Zero-knowledge Identification based on Lattices with Low Communication Costs

Rosemberg Silva[1], Pierre-Louis Cayrel[2], Richard Lindner[3]

[1]State University of Campinas (UNICAMP)
Institute of Computing
rasilva@ic.unicamp.br – Brazil

[2]Laboratoire Hubert Curien
Université de Saint-Etienne
pierre-louis.cayrel@univ-st-etienne.fr – France

[3]Technische Universität Darmstadt
Fachbereich Informatik
Kryptographie und Computeralgebra
rlindner@cdc.informatik.tu-darmstadt.de – Germany

Abstract. *In this paper we propose a new 5-pass zero-knowledge identification scheme with soundness error close to 1/2. We use the hardness of the Inhomogeneous Small Integer Solution problem as security basis. Our protocol achieves lower communication costs compared with previous lattice-based zero-knowledge identification schemes. Besides, our construction allows smaller public and secret keys by applying the use of ideal lattices. We allow the prover to possess several pairs of secret and public keys, and choose randomly which pair is to be used in a given round of execution. We also dealt with nonces in zero-knowledge schemes in a new way, lowering the number of values exchanged between the prover and the verifier. Hence, our scheme has the good features of having a zero-knowledge security proof based on a well known hard problem of lattice theory, with worst to average-case reduction, and small size of secret and public keys.*

1. Introduction

Identification schemes are cryptographic tools applied to provide access control. A common realization of such schemes comes in the form of protocols between two entities called Prover and Verifier. The former is expected to establish the validity of its identity to the latter. In order to accomplish such goal, the Prover makes used of the knowledge of a secret key, that is connected to its identity. The application of zero-knowledge constructions helps to ensure that no information about such key is leaked as the protocol is performed. The only knowledge gained by the Verifier is that the claim made by the Prover about the possession of the secret key is valid with overwhelming probability. Lattice-based instantiations of this kind of protocol have recently been proposed: [Lyubashevsky 2008], [Kawachi et al. 2008] and [Cayrel et al. 2010]. An interesting feature of some lattice-based systems in Cryptography, as seen in the seminal work from Ajtai [Ajtai 1996], is the possibility of allowing reductions from wost-cases to average-cases of well known hard problems. In the present work we use some of these results and propose an identification scheme that achieves better communication costs.

There is an standard construction of signature schemes from identification schemes through the usage of the Fiat-Shamir heuristic, secure under the random oracle model [Fiat and Shamir 1986]. It is of pivotal importance, however, that the communication costs of the identification scheme be small in order to have efficient conversion.

Among the good characteristics that our system possesses we stress the resilience to existing quantum attacks, like Shor's [Shor 1994]. In addition to that, the relatively small soundness error per round and the algorithm that prevents exchange of large vectors enables to achieve small communication costs in contrast to other lattice-based solutions, such as those seen in [Lyubashevsky 2008], [Kawachi et al. 2008] and [Cayrel et al. 2010]. This is conveyed by the Table 1, where we consider a scenario with an overall soundness error of 2^{-16}, and make the assumption that all the random choices involve the use of seeds that are 128 bits long. For a setting applying k pairs of keys, our scheme has soundness error $1/2 + 1/k^2$. Then, it suffices to run 17 rounds of the protocol in order to reach such goal. The best zero-knowledge identification scheme in term of communication cost, the CLRS scheme has soundness error of $(q+1)/q$, considering that it is based on q-ary lattices over $\mathbb{Z}_q$. Given that we are working with $q = 257$, it also requires 17 rounds in order to satisfy the requirement regarding overall soundness error. Our proposal reaches a communication cost better than CLRS'.

Scheme	Rounds	Total communication [Kbyte]
Lyubashevsky [Lyubashevsky 2008]	11	110,00
Kawachi et al. [Kawachi et al. 2008]	27	58,67
CLRS [Cayrel et al. 2010]	17	37,50
Ours	17	23,40

Table 1. Comparison with lattice-based ID schemes.

This paper is organized as follows. The concepts necessary to the understanding of our construction are are presented in Section 2. After that, we provide a detailed description of the algorithms from which our scheme is composed in Section 3. Then, we give the proofs for the zero-knowledge properties that our scheme possesses. For last, we discuss the choice of parameters in Section 4 and future lines of investigation in Section 5.

2. Preliminaries

Notation. Along the text we use bold capital letters to denote matrices and bold small letters to indicate vectors. Scalars, on their turn, are represented with regular characters.

Security Model. Our identification scheme employs a string commitment scheme that possesses the properties of computationally binding and statistically hiding. We also assume that a trusted party honestly sets up the system parameters for both the Prover and the Verifier.

A commitment scheme is said to be statistically hiding, when no computationally unbounded adversary can distinguish two commitment strings generated from two distinct

strings. It is computationally binding, if no polynomial-time adversary can imperceptibly change the committed string after sending the corresponding commitment.

We show that our construction is resilient to concurrent attacks. Thus, we allow the adversaries act as cheating verifiers prior to attempting impersonation attacks, with polynomially many different instances of the prover. Such instances share the same secret keys in each round, but use different random coins. Besides, they have their own independent state.

Zero-Knowledge. Given a language L and one of its words x, we call zero-knowledge proof of knowledge that x belongs L a protocol with the following properties:

- Completeness: any true theorem to be proved.
- Soundness: no false theorem can be proved.
- Zero-Knowledge: nothing but the truthfulness of the statement being proved is revealed. According to [Goldwasser et al. 1985], there exist the following kinds of zero-knowledge :
  - Perfect: if the simulator and the real proof produce communication tapes with exactly the same distribution.
  - Statistical: if the simulator and the real proof produce communication tapes whose statistical distributions are negligibly close.
  - Computational: if no efficient algorithm can distinguish the communication tape produced by the simulator from that corresponding to the real protocol.

Lattices. Lattices correspond to discrete additive subgroups of $\mathbb{R}^m$. One can represent them by means of a basis $\mathbf{B}$ composed of $n \leq m$ linear independent vectors in $\mathbb{R}^m$. Such basis is not unique. Given two basis $\mathbf{B}_1$ and $\mathbf{B}_2$ for the same lattice, there is a unimodular matrix $\mathbf{U}$ such that $\mathbf{B}_1 = \mathbf{B}_2\mathbf{U}$.

Given a basis $\mathbf{B}$, the corresponding lattice is generated by all combinations of vectors in $\mathbf{B}$ with integer coefficients.

There are hard problems defined over lattices that can be used as security assumptions in the construction of cryptographic schemes. We list below the problems that are related to the identification scheme proposed in this paper. We assume that the norm used throughout this document is max-norm.

Definition 2.1 (Search Shortest Vector Problem, SVP)**.** On input a lattice basis $\mathbf{B} \in \mathbb{Z}^{m\times n}$, the computational shortest vector problem (SVP) is defined as the task of obtaining a non-zero lattice vector $\mathbf{Bx}$ satisfying $\|\mathbf{Bx}\| \leq \|\mathbf{By}\|$ for any other $\mathbf{y} \in \mathbb{Z}^n \setminus \{0\}$.

As commonly happens in the study of computational complexity, one can express the problem above as optimization, decision and approximation [Micciancio and Goldwasser 2002].

Definition 2.2 (Short Integer Solution, SIS)**.** On input a prime q, a positive real number b, a matrix $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$, associated with a lattice basis, the short integer solution (SIS) problem is defined as the task of finding a non-zero vector $\mathbf{x} \in \mathbb{Z}^m$ such that $\mathbf{Ax} = \mathbf{0} \pmod q$, with $\|\mathbf{x}\| \leq b$.

The SIS has served as security assumption in several cryptographic schemes with worst-case connections. In [Ajtai 1996] and [Ajtai and Dwork 1996], Ajtai first showed how to use computationally intractable worst-case lattice problems as building blocks for cryptosystems. The parameter sizes involved, however, are not small enough to enable practical implementations.

Working towards making lattice-based schemes practical, Micciancio showed that it is possible to represent a basis, and thus public keys, with space that grows quasi-linearly in the lattice dimension [Micciancio 2007] through cyclic lattices. Further improvement was obtained in conjunction with Lyubashevsky, achieving compression functions that are both efficient and provably secure assuming the hardness of worst-case lattice problems for ideal lattices [Lyubashevsky and Micciancio 2006].

A variety of hard problems associated with lattices has been used as security basis in a number of cryptographic schemes. For example, Lyubashevsky's identification scheme is secure under active attacks, assuming the hardness of approximating SVP in all lattices of dimension n to within a factor of $\tilde{O}(n^2)$. By weakening the security assumption, on the other hand, one can achieve parameters small enough to make a practical implementation feasible, as seen in the identification scheme proposed by Kawachi et al. in [Kawachi et al. 2008]. In this later work, the authors suggest to use approximate Gap-SVP or SVP within factors $\tilde{O}(n)$. Cayrel et al. improved the results from Kawachi in the CLRS scheme [Cayrel et al. 2010].

Ideal Lattices. There is a particular class of lattices that are closed under ring multiplications. They correspond to the ideals of some polynomial quotient ring.
Definition 2.3 (Ideal lattices)**.** Let f be a monic polynomial of degree n. We call L is an *ideal lattice* if it corresponds to an ideal I in the ring $\mathbb{Z}[x]/\langle f\rangle$.

For lattices of rank n and entries from $\mathbb{Z}_q$, the storage needs for characterizing a basis is $n^2\log(q)$ bits. By applying ideal lattices, instead, this value can be reduced to $n\log(q)$ bits. Besides the gain in storage, there is also a gain in performance, given that matrix/vector multiplications can be performed in time $O(n\log(n))$ using discrete FFTs.

Both SIS and SVP restricted to ideal lattices still keep the worst-case to average-case connection, provided that f is irreducible over the integers.
Definition 2.4 (Ideal-SIS)**.** Given a monic polynomial f of degree n, the ring $R_f = \mathbb{Z}[x]/\langle f\rangle$, and m elements $a_1,\ldots,a_m \in R_f/qR_f$, we define *Ideal-SIS* the problem of finding $x_1,\ldots,x_m \in R_f$ satisfying $\sum_{i=1}^{m} a_i x_i = 0 \pmod q$ and $0 < \|(x_1,\ldots,x_m)\| \leq b$.

Canonical identification schemes Let the tuple $(\textsc{KeyGen}, P, V)$ be an identification scheme, where KeyGen is the key-generation algorithm which on input parameters outputs $(\mathrm{sk},\mathrm{pk})$, P is the prover algorithm taking input sk, V is the verifier algorithm taking inputs parameters and pk. We say that such scheme is a canonical identification scheme if it is a public-coin 3-move protocol $\{$commitment, challenge, answer$\}$ that enables P to convince V about the knowledge of sk.

Stern's Identification Scheme. The first code-based identification scheme was proposed by Harari in 1989 [Harari 1988], but it was broken by Véron in 1995 [Véron 1995]. In 1990, Girault devised a three-pass scheme [Girault 1990]. Neither of those schemes was practical, from performance perspective, though. The first practical code-based identification scheme, from performance standpoint, was proposed by Stern [Stern 1993]. Its security relies on the collision-resistance property of a hash function h and the hardness of the syndrome decoding problem. The entity to be authenticated possesses a pair of keys $(\mathbf{i}, \mathbf{s})$ related by $\mathbf{i} = \mathbf{H}^T\mathbf{s}$, where $\mathbf{H}$ is a public parity check matrix of a given code, $\mathbf{s}$ is a private binary vector of Hamming weight p, and $\mathbf{i}$ is its public syndrome. It engages in a zero-knowledge protocol with a verifier, aiming to prove the knowledge of the secret key, without revealing its value. In the same paper Stern proposed some variations of the protocol. The most efficient in terms of communication costs had soundness error of 2/3. This implies that, in order to reach a confidence level L on the authenticity of the prover, the protocol has to be repeated a number r of rounds, in order to satisfy $1 - (2/3)^r \geq L$.

Gaborit and Girault's Identification Scheme Another scheme suggested by Gaborit and Girault requires smaller storage for public data [Gaborit and Girault 2007]. Given that the schemes we have seen are dealing with codes, this usually implies that a generator matrix or a parity check matrix is needed to fully characterize them. The idea applied by Gaborit and Girault was to use double-circulant matrices for a compact representation.

In our work, we point out that a combination of these two approaches (and the one from Aguilar et al. [Aguilar et al. 2011]) can be used in the lattice context, namely ideal lattices (which allow a very compact representation, as efficient as double-circulant matrices) for an identification scheme structure with soundness error of 1/2. With this, we manage to have the lowest communication costs and lowest public data storage needs.

3. Identification Scheme

Our scheme is comprised of two algorithms: key generation, and identification protocol. The first picks uniformly at random k binary vectors with length m, Hamming weight p and disjoint supports. The corresponding public keys are obtained by multiplication of the private key by a uniformly chosen matrix $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$. Given the small norm of the private keys, it is still computationally intractable to derive them from the public data for a suitable choice of parameters with algorithms known to this date. Such task corresponds to the inhomogeneous SIS problem. In addition to that, several valid private keys correspond to a given public key. This fact will be used later on to establish the concurrent security of our scheme.

KEYGEN:
$\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{n \times m}$
Compute $\mathbf{x} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ and $\mathbf{y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_k\}$ where
$\mathbf{x}_i \xleftarrow{\$} \{0,1\}^m$, with Hamming weight p and $1 \leq i \leq k$
$\mathbf{y}_i \longleftarrow \mathbf{A}\mathbf{x}_i \mod q$ with $1 \leq i \leq k$
$\text{COM} \xleftarrow{\$} \mathcal{F}$, suitable family of commitment functions
Output $(\text{sk}, \text{pk}) = (\mathbf{x}, (\mathbf{y}, \mathbf{A}, \text{COM}))$, where the private key is sk, and the public key is pk.

Figure 1. Key generation algorithm, parameters n, m, q are public.

The second algorithm corresponds to the identification protocol. To a given user we have associated k distinct pairs of keys. In a given round of execution, the Verifier

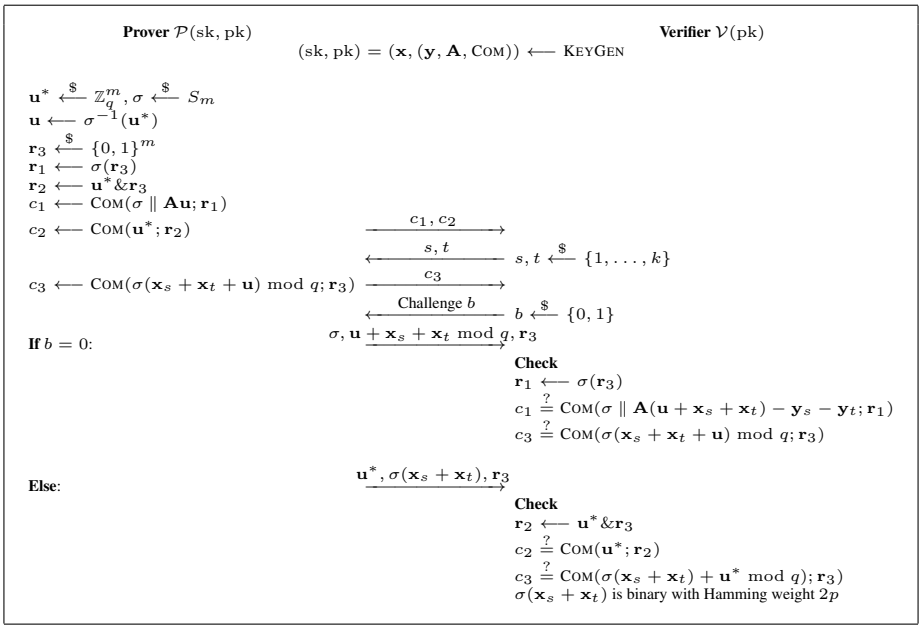

Figure 2. Identification protocol

picks two pairs of keys that the Prover is supposed to use for the interactive proof. When performing operations over the private keys, the Prover uses blinding factors in the form of permutations and vector sums with modular reduction. Both the permutation and the vector to be added to the private keys are uniformly randomly chosen. Hence, observing the outcome does not reveal any information about the private key value, given that its statistical distribution is uniform. This is applied in the demonstration of the zero-knowledge property of our scheme.

In order to help in the reduction of communication costs, the nonces that are used in conjunction with the COM functions can be generated in such a way that only one of the three values $\mathbf{r}_i$ is chosen uniformly at random. We can chose $\mathbf{r}_3$ because $\mathbf{c}_3$ is the common commitment that is checked for both possible values for the challenge b. The other two nonces may be obtained by performing operations involving the first one, either by applying a random permutation (σ) or by performing the bitwise logical AND with the appropriate number of bits of the permutation of a randomly chosen vector ($\mathbf{u}$). When replying to the challenges, the Prover would give to the Verifier all the seeds needed to reconstruct the nonces $\mathbf{r}_i$. The Prover also sends the seeds for computing σ and $\mathbf{u}^*$, depending on the challenge received from the Verifier, instead of sending matrices and vectors that define them. We are making the assumption that the utility function to derive them from the seeds are publicly known.

Regarding the computation of the blinding vector $\mathbf{u}$, we first randomly choose its image $\mathbf{u}^*$. Then, using the seed of the permutation σ, we obtain $\mathbf{u} \longleftarrow \sigma^{-1}(\mathbf{u}^*)$. This choice enables to, instead of replying the challenge $b = 1$ with a vector in $\mathbb{Z}_q^m$, send only a seed to reconstruct the value of $\mathbf{u}^*$ by the Verifier. This is the point where the improvement in terms of communication costs regarding CLRS becomes more evident.

For instantiating the commitment scheme COM, we recommend Kawachi's [Kawachi et al. 2008], whose security is based on SIS. As seen with CLRS and SWI-

FFT, the application of ideal lattices can bring improvement in performance and storage needs, without sacrificing security.

3.1. Security

We provide below proofs for the completeness, soundness and zero-knowledge properties of the identification scheme described in Figure 2. We use as assumptions the fact that the string commitment scheme COM is a statistically hiding and computationally binding commitment scheme, and also that the inhomogeneous SIS problem is hard.

3.1.1. Completeness

Given that an honest prover has knowledge of the private keys $\mathbf{x}_i$, the blending mask $\mathbf{u}$ and the permutation σ, he will always be able to derive the commitments c_1, c_2 and c_3, and reveal to the verifier the information necessary to check that they are correct. He can also show that the private keys in his possession has the appropriate Hamming weights. So, the verifier will always accept the honest prover's identity in any given round. This implies perfect completeness.

3.1.2. Zero-Knowledge

We give a demonstration of the zero-knowledge property for the identification protocol shown in Figure 2. Here, we require the commitment function COM to be statistically hiding, i.e., $\text{COM}(x; r)$ is indistinguishable from uniform for a uniform $r \in \{0,1\}^m$.

Theorem 3.1. *Let q be prime. The protocol described in Figure 2 is a statistically zero-knowledge proof of knowledge that the prover knows a set of k secret binary vectors $\mathbf{x}_i$ of Hamming weight p that satisfy $\mathbf{A}\mathbf{x}_i = \mathbf{y}_i \bmod q$, for $i \in \{1, \ldots, k\}$, if the employed commitment scheme* COM *is statistically-hiding.*

Proof. To prove the zero-knowledge property of our protocol, we construct a simulator S that, given oracle access to a verifier V (not necessarily honest), produces a communication tape that is statistically indistinguishable from a tape generated by the interaction between an honest prover P and the given verifier V. The construction of such simulated tape is described below. The simulator prepares to answer the second challenge b by guessing its value $\tilde{b}$ picked uniformly at random from $\{0,1\}$, and produces the corresponding commitments c_1 and c_2. It accesses the verifier, as an oracle, passing the commitments c_1 and c_2, obtaining as response the first challenge $\{s,t\}$. Then, it computes commitment c_3 and passes it to the verifier, obtaining the final challenge b. If b and $\tilde{b}$ match, the simulator records the interaction in the communication tape. Otherwise, it repeats the process. The number of rounds recorded r corresponds to what would be expected from a real pair (P, V) in order to reach a specified value for the overall soundness error L.

The computations of each commitment are executed as follows.

If $\tilde{b} = 0$, the simulator selects $\mathbf{u}$, σ and the nonces $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ as per protocol, computes the commitments $\{c_1, c_2\}$ and passes them to the verifier V, which responds with a challenge $\{s,t\}$. The simulator solves the equations $\mathbf{A}\mathbf{x}_t = \mathbf{y}_t \pmod q$ and $\mathbf{A}\mathbf{x}_s = \mathbf{y}_s$

$(\mathrm{mod}\ q)$ for $\mathbf{x}_t$ and $\mathbf{x}_s$, without any restriction regarding the norm of the solution. Such step is not computationally hard, and can be done in polynomial time. With these pseudo secret keys $\mathbf{x}_t$ and $\mathbf{x}_s$, the simulator computes c_3 according to the protocol and sends it to the verifier. The deviation in c_3 is not recognized because COM is statistically hiding. Upon receipt of c_3 the verifier responds with the final challenge b. If it matches the value to which the simulator had prepared to answer, namely $\tilde{b}$, then it reveals the values $\{\sigma, \mathbf{u} + \mathbf{x}_s + \mathbf{x}_t \bmod q, \mathbf{r}_1, \mathbf{r}_3\}$ to the verifier. The whole set of data exchanged between the simulator and the verifier is recorded. If there is not a match between b and $\tilde{b}$, however, the simulator does not record anything and prepares for a new round by picking another $\tilde{b}$ uniformly at random and restarts the process.

If $\tilde{b} = 1$, the simulator needs to play against the second verification branch. It selects $\mathbf{u}$, σ and the nonces $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ according to the protocol, uniformly at random. It computes the commitments $\{c_1, c_2\}$, sends them to the verifier, obtaining the answer t as result. Then, the simulator picks $\mathbf{x}_t$ and $\mathbf{x}_t$ uniformly at random as a binary vectors of dimension m, Hamming weight p and disjoint supports, without having to satisfy the restrictions $\mathbf{A}\mathbf{x}_s = \mathbf{y}_s \bmod q$ and $\mathbf{A}\mathbf{x}_t = \mathbf{y}_t \bmod q$, given that this will not be used to check the validity of the commitments, in case the guessed challenge is correct. It suffices that c_2 and c_3 can be reproduced by the verifier, and that the surrogate private keys $\mathbf{x}_t$ and $\mathbf{x}_t$ have Hamming weight p. The simulator computes commitment c_3, sends it to the verifier, and gets the final challenge as response. Again, such deviation is hidden by COM. In case the final challenge matches the guessed value, the whole interaction of this round is recorded. Otherwise, the simulator picks another $\tilde{b}$ and restarts the process. As a result, after $2r$ rounds in average, the simulator produces a communication tape statistically indistinguishable from a real one, provided that COM is statistically hiding. □

3.1.3. Soundness

Theorem 3.2. *If after r rounds of protocol execution a cheating prover is accepted with probability at least $(1/2+1/k^2)^r+\epsilon$, for any $\epsilon > 0$, either there is a knowledge extractor, which is a polynomial time probabilistic Turing machine that computes with overwhelming probability the sum of two private keys $\mathbf{x}_i$ and $\mathbf{x}_j$, with $i, j \in \{1, \ldots, k\}$, which are solutions to instances of the inhomogeneous SIS, or the binding property of the commitment scheme* COM *is broken.*

Proof. We follow the same strategy as Véron [Véron 1996] for reasoning about trees that model the probability space corresponding to the execution of the protocol. Let us suppose that a cheating prover can be accepted with such probability as $(1/2+1/k^2)^r+\epsilon$, or higher. Then, by rewinding this prover a number of times given by $1/\epsilon$, we can find with overwhelming probability a node with two sons in the execution tree associated with the protocol between the cheating prover and the verifier, corresponding to the reception of the two possible values for the challenge b. This means that such cheating prover will be able to answer all challenges for a fixed set of commitments c_1, c_2, c_3.

There are two possibilities for him to accomplish this. The first one is to have the arguments from which the commitments assuming different values for the two challenges,

but with similar images through the application of the function COM. This implies a violation of the binding property of the commitment scheme, given that a collision was found.

The second possibility is that the arguments to the function COM match. Let us call σ_0 and $\mathbf{u}_0 + \mathbf{x}_{s\_0} + \mathbf{x}_{t\_0}$ the values revealed by the cheating prover upon receipt of the challenge $b = 0$. Let us call $\sigma(\mathbf{u}_1)$ and $\sigma_1(\mathbf{x}_{s\_1} + \mathbf{x}_{t\_1})$ the answer to the challenge $b = 1$. Then, we can obtain the sum of the private key $\mathbf{x}_s + \mathbf{x}_t$ as $\sigma_0^{-1}(\sigma_1(\mathbf{x}_{t\_1} + \mathbf{x}_{t\_1}))$. This also means that we solved an arbitrary inhomogeneous SIS instance in probabilistic polynomial time, violating our assumption that such problem is hard.

The ability to find the sum of two arbitrary private keys also implies the ability to obtain each of the individual keys corresponding to such sum. In order to do that, we can use the fact that private keys are binary vectors with constant Hamming weight p. Then, using the pigeon hole principle, after finding k pair sums, we will have necessarily a non-empty intersection with the support set of the next pair sum. Such intersection corresponds to the support set of a single key. Given that the key is binary, it can be uniquely determined from its support set. Then, applying an XOR operation between each partially colliding sum and the recently computed key, we can retrieve the other two remaining private keys. This whole process can be executed over polynomially many key sums, so that all the individual private keys can be recovered.

□

In the security proofs above, we make the assumption that the distributions of $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r}_3$ are uniform, provided that $\mathbf{r}_3$ is randomly chosen from $\{0,1\}^n$, and that the other two nonces are computed as $\mathbf{r}_1 \longleftarrow \sigma(\mathbf{r}_3)$ and $\mathbf{r}_2 \longleftarrow \sigma(\mathbf{u}) \,\&\, \mathbf{r}_3$. Besides, given that COM is statistically-hiding, these choices can not be distinguished from what would have been obtained, had $\mathbf{r}_1$ and $\mathbf{r}_2$ been directly picked at random from $\{0,1\}^n$.

3.2. Security and Performance Considerations

The application of ideal lattices in this identification scheme can lead to improved and reduced memory footprint for public data. The usual restrictions regarding the choice of irreducible polynomials in the characterization of the ideal lattice, as well as its expansion factor must be observed, as discussed in [Lyubashevsky and Micciancio 2006]. This helps to ensure that finding short vectors in this kind of lattice remains hard to achieve.

Our scheme is also secure against concurrent attacks. This is a consequence of the fact that a public key corresponds to multiple secret keys, when the parameters are chosen in such a way that the pre-image set given by the private keys is much larger than the image set to which the public keys belong. The keys are related via mapping through the matrix $\mathbf{A}$.

3.3. Communication Costs

This identification scheme can benefit from the use of ideal lattices, by performing faster matrix vector multiplications with FFTs, similarly to what was done with SWIFFT hash function [Lyubashevsky et al. 2008]. The associated polynomial must be irreducible and with small expansion factor, as suggested in [Lyubashevsky and Micciancio 2006], in order to avoid known attacks.

Another efficient identification scheme, named CLRS, based on SIS was recently proposed by Cayrel *et al.*[Cayrel et al. 2010]. It possesses a soundness error of $(q+1)/2q$ per round, which is slightly higher than ours, namely 1/2. Such small difference plays an important role in terms of performance as depicted in Table 1. We provide below a comparison in terms of communication costs per round.

Our identification scheme exhibits the following message exchanges between the prover P and the verifier V. We are assuming that the seeds used to generate random elements in the protocol are 128 bits wide, and that the commitment function COMprovides results that are 256 bits wide.

- Commitments: 768 bits, corresponding to three commitment vectors.
- First challenge: $2\log_2 k$ bits
- Second challenge: 1 bit
- Answer to the second challenge: $(m/2)\log_2 q + m/2 + 256$, which is an average between
  - case challenge is 0:
    - one permutation seed: 128 bits
    - one vector in $\mathbb{Z}_q^m$: $m\log_2 q$ bits
    - one seed for the nonce $\mathbf{r}_3$: 128 bits
  - case challenge is 1:
    - one seed for reconstructing the vector $\mathbf{u}^* \in \mathbb{Z}_q^m$: 128 bits
    - one vector in $\{0,1\}^m$: m bits
    - one seed for the nonce $\mathbf{r}_3$: 128 bits

Thus, the total communication costs in bits per round of the protocol can be expressed as

$$\frac{m}{2}\log_2 q + 2\log_2 k + \frac{m}{2} + 1025.$$

Making the substitution of the parameters used in a concrete implementation, we have the 11275 bits per round. The overall cost for a scenario demanding soundness error of 2^{-16} is listed in Table 1. From there, one can see that our scheme improves the state of art (CLRS) in approximately 38%.

4. Attacks on the security assumptions

An attacker might try to break our scheme either by finding collisions in the COM function at will or by computing solutions to the SIS instances corresponding to the scheme. Given that the COM function adopted is also based on SIS [Kawachi et al. 2008], this implies the ability of successfully finding solutions for SIS defined over $\mathbb{Z}_q$, with vectors of dimension m and approximation factor $\sqrt{m}$. Breaking our instances of inhomogeneous SIS, implies the capacity to find vectors with max-norm 1 in arbitrarily chosen lattices. Neither of these attacks can be efficiently implemented with tools and techniques currently available. This does not change with the application of ideal lattices either.

In similarity with CLRS, we choose the parameters such that with overwhelming probability that there are other solutions to $\mathbf{Ax} = \mathbf{y} \bmod q$ besides the private key possessed by the Prover. This fact is of great importance for obtaining security against concurrent attacks. In order to reach this goal, q and m are chosen in such a way that the cardinality of set of the private keys is much bigger than the cardinality of set of the

public keys. This ensures that the system has the property of witness indistinguishability, which is kept under parallel composition.

As an instantiation with 100 bits of security, we suggest the parameters listed in Table 2, which are comparable to those used by the SWIFFT hash function and the CLRS identification scheme. The best combinatorial attack for finding short lattice vectors to this date, devised by Wagner [Wagner 2002], has a computational complexity above 2^{100} for such set of parameters. In addition to that, our private keys have norm below of what the best lattice reduction algorithms can obtain.

We chose the parameter k as 24 so that the soundness error per round be approximately the same as that of CLRS. Naturally, the Verifier has to load the set of public keys before the execution of the protocol from the trusted party that generates the keys. This represents an extra cost when compared to CLRS, but such operation can be executed at setup time. We are primarily concerned with the payload exchanged between the Prover and the Verifier. This can help in preserving bandwidth and battery time, which is important for constrained devices.

k	n	m	q	COM Length (bits)
24	64	2048	257	256

Table 2. Parameters for Lattice Instantiation

5. Conclusion and Further Work

We have shown in this paper a construction of a lattice-based identification scheme with worst-case connections, taking as starting point a code-based scheme. Its small soundness error results in lower communication costs than those from other lattice-based constructions over the I-SIS or SIS problems. Further improvements in computational costs can be obtained with the application of ideal lattices, without weakening the security properties.

As future work, we suggest an extension of this approach of replacing security assumptions used by cryptographic schemes to other hard problems in lattices, such as LWE (learning with errors) which has a formulation very close to that of error-correcting codes.

References

Aguilar, C., Gaborit, P., and Shreck, J. (2011). A new zero-knowledge code-based identification and signature scheme with reduced communications. *preprint*.

Ajtai, M. (1996). Generating hard instances of lattice problems. *Electronic Colloquium on Computational Complexity (ECCC)*, 3(7).

Ajtai, M. and Dwork, C. (1996). A public-key cryptosystem with worst-case/average-case equivalence. *Electronic Colloquium on Computational Complexity (ECCC)*, 3(65).

Cayrel, P.-L., Lindner, R., Rückert, M., and Silva, R. (2010). Improved zero-knowledge identification with lattices. In *Provable Security*, volume 6402 of *Lecture Notes in Computer Science*, pages 1–17. Springer.

Fiat, A. and Shamir, A. (1986). How to prove yourself: Practical solutions to identification and signature problems. In Odlyzko, A. M., editor, *CRYPTO*, volume 263 of *Lecture Notes in Computer Science*, pages 186–194. Springer.

Gaborit, P. and Girault, M. (2007). Lightweight code-based identification and signature. *IEEE Transactions on Information Theory (ISIT)*, pages 186–194.

Girault, M. (1990). A (non-practical) three-pass identification protocol using coding theory. In Seberry, J. and Pieprzyk, J., editors, *AUSCRYPT*, volume 453 of *Lecture Notes in Computer Science*, pages 265–272. Springer.

Goldwasser, S., Micali, S., and Rackoff, C. (1985). The knowledge complexity of interactive proof-systems. In *Proceedings of the seventeenth annual ACM symposium on Theory of computing*, page 304. ACM.

Harari, S. (1988). A new authentication algorithm. In Cohen, G. D. and Wolfmann, J., editors, *Coding Theory and Applications*, volume 388 of *Lecture Notes in Computer Science*, pages 91–105. Springer.

Kawachi, A., Tanaka, K., and Xagawa, K. (2008). Concurrently secure identification schemes based on the worst-case hardness of lattice problems. In *ASIACRYPT '08: Proceedings of the 14th International Conference on the Theory and Application of Cryptology and Information Security*, pages 372–389, Berlin, Heidelberg. Springer-Verlag.

Lyubashevsky, V. (2008). Lattice-based identification schemes secure under active attacks. In Cramer, R., editor, *Public Key Cryptography*, volume 4939 of *Lecture Notes in Computer Science*, pages 162–179. Springer.

Lyubashevsky, V. and Micciancio, D. (2006). Generalized compact knapsacks are collision resistant. In Bugliesi, M., Preneel, B., Sassone, V., and Wegener, I., editors, *ICALP (2)*, volume 4052 of *Lecture Notes in Computer Science*, pages 144–155. Springer.

Lyubashevsky, V., Micciancio, D., Peikert, C., and Rosen, A. (2008). Swifft: A modest proposal for fft hashing. In Nyberg, K., editor, *FSE*, volume 5086 of *Lecture Notes in Computer Science*, pages 54–72. Springer.

Micciancio, D. (2007). Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. In *Computational Complexity*. Springer.

Micciancio, D. and Goldwasser, S. (2002). *Complexity of Lattice Problems: a cryptographic perspective*, volume 671 of *The Kluwer International Series in Engineering and Computer Science*. Kluwer Academic Publishers, Boston, Massachusetts.

Shor, P. W. (1994). Polynominal time algorithms for discrete logarithms and factoring on a quantum computer. In Adleman, L. M. and Huang, M.-D. A., editors, *ANTS*, volume 877 of *Lecture Notes in Computer Science*, page 289. Springer.

Stern, J. (1993). A new identification scheme based on syndrome decoding. In Stinson, D. R., editor, *CRYPTO*, volume 773 of *Lecture Notes in Computer Science*, pages 13–21. Springer.

Véron, P. (1995). Cryptanalysis of harari's identification scheme. In Boyd, C., editor, *IMA Conf.*, volume 1025 of *Lecture Notes in Computer Science*, pages 264–269. Springer.

Véron, P. (1996). Improved identification schemes based on error-correcting codes. *Appl. Algebra Eng. Commun. Comput.*, 8(1):57–69.

Wagner, D. (2002). A generalized birthday problem. In Yung, M., editor, *CRYPTO*, volume 2442 of *Lecture Notes in Computer Science*, pages 288–303. Springer.