.3, with probability at least $1 - N^6/rs$ there is no collision, no colourful 6-cycle and fact B.4 holds. Assume that this is the case and that $|E| > 2^{28}$. We divide E into the following subsets:

$$\begin{array}{lll} L_1 & := & \{(x, y) \in E \mid x \text{ is F-dominated}\} \\ L_2 & := & \{(x, y) \in E \mid x \text{ is G-dominated}\} \\ L_3 & := & \{(x, y) \in E \mid y \text{ is F-dominated}\} \\ L_4 & := & \{(x, y) \in E \mid y \text{ is G-dominated}\} \\ L_5 & := & \{(x, y) \in E \mid x, y \text{ are colourful}\} \end{array}$$

We have $E = \bigcup_{i=1}^{5} L_i$ and $|E| < N$. Subsequently we use variables ϵ_i in the estimations which will be fixed at the end of the proof.

If $|\mathbf{L_1}| > \epsilon_1 |\mathbf{E}|$ then G contains a subgraph $G^1 = (V_1^1, V_2^1, E^1)$ so that V_1^1 is F-monochromic and $|E^1| > \frac{3}{4}\epsilon_1|E|$. If the number of F-colours $\mathbf{f}$ satisfying $|V_1^{1,F}(\mathbf{f})| > N^{1/4}$ is greater than $\epsilon_2|E^1|N^{-1/2}$ we have $v > \frac{3}{4}\epsilon_1\epsilon_2|E|N^{-1/4}$.

On the other hand if the number of those colours is at most $\epsilon_2|E^1|N^{-1/2}$ using Lemma B.5 we see that G^1 contains a subgraph $G^2 \subseteq (V_1^2, V_2^2, E^2)$ so that (11) holds for $x \in V_1^2$, for all F-colours $\mathbf{f}$ it holds that $|V_1^{2,F}(\mathbf{f})| \leq N^{1/4}$ and $|E^2| > (1-\epsilon_2)|E^1| - 4v\log_2 v = \alpha_1|E|$. Let $U_2 := \{x \in V_2^2 \mid x \text{ is F-dominated}\}$. We distinguish two cases:

1. If $E^2(V_1^2, U_2) > \epsilon_3|E^2|$ then G^2 contains a subgraph $\tilde{G} = (\tilde{V}_1, \tilde{V}_2, \tilde{E})$ so that $\tilde{V}_1$ and $\tilde{V}_2$ are F-monochromic and $|\tilde{E}| > \frac{3}{4}\epsilon_3|E|$. The number of it's vertices $|\tilde{V}_1| + |\tilde{V}_2|$ is at least $2\sum_{\mathbf{f}}\sqrt{|\tilde{V}_1^F(\mathbf{f})|\cdot|\tilde{V}_2^F(\mathbf{f})|}$ which is no more than $2N^{-1/4}\sum_{\mathbf{f}}|\tilde{V}_1^F(\mathbf{f})|\cdot|\tilde{V}_2^F(\mathbf{f})|$. Since $|\tilde{E}| \leq \sum_{\mathbf{f}}|\tilde{V}_1^F(\mathbf{f})|\cdot|\tilde{V}_2^F(\mathbf{f})|$ we get
$|V| > \frac{3}{4}\epsilon_3\alpha_1|E|N^{-1/4}$.

2. If $|E^2(V_1^2, V_2^2 - U_2)| > (1-\epsilon_3)|E^2|$ we set $\hat{G} = (V_1^2, V_2^2 - U_2, E^2(V_1^2, V_2^2 - U_2))$. Since $\hat{V}_1$ is F-monochromic and $|\hat{V}_1^F(\mathbf{f})| \leq N^{1/4}$ holds for all F-colours $\mathbf{f}$, using fact B.4 we see that the number of 4-paths (x_1, a, b, c, x_2) with $x_1, x_2 \in \hat{V}_1$, $F(x_1, a) \neq F(a, b)$ and $F(b, c) \neq F(c, x_2)$ is bounded by
$\sum_{x_2 \in \hat{V}_1}\left(2N^{1/4}|V| + 2d(x_2)N^{1/4}\right)$, which is at most $4|V|^2N^{1/4}$.

On the other hand since $\hat{V}_1$ is F-monochromic the number of those 4-paths is at least

$$\sum_{(w,x)\in\hat{E}} \sum_{\substack{y\in N(x)\\ F(w,x)\neq F(y,x)}} \sum_{z\in N(y)\setminus\{x\}} d(z) - d_{F(y,z)}(z).$$

Since (11) holds for $x \in \hat{V}_1$, we can estimate this by

$$\begin{aligned} & 1/2 \sum_{(w,x)\in\hat{E}} \sum_{\substack{y\in N(x)\\ F(w,x)\neq F(y,x)}} \sum_{z\in N(y)} d(z) - d_{F(y,z)}(z) \\ = \; & 1/2 \sum_{y\in\hat{V}_1}\Big(\sum_{z\in N(y)} d(z) - d_{F(y,z)}(z)\Big)^2. \end{aligned}$$

Using the Cauchy-Schwarz inequality and that all $z \in \hat{V}_2$ are not F-dominated this is at least $2^{-5}|V|^{-1}\Big(\sum_{(y,z)\in\hat{E}} d(z)\Big)^2$, which equals $2^{-5}|V|^{-1}\Big(\sum_{z\in\hat{V}_2} d(z)^2\Big)^2$. Finally we get the bound $2^{-5}|V|^{-3}|\hat{E}|^4$ which yields $|V| > 2^{-7/5}(1-\epsilon_3)^{4/5}\alpha_1^{4/5}|E|N^{-1/4}$.

If $|\mathbf{L_i}| \geq \epsilon_1|\mathbf{E_0}|$ for any $\mathbf{i} \in \{\mathbf{2}, \mathbf{3}, \mathbf{4}\}$ we get the same estimations analogously.

Now let $|\mathbf{L_5}| \geq (\mathbf{1} - \mathbf{4}\epsilon_1)|\mathbf{E_0}|$. Set $G^4 := (V_1, V_2, L_5)$ and let E_4^0 denote the set of edges $(x, y) \in E^4$ with $|A_x(y)| \leq 7/8\, d(x)$ and $|A_y(x)| \leq 7/8\, d(y)$. By Lemma A.8 we get $|E_4^0| > |E_4| - 2|V|$.

Since there are no colourful 6-cycles in G^4 the number of coloured 3-paths, i.e. the paths (w, x, y, z) in G^4 with $w \in V_2^4$, $F(w,x) \neq F(x,y) \neq F(y,z)$ and $G(w,x) \neq G(x,y) \neq G(y,z)$, is at most $|V|^2$.

On the other hand it is at least

$$\sum_{(x,y)\in E_0^4} \Big(d(x) - |A_x(y)|\Big)\Big(d(y) - |A_y(x)|\Big)$$

which is greater than $2^{-6}\sum_{(x,y)\in E_0^4} d(x)d(y)$. Using the identity $\sum_{(x,y)} d(x)^{-1} = \sum_{(x,y)} d(y)^{-1} = |V|$ we can estimate this by $2^{-6}\min\left(\sum a_i b_i \mid \sum a_i^{-1} + b_i^{-1} \leq 2|V|\right)$, where the minimum is taken over all $\vec{a}, \vec{b} \in \mathbf{R}^{|E_0^4|} - \{\vec{0}\}$. The minimum occurs, if all a_i and b_i are equal. Setting $|E_0^4| \geq |E^4| - 2|V| = \alpha_2|E|$ we get the bound $2^{-6}\alpha_2^3|E|^3|V|^{-2}$ which yields $|V| > 2^{-3/2}\alpha_2^{3/4}|E|N^{-1/4}$.

Assume that $|V| < 2^{-4.75}|E|N^{-1/4}$ and $|E| > 2^{28}$. Set $\epsilon_1 = 0.2$, $\epsilon_2 = 0.3$, $\epsilon_3 = 0.5$. Since $N \geq 2^{92}$ we get $\alpha_1 \geq 0.11$, $\alpha_2 \geq 0.2$ and finally $|V| > 2^{-4.75}|E|N^{-1/4}$ which is a contradiction. This completes our proof.

■

References

[And92] R. J. Anderson. Attack on server assisted authentication protocols. *Elect. Lett.*, 28(15):1473, 1992.

[Bol78] B. Bollobas. *Extremal graph theory*, volume 11 of *L. M. S. Monographs.* Academic Press. XX, London-New York-San Francisco, 1978.

[Kaw95] S. Kawamura. Information leakage meassurement in a distributed computation protocol. *IEICE Trans. Fundamentals*, E78-A(11):59–66, 1995.

[KS93] S. Kawamura and A. Shimbo. Fast server-aided secret computation protocols for modular exponentiaion. *IEEE JSAC*, 11(5):778–784, 1993.

[LL95] C. H. Lim and P. J. Lee. Security and performance of server-aided RSA computation protocolls. In *Advances in Cryptology - Proceedings of Crypto'95*, volume 963 of *Lecture Notes in Computer Science*, pages 70–83. Springer Verlag, 1995.

[LW92] J. H.van Lint and R. M. Wilson. *A course in combinatorics.* Cambridge University Press, 1992.

[MILY93] T. Matsumoto, H. Imai, C. S. Laih, and S. M. Yen. On verifiable implicit asking protocols for RSA computation. In *Advances in Cryptology - Proceedings of Auscrypt'92*, volume 718 of *Lecture Notes in Computer Science*, pages 296–307. Springer Verlag, 1993.

[MKI89] T. Matsumoto, K. Kato, and H. Imai. Speeding up computation with insecure auxiliary devices. In *Advances in Cryptology - Proceedings of Crypto'88*, volume 403 of *Lecture Notes in Computer Science*, pages 497–506. Springer Verlag, 1989.

[Nec94] V. I. Nechaev. Complexity of a determinate algorithm for the discrete logarithm. *Mathematical Notes*, 55(2):165–172, 1994.

[PW93] B. Pfitzmann and M. Waidner. Attacks on protocols for server-aided RSA computation. In *Advances in Cryptology - Proceedings of Eurocrypt'92*, volume 658 of *Lecture Notes in Computer Science*, pages 153–162. Springer Verlag, 1993.

[QS91] J. J. Quisquater and M. De Soete. Speeding up smart card RSA computation with insecure coprocessors. In *Proceedings of Smart Card 2000*, pages 191–197. North Holland, 1991.

[Sho] V. Shoup. Lower bounds for discrete logarithms and related problems. In *Advances in Cryptology - Proceedings of Eurocrypt'97*.

[vOW96] P. van Oorschot and M. Wiener. Improving implementable meet-in-the-middle attacks by orders of magnitude. In *Advances in Cryptology - Proceedings of Crypto'96*, volume 1109 of *Lecture Notes in Computer Science*, pages 229–236. Springer Verlag, 1996.