Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Universität Hamburg
Fachbereich Mathematik

Bundesstraße 55
20146 Hamburg

DIPLOMARBEIT

Identification Schemes from Restricted Collision-Resistant Linear Hash Functions

Identifikationsverfahren aus eingeschränkt kollisionsresistenten linearen Streufunktionen

Sven Moritz Hallberg

SS 2015

Betreut durch Dr. Hubert Kiechle

Ich versichere an Eides Statt, dass ich die Arbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe.

.....

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit der Konstruktion kryptographischer Identifikationsverfahren aus Streufunktionen einer bestimmten Klasse. Sie bezieht sich auf Arbeiten von Lyubashevsky und Micciancio zu gitterbasierten Verfahren, um allgemeine Muster herauszuarbeiten und Details zu erläutern. Ein Begriff "eingeschränkter" Kollisionsresistenz wird definiert, um die zentrale Problemstellung in Lyubashevskys Arbeit im allgemeinen Fall zu erfassen; nämlich, dass die gegebene Streufunktion nur eingeschränkt auf eine Teilmenge ihres Definitionsbereichs kollisionsresistent wird. Das Ergebnis ist eine Konstruktion kanonischer Identifikationsverfahren im Rahmen von Ringen und Moduln. Diese wird motiviert durch eine entsprechende abstrakte Bearbeitung von Lyubashevskys verwandtem Einmal-Signaturverfahren.

Die resultierenden Identifikationsverfahren können mit gewisser Wahrscheinlichkeit falsch-negative Ergebnisse liefern. Parallele Varianten, die diesem begegnen, werden diskutiert und Beweise dafür angegeben, dass ihre Sicherheit erhalten bleibt.

Schließlich folgt eine kurze Einführung zu Sicherheit auf Basis von Komplexitätsannahmen in Idealgittern und Micciancios Gitter-Streufunktion wird vorgestellt. Es wird gezeigt, wie die allgmeinen Konstruktionen instanziiert mit dieser Funktion Lyubashevskys ursprüngliche Verfahren ergeben. Abschätzungen für zulässige Werte der beteiligten Parameter werden hergeleitet und besprochen.

Abstract

This work is concerned with the construction of cryptographic identification schemes from a certain class of hash functions. It revisits prior work of Lyubashevsky and Micciancio to extract general patterns and clarify details. A notion of "restricted" collision resistance is defined to capture in the general case the central problem in Lyubashevsky's work; namely that the given hash function only becomes collision-resistant when restricted to a subset of its domain. The result is a construction of canonical identification schemes in the setting of rings and modules. It is motivated by a corresponding abstract treatment of Lyubashevsky's related one-time signature scheme.

The resulting identification schemes may yield false negatives with some probability. Parallelized variants that mitigate this are discussed and proofs provided that their security remains intact.

Finally, a short introduction to security based on hardness assumptions in ideal lattices is given and Micciancio's lattice hash function is presented. It is shown how the general constructions yield Lyubashevsky's original schemes when instantiated with this function. Bounds for valid values of the associated parameters are derived and discussed.

Contents

1 Introduction

Identification schemes are cryptographic protocols that allow a user to prove their identity without having to trust the verifying party. They are part of the wider class of *interactive proof systems*. Uses include electronic passports, payment cards, and military friend-or-foe systems. These uses are to be distinguished from systems such as simple doorlock keycards where verification is always performed by the issuing party and presenting a shared secret suffices. By contrast, a passport must not become forgeable by any checkpoint it passes through. Fiat and Shamir therefore introduce identification schemes in the following hierarchy with authentication and signatures [FS87]:

> An *authentication scheme* allows Alice and only Alice to prove her identity to Bob.
>
> An *identification scheme* is an authentication scheme where in addition to the above, Bob cannot impersonate Alice to someone else.
>
> In a *signature scheme*, Bob cannot even fake a successful interaction with Alice to himself. The security condition in this case is usually referred to as *unforgeability*.

Like the security notions build on each other, complex cryptographic constructions regularly derive from more basic components. Hash functions play a fundamental role. While they are often used in practice for their *compression* property, mapping a large domain to a limited range, many types of cryptographic schemes can be shown to be constructible from just hash functions with suitable properties. Typically required are *one-way-ness* and *collision resistance*. These are *hardness assumptions* from which security properties are derived by reduction: A hypothetical attack on the system is used to solve a problem such as inverting a function or finding collisions.

There are many cases where hardness is only conjectured. The RSA problem is a prominent example; it appears to be as hard as factoring but no proof has been dicovered. In fact, factoring itself is only believed to be hard because no efficient algorithm is known. Classic encryption algorithms as well as hash functions such as SHA have historically been proposed without security proofs, confidence in them based on failure to find evidence to the contrary. Basing schemes directly on reductions to hard problems is therefore an area of active research, sometimes referred to in a slight stretch of language as *provable security*.

The identification scheme we are interested in was developed by Lyubashevsky in [Lyu08a] and later refined in [Lyu08b]. It takes the form of a canonical *challenge-response* protocol where the prover "signs" each random challenge using a *one-time* signature scheme. The construction hides secret keys behind a linear hash function whose collision resistance reduces to a form of the *shortest vector problem* in lattices.

A lattice is a subgroup of $\mathbb{R}^n$ consisting of integer combinations of some basis. The study of lattices is connected with coding theory and both fields are the subject of current research into *post-quantum* cryptographic schemes that resist attacks on a (hypothetical) quantum computer. Micciancio gives a survey of lattice-based cryptography in [NV09]. See [CS99] by Conway and Sloane for an introduction to lattices and coding theory.

The lattice-based schemes presented here involve several interdependent parameters and Lyubashevsky's proofs contain some technicalities as a result. The main goal of this work is to split the constructions into an abstract and a lattice-specific part and to improve clarity by expanding the presentation. Some bounds for valid parameters are derived more explicitly than in [Lyu08b].

Chapters 3 and 4 develop the abstract one-time signature and identification schemes assuming a linear hash function over some module and appropriate properties. Chapter 5 introduces lattices and the concrete hash function before chapters 6 and 7 apply it in the general constructions.

2 Preliminaries

This chapter gives a brief summary of the notions of signature and identification schemes, providing the basis and starting point for the main topic. It also serves to establish notational conventions used throughout the rest of this work.

- Module elements will be denoted by bold letters, e.g. $\mathbf{x}$, $\mathbf{y}$.
- Algorithms and Turing machines are set in a typewriter font, e.g. $\mathtt{A}$, $\mathtt{V}$, and are generally considered probabilistic polynomial-time machines. Subscripts are used to denote auxiliary inputs, e.g. $\mathtt{S}_k$.
- The output of an *interactive* Turing machine $\mathtt{B}$ after interaction with $\mathtt{A}$ on common input x is denoted by $\langle \mathtt{A}, \mathtt{B} \rangle(x)$.
- Variable assignment is denoted as $x \leftarrow \mathtt{A}$.
- Throughout this work, n refers to the primary security parameter of the scheme under discussion, which is formally passed as 1^n to key generation algorithms. We will omit it from the presentation for convenience, considering it an implicit parameter of any algorithm.
- Choosing an element x uniformly at random from a set D is denoted by $x \xleftarrow{\$} D$.
- Probabilities are denoted as $\mathrm{P}(\dots)$. The probability of an event A under a condition B is denoted by $\mathrm{P}(A \mid B)$.
- Landau "Big-O" notation is used to denote asymptotic bounds: If for any constant $C > 0$ there exists an N such that $f(n) \leqslant C \cdot g(n)$ for all $n > N$, write

  $$f(n) = O(g(n)) \quad .$$

  Also, "soft-O" notation will be used to summarize results without regard for polylogarithmic factors.

  $$f(n) = \tilde{O}(g(n)) \;:\Leftrightarrow\; \exists k.\, f(n) = O(g(n) \cdot \log^k n)$$

Definition 2.1. A function $f(n)$ is called *negligible* if it diminishes super-polynomially, i.e. for any polynomial p, there exists an N such that $f(n) \leqslant 1/p(n)$ for all $n > N$. We will write simply

$$f(n) = O(n^{-k}) \quad .$$

When referring to probabilities, the complementary notion

$$f(n) = 1 - O(n^{-k})$$

will be called *overwhelming*.

We now define (length-restricted) signature schemes as well as identification schemes and their security following the presentation in Goldreich [Gol01, Gol04].

Definition 2.2. A *signature scheme* is a triple $(\mathtt{G}, \mathtt{S}, \mathtt{V})$ of probabilistic polynomial-time algorithms where:

1. $\mathtt{G}$ is the *key generator*; it outputs a pair of keys (k, K) which are used for *signing* and *verification* by $\mathtt{S}$ and $\mathtt{V}$, respectively.

2. *(Soundness)* All signatures produced by $\mathtt{S}_k$ are accepted by $\mathtt{V}_K$.

3. *(Length restriction)* The set of valid inputs to $\mathtt{S}$ is finite.

Note. The above definition covers both private and public key schemes which differ only in their security notions. In a private key scheme, k will generally equal K. That said, we will only concern ourselves with the public key case.

Definition 2.3. A public-key signature scheme $(\mathtt{G}, \mathtt{S}, \mathtt{V})$ is called *secure* or *unforgeable* if for every polynomial-time adversary $\mathtt{A}$

$$\mathrm{P}\left(\mathtt{V}_K(c,z) \wedge c \notin \mathrm{Q}_{\mathtt{A}}^{\mathtt{S}_k}(K) \text{ where } (k,K) \leftarrow \mathtt{G},\ (c,z) \leftarrow \mathtt{A}(K)\right) = O(n^{-k})$$

where $\mathrm{Q}_{\mathtt{A}}^{\mathtt{S}_k}(K)$ refers to the set of queries made by $\mathtt{A}$ on input K to an oracle for $\mathtt{S}_k$.

Definition 2.4. A public-key signature scheme is *strongly unforgeable* if in addition to definition 2.3, given a valid signature z for some document, an adversary has negligible chance to produce a second valid signature $z' \neq z$ for the same document.

Definition 2.5. An *identification scheme* consists of an algorithm $\mathtt{I}$ and a pair of interactive Turing machines $(\mathtt{P}, \mathtt{V})$ where

1. $\mathtt{I}$ generates key pairs (α, a) to serve as identifying information.

2. $\mathtt{P}$ and $\mathtt{V}$ implement the protocol between *prover* and *verifier*. A public key a is passed as common input to the protocol while $\mathtt{P}$ receives the private key α as auxiliary input.

3. *(Soundness)* If (α, a) is a key pair generated by $\mathtt{I}$, the interaction of $\mathtt{P}$ and $\mathtt{V}$ yields success with overwhelming probability.

$$\mathrm{P}\left(\langle \mathtt{P}_\alpha, \mathtt{V}\rangle(a) = 1\right) \;=\; 1 - O(n^{-k})$$

Definition 2.6. An identification scheme $(\mathtt{I}, \mathtt{P}, \mathtt{V})$ is *secure* (under active attack) if the success probability of any polynomial-time adversary $\mathtt{A}$ is negligible even after $\mathtt{A}$ has interacted with $\mathtt{P}_\alpha$ a polynomial number of times; let T denote, as a random variable, the result of these interactions.

$$\mathrm{P}\left(\langle \mathtt{A}_T, \mathtt{V} \rangle(a) = 1\right) \;=\; O(n^{-k})$$

A security property of general interest with protocols such as identification schemes is *witness-indistinguishability*, first introduced by Feige and Shamir [FS90] as an alternative to *zero-knowledge* that is preserved under parallel composition. The term "witness" refers to proof systems characterized by a witness relation; i.e. for each instance there exists an element that enables its solution. The scheme(s) presented in this work actually satisfy the variant of being fully *witness-independent* [Gol01] (which is still preserved under parallel composition).

Note that there could be multiple witnesses for the same instance. In the context of an identification scheme, there will in general be potential private keys beside the generated one that allow the prover to succeed with respect to the same public key.

Definition 2.7. Call an identification scheme *witness-independent* if, for any Turing machine $\mathtt{V}^*$ and two private keys α_1, α_2 that fit public key a, the results of $\langle \mathtt{P}_{\alpha_1}, \mathtt{V}^* \rangle(a)$ and $\langle \mathtt{P}_{\alpha_2}, \mathtt{V}^* \rangle(a)$ are identically distributed.

3 Hash Function One-Time Signatures

This chapter introduces the concept of *one-time* signatures before a common construction for (public-key) one-time signatures is motivated and defined in a general setting.

The notion of one-time signatures is originally due to Lamport [Lam79, DH76]. In it a different key is used for every document signed. Formally, a forging attacker is allowed access to only one valid signature [Gol04].

Apart from the relaxed security requirement, a one-time signature scheme takes the same form as an ordinary signature scheme: If $\mathtt{G}$ is used to generate signing key k and verification key K, $\mathtt{S}_k(c)$ yields a signature z for the document c. $\mathtt{V}_K(c, z)$ yields 1 if z is a valid signature for c under the key k.

Definition 3.1. A signature scheme $(\mathtt{G}, \mathtt{S}, \mathtt{V})$ is secure as a (public-key) *one-time signature scheme* if and only if it is secure against attackers that make at most one query to the signing oracle.

3.1 Motivation

Like ordinary signature schemes, one-time signatures can be defined in a private or public-key setting. The private-key formulation where $k = K$ will be useful as a step in motivating the public-key scheme defined in section 3.2. The verifier, knowing the secret key, could compute the correct signature himself and perform verification by comparison. The signing algorithm can be considered a collection of functions parameterized over the keyspace. Key generation means picking a function from the collection at random. Should the key be an element of an algebraic structure like a group, the most obvious way of defining such a collection of functions is by the structure's operation.

$$f_x(a) := xa$$

This would yield a valid signature scheme according to definition 2.2 but it would generally be insecure. If the operation is invertible, the secret parameter can be computed from a and $f_x(a)$. The idea employed below is to chain two such functions, making $f_x(a)$ an intermediate value that is unknown to the attacker.

$$(g_y \circ f_x)(a) = g_y(f_x(a))$$

Of course, the overall security of the system so constructed remains to be proven. In particular, the two functions must not be algebraically "related", i.e. there must not be a

way of collapsing them to avoid compution of the intermediate. In the setting of a ring or module, two operations are readily available that might serve this purpose. E.g.:

$$\begin{aligned} f_x(a) &:= xa \\ g_y(b) &:= b + y \\ S_{(x,y)}(c) &:= xc + y \end{aligned}$$

The key consists of the two elements x and y. This "split key" structure (cf. section 3.3) will be convenient in the derivation of an identification scheme in chapter 4.

To turn the above concept into one for a public key scheme, the idea is to hide the private key values x and y behind a one-way function h. The verifier is given only $h(x)$ and $h(y)$. If h is a (ring or module) homomorphism, it allows the verification to be moved "onto the other side" of h. The simple comparison condition

$$z = xc + y$$

of the private key setting is replaced by

$$\begin{aligned} h(z) &= h(xc + y) \\ &= h(x)c + h(y) \quad . \end{aligned} \tag{$*$}$$

From equation $(*)$ it is already obvious that in order to avoid forgeries, h must be collision-resistant. The following section will concern itself with formally defining a scheme as motivated above and identifying conditions under which it remains secure.

3.2 One-Time Signatures from Linear RCRHF

A major concern of this work is how to construct its schemes when the underlying hash function can only be proven collision-resistant on a *subset* of its actual domain. It is important to clarify that notion.

Definition 3.2 (restricted collision resistance)**.** A function $h : A \to B$ is called *(restricted) collision-resistant* on a subset $D \subset A$ if it is computationally infeasible to find $x, y \in D$ such that

$$x \neq y \ \wedge \ h(x) = h(y) \quad .$$

The computational problem of finding a collision in h with elements from D will be refered to as $\mathrm{Col}(h, D)$.

Note that the above definition pertains to both x and y lying in D. In particular, no statement is made about the chance to produce, given one preimage $x \in D$, a second preimage y that is allowed to be taken from the whole of A. It will be important to exclude this possibility in the signature scheme below.

Scheme 3.1. Let R be a ring, $D_c \subset R$ a subset representing documents, M an R-module, $h : M \to R$ an R-module-homomorphism, and $D \subset M$ a subset representing valid signatures. Let $\mathtt{G}'$ be a probabilistic polynomial-time algorithm that outputs an element $(\mathbf{x}, \mathbf{y})$ of $M \times M$. Define the following signature scheme.

$$\mathtt{G} := \big((\mathbf{x}, \mathbf{y}),\ (h(\mathbf{x}), h(\mathbf{y})) \big) \qquad , \ (\mathbf{x}, \mathbf{y}) \leftarrow \mathtt{G}'$$

$$S_{(\mathbf{x},\mathbf{y})}(c) := \begin{cases} \mathbf{x}c + \mathbf{y} & \text{if } \mathbf{x}c + \mathbf{y} \in D \\ \perp & \text{otherwise} \end{cases} \qquad , \ c \in D_c$$

$$V_{(X,Y)}(c, \mathbf{z}) := \begin{cases} 1 & \text{if } \mathbf{z} \in D \ \wedge \ h(\mathbf{z}) = Xc + Y \\ 0 & \text{otherwise} \end{cases}$$

Where convenient, $\mathbf{x}$ and $\mathbf{y}$ as given above will be called the *primary* and *secondary* (signing) keys. Analogously, $h(\mathbf{x})$ and $h(\mathbf{y})$ are called primary and secondary verification keys.

Note that in order to accomodate a restricted collision resistance of h, the definition permits a partial signing function S. Not every document must admit a signature under a given key and verification only succeeds if the signature lies in D. Thus, as will be shown below, a successful attacker is forced to surmount collision resistance of h on D.

Definition 3.3 (completeness)**.** A signature scheme is called *complete* if the result of $S_k(c)$ is defined (not $\perp$) for all $c \in D_c$ and private keys k.

It is worth noting that contrary to what one might assume, it is plausible that completeness is not strictly required for one-time signature schemes. Because new keys must be arranged for each message, it can be expected to be possible to switch to a new key for the same message if a signature appears undefined. This is exactly what will be done to make the identification scheme of chapter 4 (statistically) complete.

Theorem 3.1 (soundness)**.** *Every signature successfully created (not $\perp$) by scheme 3.1 is accepted under the corresponding public key.*

Proof. Let $(\mathbf{x}, \mathbf{y})$ denote the private key as above. Let $X = h(\mathbf{x})$, $Y = h(\mathbf{y})$ be the corresponding public key. The linearity of h ensures that, for any signature $\mathbf{z} = \mathtt{S}_{(\mathbf{x},\mathbf{y})}(c)$, the defining condition of $\mathtt{V}_{(X,Y)}(c)$ is satisfied.

$$\begin{aligned} h(\mathbf{z}) &= h(\mathtt{S}_{(\mathbf{x},\mathbf{y})}(c)) = h(\mathbf{x}c + \mathbf{y}) = h(\mathbf{x})c + h(\mathbf{y}) \\ &= Xc + Y \end{aligned}$$

□

Theorem 3.2 (security)**.** *If the following conditions hold, a successful attack against scheme 3.1 is at least as hard as $Col(h, D)$.*

1. *For a primary key $\mathbf{x}$ as generated by $\mathtt{G}'$ and given documents $c_1 \neq c_2$, the expression $\mathbf{x}(c_1 - c_2)$ uniquely determines the value of $\mathbf{x}$.*

   NB. *This condition is trivially satisfied if R is an integral domain.*

2. *The probability that the key generator produces a particular key* $(\mathbf{x}, \mathbf{y})$ *is negligible, even under knowledge of the associated public key and a signature for some document* $c \in D_c$*:*

$$\mathrm{P}\left((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y}) \mid h(\tilde{\mathbf{x}}) = X,\, h(\tilde{\mathbf{y}}) = Y,\, \tilde{\mathbf{x}}c + \tilde{\mathbf{y}} = \mathbf{z}\right) = O(n^{-k})$$

where the probability is taken over the coin tosses of $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \leftarrow \mathtt{G}'$ *and where* $X = h(\mathbf{x})$*,* $Y = h(\mathbf{y})$*, and* $\mathbf{z} = \mathbf{x}c + \mathbf{y}$*.*

Proof. An attack on the scheme takes the following form. Let $(\mathbf{x}, \mathbf{y})$ be the generated secret key. The attacker is allowed one signature query. It submits a document c_1 and is granted

$$\mathbf{z}_1 := \mathbf{x}c_1 + \mathbf{y} \quad .$$

It must now produce c_2 and $\mathbf{z}_2 \neq \mathbf{z}_1$ such that

$$h(\mathbf{z}_2) = Xc_2 + Y \quad .$$

Consider the following two cases.

$\underline{\mathbf{z}_2 = \mathbf{x}c_2 + \mathbf{y}}$: The attacker has guessed the real signature. It will be shown that this can only happen with negligible probability. Note that in this case, $\mathbf{z}_2 \neq \mathbf{z}_1$ implies $c_2 \neq c_1$. Consider

$$\mathbf{z}_1 - \mathbf{z}_2 = \mathbf{x}(c_1 - c_2) \quad .$$

By assumption 1, $\mathbf{x}$ and in turn $\mathbf{y}$ are uniquely determined by the above equation. Therefore the probability of this case is at most the probability of $\mathtt{G}'$ generating that key, under any conditions implied by the attacker's knowledge. This conditional probability is negligible by assumption 2.

$\underline{\mathbf{z}_2 \neq \mathbf{x}c_2 + \mathbf{y}}$: The attacker has found a fake signature that is accepted as valid. This yields a collision in h:

$$h(\mathbf{z}_2) = h(\mathbf{x})c_2 + h(\mathbf{y}) = h(\mathbf{x}c_2 + \mathbf{y})$$

Thus, a successful attack on $(\mathtt{G}, \mathtt{S}, \mathtt{V})$ is at least as hard as finding a collision in h. □

Corollary 3.3. *If the conditions of theorem 3.2 hold and h is collision-resistant on D, definition 3.1 describes an unforgeable one-time signature scheme.* □

Theorem 3.4. *If the one-time signature scheme 3.1 is unforgeable, it is also strongly unforgeable.*

Proof. Let $\mathbf{z}_1$ and $\mathbf{z}_2$ be two different valid signatures for the same document c. It follows directly that both hash to the same value.

$$h(\mathbf{z}_1) = h(\mathbf{z}_2) = Xc + Y$$ □

3.3 Split-Key Signatures

The keys of the one-time signature scheme developed in this chapter naturally consist of two parts. This structure will lend itself to the construction of the identification scheme of chapter 4. Split-key signatures represent a step up from one-time signatures where one part of the key is re-used long-term as in an ordinary signature scheme.

Definition 3.4. Consider a one-time signature scheme $(\mathtt{G}, \mathtt{S}, \mathtt{V})$. Call it a *split-key signature scheme* if there exist independent algorithms $\mathtt{G}_1$ and $\mathtt{G}_2$ such that $\mathtt{G}$ yields a key of the form $((x, y), (X, Y))$ where

$$(x, X) \leftarrow \mathtt{G}_1 \quad \text{and}$$
$$(y, Y) \leftarrow \mathtt{G}_2 \quad .$$

By convention (x, X) is considered a *long-term key* whereas (y, Y) will be assumed one-time.

Shoup and Belare [BS07] introduce a similar concept called *two-tier* signatures. They are able to give a general construction of two-tier signatures from any identification scheme, whereas we intend to go in the other direction. It is of note that their definition allows $\mathtt{G}_2$ to depend on the result of $\mathtt{G}_1$ which is not suitable for our case: Our proof of the witness-independence of the identification scheme will depend on the verification and signing keys being chosen independently.

4 Hash Function Identification Schemes

This chapter draws on the previous to build an identification scheme based on a linear hash function. Section 4.1 motivates the construction from split-key signatures. Section 4.2 defines and proves secure the basic construction. In addition, witness-independence of the scheme is proved. Section 4.3 provides parallel variants to mitigate false negatives that may arise with a restricted collision-resistant hash function.

4.1 Identification Schemes and One−time Signatures

Many identification schemes follow the same *canonical* [AABN02] three-move pattern shown in figure 4.1. It consists of the algorithms `KeyGen`, `Commit`, `Respond`, and `Verify`. In an initial step, `KeyGen` is used to generate the prover's private key x and the corresponding public key X which is communicated to the verifier. Note that this step is only performed once and thus not part of the protocol proper. It is included in the diagram mainly to introduce x and X. In fact, the key generation need not be performed by the prover but could be delegated to a trusted third party. It is assumed that the verifier has correctly received X.

In the first step of the protocol, the prover computes the *commitment* Y which it sends to the verifier. The commitment is so called because it generally corresponds to a secret y and revealing Y "commits" the prover to using this particular y in later steps. As indicated in the figure, the commitment may theoretically depend on the previously-generated key, but it usually does not. Usually y is randomly chosen and Y derived from it by means of a one-way function.

In the second step, the verifier sends a random challenge c. Here, the set of choices has been denoted D_c to signify that it may depend on the scheme. In general one could think of a random bit-string, but it will be convenient in the case of section 7 to directly use (a subset of) the ring over which the scheme operates.

The prover computes his *response* to the challenge using his knowledge of the secrets x and y. Finally, the verifier uses his knowledge of X and Y to compute the *decision bit* d and accepts if and only if it is 1.

The form of canonical three-move identification schemes arises naturally with schemes constructed from (split-key) one-time signatures. It appears noteworthy that relaxing the notion of *signature* in this way allows us to take a step down the hierarchy and build both identification schemes and proper signatures up from there.

To construct an identification scheme, recall that with split-key signatures, the algorithm `G` consist of two parts $\mathtt{G}_1$ and $\mathtt{G}_2$ such that if $\mathtt{G}_1$ yields (x, X) and $\mathtt{G}_2$ yields (y, Y), then $((x, y), (X, Y))$ forms the result of `G`. Looking at figure 4.1, it is easy to see how to

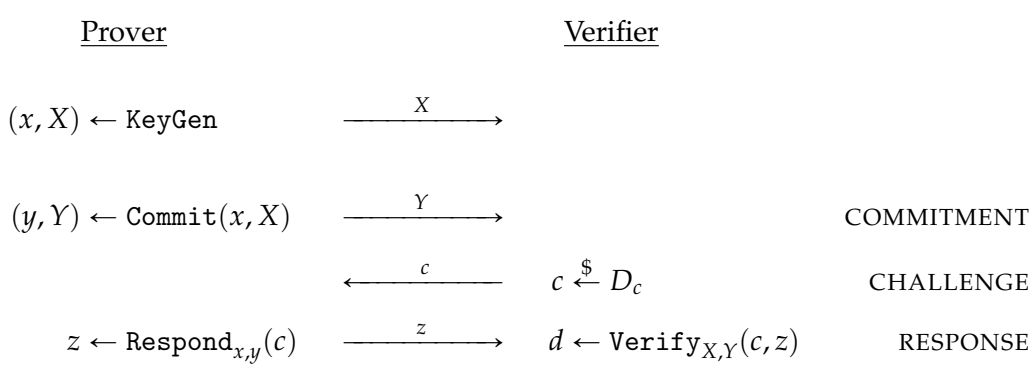

Figure 4.1: Canonical three-move identification scheme

use the algorithms of a split-key one-time signature to form an identification scheme: Respond by signing the challenge using a fresh secondary key that is generated during the commitment phase. It is important to note that this is a derivation in *Gestalt* only; the security proofs remain separate. Indeed, the instantiations of chapters 6 and 7 use different parameters to support their individual proofs of security.

4.2 Identification Schemes from Linear RCRHF

We will apply the construction described above to the hash function one-time signature of scheme 3.1.

Protocol 4.1. Consider scheme 3.1, consisting of algorithms $(\mathtt{G}, \mathtt{S}, \mathtt{V})$ such that $\mathtt{G}$ yields $((\mathbf{x}, \mathbf{y}), (h(\mathbf{x}), h(\mathbf{y})))$ and the private key $(\mathbf{x}, \mathbf{y})$ is generated by an algorithm $\mathtt{G}'$. If $\mathtt{G}'$ consists of independent algorithms $\mathtt{G}'_1$ and $\mathtt{G}'_2$ that produce $\mathbf{x}$ and $\mathbf{y}$, respectively, define an identification scheme

$$\begin{aligned} \texttt{KeyGen} &:= \mathtt{G}_1 & \texttt{Respond} &:= \mathtt{S} \\ \texttt{Commit} &:= \mathtt{G}_2 & \texttt{Verify} &:= \mathtt{V} \end{aligned}$$

where $\mathtt{G}_1$ uses $\mathtt{G}'_1$ to produce $(\mathbf{x}, h(\mathbf{x}))$ and $\mathtt{G}_2$ uses $\mathtt{G}'_2$ to produce $(\mathbf{y}, h(\mathbf{y}))$.

Figure 4.2 shows the protocol in detail. As alluded to in the description of split-key signatures, the first half of the key is used as the (long-term) key for the identification scheme. The second half is generated as part of the protocol and revealing its public (verification) part forms the commitment. With the keys thus exchanged, signing a random challenge serves as proof of identity. The signing function $\mathtt{S}$ is allowed to be partial, i.e. to yield a result of $\perp$ which causes verification to fail. This is interpreted as an abort of the protocol. The following definition states a condition that will serve us in determining the probability of aborts as well as in the security proofs.

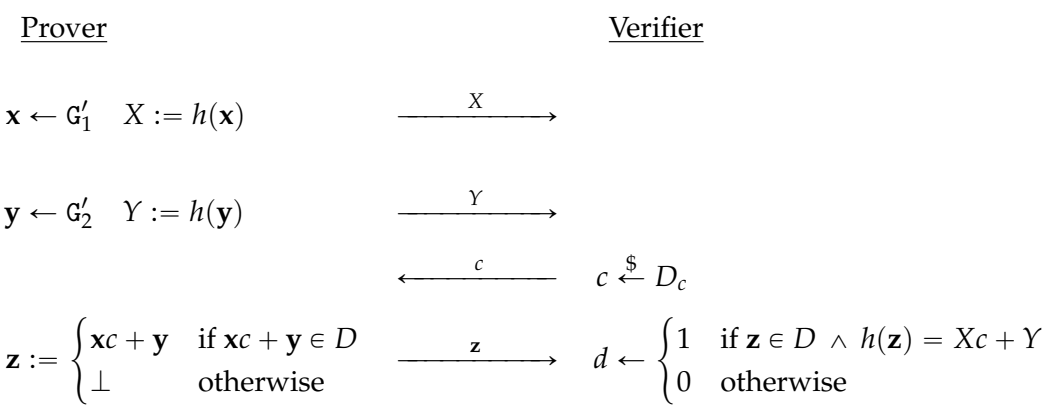

Figure 4.2: Hash function identification scheme

Definition 4.1 (targetability)**.** Call a result $\mathbf{z} \in D$ of protocol 4.1 *targetable* by witness $\mathbf{x}$ if $\mathbf{z} - \mathbf{x}c \in \mathrm{Rg}(\mathtt{G}_2')$ for all challenges $c \in D_c$. Call the identification scheme targetable if all results $\mathbf{z} \in D$ are targetable by all witnesses $\mathbf{x} \in \mathrm{Rg}(\mathtt{G}_1')$.

Theorem 4.1 (aborts)**.** *If protocol 4.1 is targetable and the output of $\mathtt{G}_2'$ is uniformly distributed, then the scheme aborts with probability*

$$1 - \frac{|D|}{|\mathrm{Rg}(\mathtt{G}_2')|} \quad .$$

Proof. Let $\mathbf{z}$ be the prover's response. The probability that $\mathbf{z}$ lies in D is determined by the number of $\mathbf{y} \in \mathrm{Rg}(\mathtt{G}_2')$ such that $\mathbf{x}c + \mathbf{y} \in D \ \Leftrightarrow \ \mathbf{y} \in D - \mathbf{x}c$. The set $D - \mathbf{x}c$ is a subset of $\mathrm{Rg}(\mathtt{G}_2')$ by the assumption of targetability and we have

$$|\mathrm{Rg}(\mathtt{G}_2') \ \cap \ (D - \mathbf{x}c)| = |D - \mathbf{x}c| = |D| \quad .$$

Thus, the probability for success is $|D| \,/\, |\mathrm{Rg}(\mathtt{G}_2')|$ as required. □

Note that the probability for an abort does not depend on the value of the private key, or witness, $\mathbf{x}$. In fact, we can show the protocol as a whole to be witness-independent.

Theorem 4.2 (witness-independence)**.** *If protocol 4.1 is targetable and the output of $\mathtt{G}_2'$ is uniformly distributed, then the scheme is witness-independent.*

Proof. Recall the definition of witness-independence. We must show that the distribution of the output of any machine in the role of the verifier does not depend on the value of $\mathbf{x}$, where $h(\mathbf{x}) = X$. For this, it suffices to show that any *inputs* received by such a machine during the protocol are independent of $\mathbf{x}$.

A verifier of the protocol learns Y and $\mathbf{z}$. The commitment Y is produced by the prover as a function of $\mathbf{y}$ which is chosen independently of $\mathbf{x}$.

To show that the distribution of the response $\mathbf{z}$ is independent of $\mathbf{x}$, recall that $\mathbf{z}$ can either lie in D or be $\perp$. Proceed by case analysis:

$\underline{\mathbf{z} \in D}$: For any challenge $c \in D_c$ and witness $\mathbf{x}$, there exists exactly one $\mathbf{y}$ such that $\mathbf{x}c + \mathbf{y} = \mathbf{z}$, namely $\mathbf{z} - \mathbf{x}c$. The latter is an element of $\mathrm{Rg}(\mathtt{G}'_2)$ by the assumption of targetability and hashes to Y:

$$h(\mathbf{z} - \mathbf{x}c) = h(z) - h(\mathbf{x})c = Xc + Y - Xc = Y$$

Therefore, it is a valid and the only choice for $\mathbf{y}$, so the probability of $\mathbf{z}$ resulting from a given $\mathbf{x}$ is always

$$\frac{1}{|h^{-1}(Y) \cap \mathrm{Rg}(\mathtt{G}'_2)|}$$

which is independent of $\mathbf{x}$.

$\underline{\mathbf{z} = \perp}$: The probability of this case is independent of $\mathbf{x}$ by theorem 4.1. □

Theorem 4.3 (security)**.** *If protocol 4.1 is targetable and the following conditions hold, a successful attack can be used to construct a solution to* $Col(h, \mathrm{Rg}(\mathtt{G}'_2))$*. Let* $\mathbf{x}$ *be any private key as generated by* $\mathtt{G}'_1$*.*

1. *Given two challenges* $c_1 \neq c_2$*, the value of* $\mathbf{x}(c_1 - c_2)$ *uniquely determines* $\mathbf{x}$*.*
2. *There exists with overwhelming probability another key* $\mathbf{x}' \neq \mathbf{x}$ *such that* $h(\mathbf{x}') = h(\mathbf{x})$*.*
3. *The key* $\mathbf{x}$ *is chosen uniformly at random by* $\mathtt{G}'_1$ *or the identification scheme is witness-independent.*

Proof. To construct a collision in h, generate $\mathbf{x}$ as required by the protocol and act as an honest prover to the adversary in the first step. In the second phase, acting as the honest verifier, note that the state of the adversary (a Turing machine) can be saved and the machine rerun from that state. Save the adversary's state after receiving Y and run the rest of the protocol twice, using two challenges $c_1 \neq c_2$.[1] Call the adversary's corresponding answers $\mathbf{z}_1$ and $\mathbf{z}_2$, respectively. If both answers are valid, i.e.

$$\begin{aligned} & h(\mathbf{z}_1) = Xc_1 + Y \\ \wedge\ & h(\mathbf{z}_2) = Xc_2 + Y \quad , \end{aligned}$$

we can derive

$$\begin{aligned} & h(\mathbf{z}_1) - Xc_1 = h(\mathbf{z}_2) - Xc_2 \\ \Rightarrow\ & h(\mathbf{z}_1 - \mathbf{x}c_1) = h(\mathbf{z}_2 - \mathbf{x}c_2) \quad . \end{aligned}$$

This means that we have a collision (in $\mathrm{Rg}(\mathtt{G}'_2)$ by virtue of targetability) if

$$\begin{aligned} & \mathbf{z}_1 - \mathbf{x}c_1 \ \neq\ \mathbf{z}_2 - \mathbf{x}c_2 \\ \Leftrightarrow\ & \mathbf{z}_1 - \mathbf{z}_2 \ \neq\ \mathbf{x}(c_1 - c_2) \quad . \end{aligned}$$

[1] Formally we are acting as the honest verifier and must pick the challenges uniformly at random. The chance to pick $c_1 = c_2$, however, can be made negligibly small even if $|D_c|$ is constant, by trying n times. The attack will still run in polynomial time.

When c_1, c_2, $\mathbf{z}_1$, and $\mathbf{z}_2$ are given, the above is a condition on $\mathbf{x}$ which we will denote as an event $\mathfrak{K}(\mathbf{x})$. The opposite

$$\mathbf{z}_1 - \mathbf{z}_2 = \mathbf{x}(c_1 - c_2)$$

holds for at most one $\mathbf{x}$ by assumption 1.

If $\mathbf{x}$ is chosen uniformly at random, the probability for $\mathfrak{K}(\mathbf{x})$ is proportional to the number of choices for $\mathbf{x}$ that satisfy the condition. By assumption 2 we can assume at least one other choice for $\mathbf{x}$ to exist and obtain

$$\mathrm{P}\left(\mathbf{z}_1 - \mathbf{z}_2 \neq \mathbf{x}(c_1 - c_2)\right) \geqslant \frac{1}{2} \quad ,$$

completing the proof in this case.

If we wish to avoid assumptions about the distribution of $\mathbf{x}$, we can base the proof on witness-independence. Let $\mathbf{x} \neq \mathbf{x}'$ be two witnesses with $h(\mathbf{x}) = h(\mathbf{x}') = X$ such as we may assume exist by assumption 2. As before we know that at most one of them can fail to satisfy $\mathfrak{K}$ and thus observe

$$\mathrm{P}\left(\mathfrak{K}(\mathbf{x}') \;\middle|\; \neg\mathfrak{K}(\mathbf{x})\right) = 1 \quad .$$

Let $p = \mathrm{P}\left(\mathfrak{K}(\mathbf{x})\right)$ and apply the law of total probability to obtain

$$\begin{aligned}
\mathrm{P}\left(\mathfrak{K}(\mathbf{x}')\right) &= \mathrm{P}\left(\mathfrak{K}(\mathbf{x}') \;\middle|\; \mathfrak{K}(\mathbf{x})\right) \cdot p \;+\; \mathrm{P}\left(\mathfrak{K}(\mathbf{x}') \;\middle|\; \neg\mathfrak{K}(\mathbf{x})\right) \cdot (1 - p) \\
&= \mathrm{P}\left(\mathfrak{K}(\mathbf{x}') \;\middle|\; \mathfrak{K}(\mathbf{x})\right) \cdot p \;+\; 1 - p \\
&\geqslant 1 - p \quad .
\end{aligned}$$

By witness-independence we have $\mathrm{P}\left(\mathfrak{K}(\mathbf{x})\right) = \mathrm{P}\left(\mathfrak{K}(\mathbf{x}')\right)$ so the above yields

$$p \geqslant 1 - p \;\Leftrightarrow\; p \geqslant \frac{1}{2} \quad ,$$

proving the theorem. □

Corollary 4.4. *If h is collision-resistant on $\mathrm{Rg}(\mathtt{G}'_2)$ and the conditions of theorem 4.3 hold, the identification scheme of definition 4.1 is secure in the active attack model.* □

4.3 Multiple Parallel Executions

Recall that if the one-time signing function $\mathtt{S}$ is partial, protocol 4.1 may produce a false negative result with some non-negligible probability as in theorem 4.1. Specifically, $\mathbf{z} \in D$ need not be satisfied which, were the protocol allowed to proceed, would break our reduction to finding a collision in h. The obvious approach to counter this is to retry until the scheme succeeds or the verifier runs out of patience. However, in order to save on the number of messages exchanged, it is desirable to perform multiple runs in parallel: succeed if any one run succeeds, fail otherwise. Figure 4.3 shows the modified scheme.

The following theorem gives an asymptotic measure for the number of rounds required to ensure success.

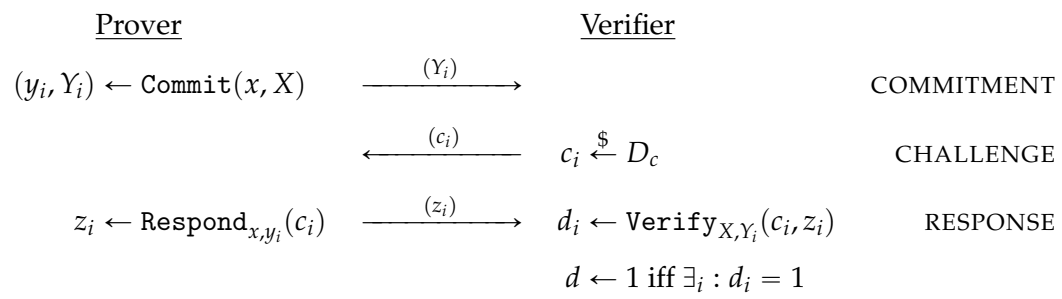

Figure 4.3: Multiple parallel rounds of a canonical identification scheme

Theorem 4.5 (aborts)**.** *If the probability $q(n)$ for failure of a single round is bounded by a constant $\delta < 1$, then a number of rounds $t(n) = \omega(\log n)$ makes the chance of overall failure negligible.*

Note. Following Knuth [Knu76], we use ω to denote asymptotic dominance. That is, define $f(n) = \omega(g(n))$ to mean that for *all* constants $C > 0$, there exists an N such that $f(n) \geqslant C \cdot g(n)$ for all $n > N$. We say $f(n)$ (asymptotically) *dominates* $g(n)$.

Proof. If $t(n)$ dominates $\log n$, we have that in particular, for any $k \in \mathbb{N}$,

$$t(n) \;\geqslant\; \frac{k}{-\log \delta} \cdot \log n \;>\; \frac{k \cdot \log n}{-\log q(n)} \;=\; \frac{-\log n^k}{\log q(n)}$$

which implies $t(n) \cdot \log q(n) < -\log n^k$ and thus

$$q(n)^{t(n)} < \frac{1}{n^k}$$

as desired. □

Turning to security properties, note that witness-independence is preserved under parallel composition [FS90]. However, we have to revisit the claim of security under active attack.

Theorem 4.6 (security)**.** *Under the assumptions of theorem 4.3 and with at most a polynomial number of rounds (in n), a successful (active) attack on the parallel scheme shown in figure 4.3 can be used to construct a solution to $Col(h, \mathrm{Rg}(\mathtt{G}_2'))$.*

Proof. Proceed analogously to the single-round case. Generate $\mathbf{x}$ and let the adversary emit its commitments Y_i. Run the rest of the protocol twice for two families of challenges (c_i) and (d_i) generating answers (z_i) and (ζ_i), respectively. We must prove that there exists with non-negligible probability an i such that $c_i \neq d_i$ and both z_i and ζ_i are valid. Then with non-negligible probability

$$z_i - \mathbf{x}c_i \neq \zeta_i - \mathbf{x}d_i$$

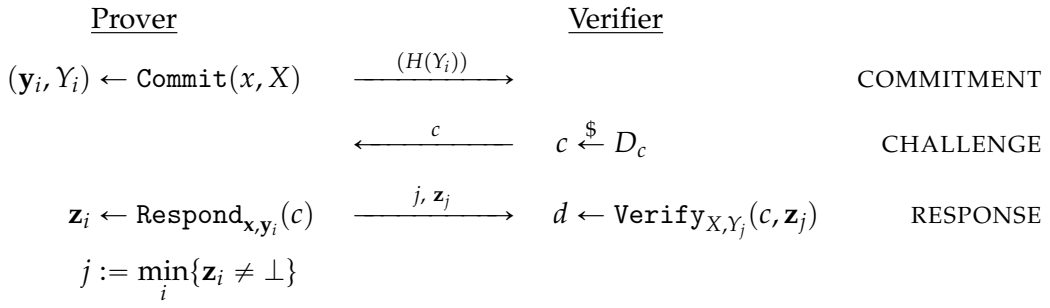

Figure 4.4: Parallel scheme modified to reduce communication complexity

holds, providing a collision, by the same argument as before: There exists at most one $\mathbf{x}$ for which the above fails and either $\mathbf{x}$ being chosen uniformly at random or witness-independence implies a high probability for collision.

Now let $\hat{p}$ be the probability for the attacker to succeed in the parallel scheme, i.e. to succeed with at least one of t responses. Given $t = O(n^k)$, some run i must succeed with non-negligible probability or $\hat{p}$ would be negligible. Then the probability that run i succeeds for both c_i and d_i is also non-negligible. As before, the possiblity that $c_i = d_i$ can be made negligible. □

Slight modifications to the protocol allow a further reduction in communication complexity without affecting security:

1. If the values Y_i are relatively large, it is a standard trick to send $H(Y_i)$ instead, where H is any collision-resistant hash function with a smaller codomain.
2. Instead of a family (c_i) it is possible to use the same challenge c in all parallel runs.
3. It is sufficient to respond with a single successful result.

The scheme incorporating the above modifications is shown in figure 4.4.

Firstly, we prove that the (asymptotical) number of runs required to ensure success remains the same as before.

Theorem 4.7 (aborts)**.** *If the probability for failure of the single-round scheme is bounded by a constant $\delta < 1$ for any challenge c, then the modified scheme of figure 4.4 with $\omega(\log n)$ rounds fails with the same (negligible) probability as the unmodified scheme (figure 4.3).*

Proof. If the bound δ is independent of the choice of c, the individual results $\mathbf{z}_i$ are valid with the same probability as in the unmodified case. The modified variant succeeds with j well-defined if and only if the unmodified parallel scheme would succeed. □

However, since this scheme consists no longer of simple parallel executions of the base protocol, we must formally revisit the proofs of both witness-independence and security.

Theorem 4.8 (witness-independence)**.** *If the single-round protocol 4.1 is targetable and the output of $\mathtt{G}_2'$ is uniformly distributed, then the modified parallel scheme as shown in figure 4.4 is witness-independent.*

Proof. Exactly as in the single-round case the results $\mathbf{z}_i$ are independent of $\mathbf{x}$, including the case that they are $\perp$. The modified protocol additionally reveals the choice j which as a function of the family $(\mathbf{z}_i)$ is also independent of $\mathbf{x}$. □

Theorem 4.9 (security)**.** *Under the assumptions of theorem 4.3 and with at most a polynomial number of rounds (in n), a successful attack on the modified scheme shown in figure 4.4 can be used to construct a solution to $Col(h, \mathrm{Rg}(\mathtt{G}_2'))$.*

Proof. Generate $\mathbf{x}$ and let the adversary emit its commitments Y_i. Run the rest of the protocol twice for two challenges $c \neq c'$ generating answers $(j, \mathbf{z})$ and $(j', \mathbf{z}')$, respectively. We must prove that with non-negligible probability $j = j'$ and both replies are valid to obtain a collision from

$$\mathbf{z} - \mathbf{x}c \neq \mathbf{z}' - \mathbf{x}c'$$

as before.

Let $\hat{p}$ be the probability for the attacker to succeed. Given $t = O(n^k)$, a valid response must appear with non-negligible probability for some j or $\hat{p}$ would be negligible. Then the probability that j results for both c and c' is also non-negligible. □

5 Lattice Foundations

The following chapters will apply the general constructions of the previous chapters to produce specific schemes. This chapter collects some prerequisites.

Section 5.2 introduces the lattice-based hash function due to Micciancio [Mic07]. It is one-way and restricted collision-resistant, relying on the conjectured hardness of a well-known lattice problem. Chapter 6 uses it to derive Lyubashevsky's one-time signature [LM08] and chapter 7 similarly arrives at Lyubashevsky's identification scheme [Lyu08a, Lyu08b].

Section 5.3 provides an important lemma that satisfies a precondition for the security proofs.

5.1 Ideal Lattices and the Shortest Vector Problem

A *lattice* is a free $\mathbb{Z}$-submodule of $\mathbb{R}^n$, i.e. a set

$$\sum_i^{n'} \mathbb{Z}\mathbf{v}_i$$

of integer combinations of linearly independent vectors $\mathbf{v}_i \in \mathbb{R}^n$. A lattice is called *full* or *full-rank* if $n' = n$. An *integer lattice* is a (free) $\mathbb{Z}$-submodule of $\mathbb{Z}^n$, i.e. a lattice where the $\mathbf{v}_i$ and in turn all lattice elements have integer coordinates. An *ideal lattice* is an integer lattice that forms an ideal under some ring structure defined on $\mathbb{Z}^n$.

Definition 5.1. The (approximate) *shortest vector problem* $\mathrm{SVP}_\gamma^p(\Lambda)$ asks to find a non-zero element $\mathbf{v}$ of the lattice Λ such that, with respect to the ℓ_p-norm,

$$\|\mathbf{v}\| \leqslant \gamma \cdot \lambda_1$$

where λ_1 is the length of a shortest non-zero element of Λ.

The actual difficulty of the problem depends on the lattice. If Λ is chosen at random, we speak of hardness in the *average case*. By contrast, with *worst case* hardness, a solution is sought for all lattices of a certain class. The shortest vector problem restricted to ideal lattices is denoted as Ideal-SVP. Up to certain magnitudes of approximation, general $\mathrm{SVP}_\gamma^\infty$ is known to be NP-hard in the worst case. It is conjectured that Ideal-$\mathrm{SVP}_\gamma^\infty$ and more specifically Ideal-$\mathrm{SVP}_\gamma^\infty$ restricted to particular rings are also hard in the worst case.

By Ajtai's central result, average-case hardness of SVP can be reduced to hardness in the worst case [Ajt96].

5.2 Micciancio's Lattice Hash Function

This section introduces the hash function used in the following chapters and with it the context in which all operations take place.

Consider a ring R, a natural number m, and define for any $\mathbf{a} \in R^m$

$$\begin{aligned} h_{\mathbf{a}} &: R^m \to R \\ h_{\mathbf{a}}(\mathbf{z}) &:= \mathbf{a} \cdot \mathbf{z} \end{aligned}$$

where $\mathbf{a} \cdot \mathbf{z}$ denotes the the standard inner ("dot") product. The function $h_{\mathbf{a}}$ is R-linear:

$$\begin{aligned} h_{\mathbf{a}}(\mathbf{x} + \mathbf{y}) &= h_{\mathbf{a}}(\mathbf{x}) + h_{\mathbf{a}}(\mathbf{y}) && , \mathbf{x}, \mathbf{y} \in R^m \\ h_{\mathbf{a}}(\mathbf{x}c) &= h_{\mathbf{a}}(\mathbf{x})c && , c \in R, \mathbf{x} \in R^m \end{aligned}$$

For convenience, the subscript $\mathbf{a}$ will be dropped in the following, h always considered an element of the set

$$\mathcal{H} := \{ h_{\mathbf{a}} \mid \mathbf{a} \in R^m \} \quad .$$

The ring we will be using is the ring of polynomials over $\mathbb{Z}_p$ modulo $x^n + 1$ where p is a prime number and n is a power of 2.

$$\begin{aligned} p &\quad \text{prime} \\ n &\quad \text{power of 2} \end{aligned}$$

$$\begin{aligned} f &:= x^n + 1 \\ R &:= \mathbb{Z}_p[x]/_f \end{aligned}$$

Before proceeding, we establish a prerequisite for the following theorems and the reason to choose n as a power of 2.

Lemma 5.1. *The polynomial $f = x^n + 1$ is irreducible in $\mathbb{Z}[x]$ if and only if n is a power of 2.*

Proof. The $2n$-th cyclotomic polynomial Φ_{2n} divides f since $(x^n + 1)(x^n - 1) = x^{2n} - 1$ and Φ_{2n} cannot divide $x^n - 1$. This means that f is irreducible if and only if it equals Φ_{2n} ; in fact if and only if $\deg(\Phi_{2n}) = \varphi(2n) = n$. Note that at most the n odd numbers less than $2n$ can be coprime to $2n$. Therefore $\varphi(2n) = n$ if and only if $2n$ — and thus n — has no odd factors. □

Besides irreducibility, the choice of f is related to the practical hardness of finding collisions via its *expansion factor* [Lyu08b]. The expansion factor, to be defined below, measures how the modular polynomial multiplication in R changes the "size" of its operands. To define what is meant by size in R, we take the obvious approach of using a norm on $\mathbb{Z}^n$ with respect to some set of representatives. Note that in R, due to being based on $\mathbb{Z}_p$, this is not actually a norm, though we will abuse the term when convenient.

Definition 5.2. Choose the following set of representatives for $\mathbb{Z}_p$

$$\left\{-\frac{p-1}{2}, \ldots, 0, \ldots, \frac{p-1}{2}\right\} \subset \mathbb{Z}$$

and let ρ be the function that maps an element of $\mathbb{Z}_p$ to its representative in the above set. Identify ρ with its natural (component-wise) extension to R

$$\rho : \mathbb{Z}_p[x]/f \to \mathbb{Z}[x]/f$$

and define for any $a \in R$ the pseudonorm

$$\|a\|_\infty := \|\rho(a)\|_\infty = \max_{1 \leqslant i \leqslant n} |\rho(a_i)|$$

where a_i are the coefficients of a. Likewise extend $\|\cdot\|_\infty$ to R^m; for any $\mathbf{b} \in R^m$, let

$$\|\mathbf{b}\|_\infty := \max_{1 \leqslant j \leqslant m} \|b_j\|_\infty$$

where b_j are the components of $\mathbf{b}$. This norm will be used throughout, so we drop the subscript ∞ in the following.

We state the following straightforward properties of the pseudonorm without proof.

Theorem 5.2. *For $a, b \in R$, the triangle inequality $\|a + b\| \leqslant \|a\| + \|b\|$ holds.* □

Theorem 5.3. *For $a \in R$, $s \in \mathbb{Z}$, the inequality $\|sa\| \leqslant |s| \cdot \|a\|$ holds.* □

As mentioned above, multiplication of elements in R affects the norm by an expansion factor which, perhaps surprisingly, may be considered a property of the polynomial f.

Definition 5.3. Let $f \in \mathbb{Z}_p[x]$ be a polynomial of degree n. Define the *expansion factor* ϕ of f as the smallest natural number such that

$$\|a \cdot x^i\| \leqslant \phi \cdot \|a\|$$

for all $a \in \mathbb{Z}_p[x]/f$ and $i \in \mathbb{N}$.

Theorem 5.4. *The expansion factor of $f = x^n + 1$ is 1.*

Proof. Multiplication with x, and by extension x^i, modulo $x^n + 1$ only rearranges the coefficients (up to sign), therefore leaving the norm unchanged. □

Lyubashevsky gives expansion factors for further polynomials. Since this work is only concerned with the case $f = x^n + 1$ where $\phi = 1$, we could decide to elide the expansion factor from our calculations but leave it in to keep the presentation general. Again we state without proof the following inequality which follows easily from the definitions; cf. [Lyu08b].

Theorem 5.5. *For $a, b \in R$, the inequality $\|ab\| \leqslant \phi n \cdot \|a\| \cdot \|b\|$ holds.* □

We can now formulate the main result. The hash function h is collision-resistant but only when restricted to a bounded subset $D \subset R^m$, which motivated much of the complexity in previous chapters. Choosing the area D is dependent on the choices of parameters m, n, and p; the following theorem from [LM06] makes the rather involved dependencies explicit. Note that it relates the average-case hardness of $\mathrm{Col}(h, D)$ to the worst-case hardness of a restriction of Ideal-SVP.

Theorem 5.6. *Let $D = \{\mathbf{y} \in R^m \mid \|\mathbf{y}\| \leqslant d\}$ for some integer d. Let $m > \log p / \log 2d$ and $p \geqslant 4\phi^2 dmn^{1.5} \log n$. If there is a polynomial-time algorithm that solves $Col(h, D)$ for a random $h \in \mathcal{H}$ with some non-negligible probability, then there is a polynomial-time algorithm that can solve SVP_γ^∞ for every ideal lattice in $\mathbb{Z}[x]/f$, where $\gamma = 16\phi^2 dmn(\log n)^2$.*

Proof. See [Lyu08b, LM06]. The irreducibility of f is required for this proof. □

NB. Cf. [PR06] for a similar result involving the polynomial $x^n - 1$.

Definition 5.4. Denote the problem of solving $\mathrm{SVP}_\gamma^\infty$ for every ideal lattice in $\mathbb{Z}[x]/f$ as f-$\mathrm{SVP}_\gamma^\infty$.

5.3 Cancellation in R

The following lemma provides a cancellation rule in R and R^m that will be used in the security reductions of both the following chapters. It requires the respective values to be sufficiently small and this forms one of the constraints that will lead to the chosen parameter ranges.

Lemma 5.7 (cancellation)**.** *Let $x_1, x_2, c \in R$ be such that (for $i = 1, 2$)*

$$2\phi n \cdot \|x_i\| \cdot \|c\| < p \quad ,$$

then $x_1 c = x_2 c$ implies $x_1 = x_2$ or $c = 0$.

Proof. This proposition would be trivially satisfied were R an integral domain. This is not the case here since $f = x^n + 1$ is not generally irreducible over the coefficient ring $\mathbb{Z}_p$. However, with n a power of 2, f is irreducible in $\mathbb{Z}[x]$ and so the ring

$$\mathbb{Z}[x]/f$$

of "unreduced" polynomials is an integral domain.[1] It can be proved that if calculating $x_i c$ over $\mathbb{Z}$ "involves no reduction", the desired cancellation property carries over into R. Formally, recall the system of representatives of $\mathbb{Z}_p$ upon which the pseudonorm $\|\cdot\|$ is based:

$$\left\{ -\frac{p-1}{2}, \ldots, 0, \ldots, \frac{p-1}{2} \right\}$$

[1] The quotient of a commutative ring ($\mathbb{Z}[x]$) modulo a prime element (f) is an integral domain.

Let ρ denote the induced representation of R.

$$\rho \ : \ R \to \mathbb{Z}[x]/_f$$

Also let π denote the (induced) natural projection onto R.

$$\pi \ : \ \mathbb{Z}[x]/_f \to R$$

Note that these are by definition inverse on the set of representatives.

$$\begin{aligned} \pi(\rho(x)) &= x \quad \forall x \in R \\ \rho(\pi(\xi)) &= \xi \quad \forall \xi \in \rho(R) \end{aligned}$$

Using $\pi\rho$ as a shorthand for $(\pi \circ \rho)$, rewrite $x_1 c = x_2 c$ as

$$\begin{aligned} \pi\rho(x_1)\pi\rho(c) &= \pi\rho(x_2)\pi\rho(c) \\ \Rightarrow \ \pi\left(\rho(x_1)\rho(c)\right) &= \pi\left(\rho(x_2)\rho(c)\right) \\ \Rightarrow \ \rho\pi\left(\rho(x_1)\rho(c)\right) &= \rho\pi\left(\rho(x_2)\rho(c)\right) \quad . \end{aligned}$$

Now consider

$$\|\rho(x_i)\rho(c)\| \ \leqslant \ \phi n \cdot \|\rho(x_i)\| \cdot \|\rho(c)\| \ = \ \phi n \cdot \|x_i\| \cdot \|c\| \quad .$$

By the assumption of the lemma, this implies $\|\rho(x_i)\rho(c)\| \leqslant \frac{p-1}{2}$ and therefore

$$\rho(x_i)\rho(c) \in \rho(R) \quad .$$

It follows that

$$\rho(x_1)\rho(c) = \rho(x_2)\rho(c)$$

and since this is an equation over coefficients in $\mathbb{Z}$, it implies $\rho(x_1) = \rho(x_2)$ or $\rho(c) = 0$ and thus $x_1 = x_2$ or $c = 0$ as claimed. □

The lemma easily extends componentwise to elements $\mathbf{x}_i$ of R^m.

Corollary 5.8. *Let $\mathbf{x}_1, \mathbf{x}_2 \in R^m$ and $c \in R$ be such that (for $i = 1, 2$)*

$$2\phi n \cdot \|\mathbf{x}_i\| \cdot \|c\| < p \quad ,$$

then $\mathbf{x}_1 c = \mathbf{x}_2 c$ implies $\mathbf{x}_1 = \mathbf{x}_2$ or $c = 0$. □

6 Lattice-Based One-Time Signature Scheme

In order to cast from the hash function (family) described in chapter 5 a one-time signature scheme according to section 3.2, we will supply a suitable private key generation algorithm $\mathtt{G}'$ and prove another lemma to satisfy the conditions of theorem 3.2. In addition to this, the particular construction of the key generator allows a proof of the *completeness* of the scheme, i.e. that (under any given key) every document has a valid signature.

6.1 Parameters

Given the security parameter n, theorem 5.6 requires further parameters p, m, and d satisfying certain conditions. Recall that p equals the number of possible coefficient values for the polynomials of R and d is the radius of D, the (assumed) area of collision resistance of h; m gives the dimension of R^m, the module in which operations are performed.

Let ϕ denote the expansion factor of f and set $m := \log n$. Assume $n \geqslant 4$ so that $m > 1$ and choose

$$p \geqslant \left(40 \cdot \phi^3 n^{2.5} \cdot m^4\right)^{\frac{m}{m-1}} \quad .$$

Define a constant $r := \lfloor 5p^{1/m} \rfloor$ and set

$$d := 2 \cdot \phi n \cdot r m^2 \quad .$$

Finally, let the set of "documents" be the polynomials in R with coefficients less than or equal to 1.

$$D_c = \{\, c \in R \mid \|c\| \leqslant 1 \,\}$$

As a note towards explanation: r serves as the *base radius* of the nested key sets B_i to be defined below. Its particular value becomes significant in the proof of lemma 6.3. The bound on p is chosen to satisfy the cancellation lemma and the assumptions of theorem 5.6. The value of d derives from the arithmetic operations of the signing function, ensuring that results lie in D.

6.2 Key Generation

Recall from definition 3.1 that the key generation function $\mathtt{G}$ is defined by

$$\mathtt{G} := \big((\mathbf{x}, \mathbf{y}),\ (h(\mathbf{x}), h(\mathbf{y})) \big) \qquad \text{where } (\mathbf{x}, \mathbf{y}) \leftarrow \mathtt{G}'.$$

To complete the scheme we provide the generator $\mathtt{G}'$ for the primary ($\mathbf{x}$) and secondary ($\mathbf{y}$) signing keys. The resulting scheme is sound by proposition 3.1.

In essence, the keys should be picked randomly from R^m; however, to ensure that signatures always end up in D (completeness), the sets of possible values for $\mathbf{x}$ and $\mathbf{y}$ must be restricted. In addition, a custom distribution is employed that is taylored to fit the security proof. Define the nested family of sets

$$B_\alpha = \{\, \mathbf{x} \in R^m \mid \|\mathbf{x}\| \leqslant r\alpha \,\} \quad \text{where } 1 \leqslant \alpha \leqslant m^2.$$

Now choose an index l as follows. Pick

$$\beta \xleftarrow{\$} \{0,1\}^{m^2-1}$$

uniformly at random. Set l to the index of the first 1 in β or $l = m^2$ when $\beta = 0$. Finally, pick $\mathbf{x}$ and $\mathbf{y}$ uniformly at random from B_l and $B_{l\phi n}$, respectively.

$$\mathbf{x} \xleftarrow{\$} B_l$$
$$\mathbf{y} \xleftarrow{\$} B_{l\phi n}$$

NB. The tighter restriction on $\mathbf{x}$ (by a factor of ϕn) accounts for its multiplication with c in the signing operation.

6.3 Completeness and Security

Given the above definitions, it is easy to see that signatures produced by the scheme are always valid. Recall:

$$S_{(\mathbf{x},\mathbf{y})}(c) := \begin{cases} \mathbf{x}c + \mathbf{y} & \text{if } \mathbf{x}c + \mathbf{y} \in D \\ \perp & \text{otherwise} \end{cases}$$

Theorem 6.1 (completeness)**.** *The lattice-based one-time signature scheme obtained by instantiating definition 3.1 with $\mathtt{G}'$ and associated parameters as defined above always produces valid signatures.*

Proof. Let $(\mathbf{x}, \mathbf{y})$ be produced by $\mathtt{G}'$ as defined above and $c \in D_c$. We will show

$$\mathbf{x}c + \mathbf{y} \in D \quad .$$

From the definition of $\mathtt{G}'$ and D_c we have

$$\|c\| \leqslant 1 \quad , \quad \|\mathbf{x}\| \leqslant rm^2 \quad , \quad \|\mathbf{y}\| \leqslant \phi nrm^2 \quad .$$

With the properties of $\|\cdot\|$ established in chapter 5 we deduce

$$\begin{aligned} \|\mathbf{x}c + \mathbf{y}\| &\leqslant \|\mathbf{x}c\| + \|\mathbf{y}\| \\ &\leqslant \phi n\|\mathbf{x}\|\|c\| + \|\mathbf{y}\| \\ &\leqslant 2\phi nrm^2 = d \quad . \end{aligned}$$

and thus $\mathbf{x}c + \mathbf{y}$ lies within the bounds of D. □

For the proof of security, we must prove the conditions of theorem 3.2. We have already established the cancellation property in a general form with corollary 5.8. For the second condition regarding "key probabilities" we will first establish some groundwork.

Consider the sets B_α and let

$$N_i := |B_i \times B_{i\phi n}| = |B_i| \cdot |B_{i\phi n}|$$

denote the number of private keys found at "level" i.

Lemma 6.2. *Given N_i defined as above, the ratio between the number of keys at consecutive levels is bounded as follows:*

$$\frac{N_{i+1}}{N_i} < 4^{mn}$$

Proof. From the definition of B_α we can see that

$$N_i = (2ri+1)^{mn} \cdot (2ri\phi n+1)^{mn}$$

and this yields

$$\frac{N_{i+1}}{N_i} = \left(\frac{2r(i+1)+1}{2ri+1} \cdot \frac{2r(i+1)\phi n+1}{2ri\phi n+1}\right)^{mn} \quad .$$

We observe $2r(i+1)+1 = (2ri+1)+2r$ and obtain

$$\frac{2r(i+1)+1}{2ri+1} = \left(1+\frac{2r}{2ri+1}\right) < 2$$

as well as

$$\frac{2r(i+1)\phi n+1}{2ri\phi n+1} = \left(1+\frac{2r\phi n}{2ri\phi n+1}\right) < 2$$

by analogue. Thus

$$\frac{N_{i+1}}{N_i} = \left(1+\frac{2r}{2ri+1}\right)^{mn}\left(1+\frac{2r\phi n}{2ri\phi n+1}\right)^{mn} < 4^{mn} \quad .$$

□

Now that we have an idea of how many more keys there are between levels, we will examine how many elements of the same hash value we can expect around any given private key. Together, these bounds will allow us to show that there are always such large sets of equivalent keys that picking out any particular one private key becomes negligibly unlikely. So consider the set

$$\begin{aligned} K &:= B_1 \ \cap \ \ker(h) \\ &= \{\mathbf{v} \in R^m \mid \|\mathbf{v}\| \leqslant \lfloor 5p^{1/m} \rfloor \ \wedge \ h(\mathbf{v}) = 0\} \end{aligned}$$

of "short elements" in the kernel of h.

Lemma 6.3. *The set $B_1 \cap \ker(h)$ contains more than 5^{mn} elements.*

$$|K| > 5^{mn}$$

Proof. Consider the subset of B_1 consisting of only those elements with all non-negative coefficients (with respect to our chosen representation).

$$B_1^+ := \{\, \mathbf{x} \in B_1 \mid \mathbf{x} \geqslant 0 \,\}$$

It is clear from the definitions that $|B_1^+| = (\lfloor 5p^{1/m} \rfloor + 1)^{mn}$. Noting that $p^{1/m}$ will be irrational, it follows that

$$\begin{aligned} |B_1^+| &= \lceil 5p^{1/m} \rceil^{mn} \\ &> (5p^{1/m})^{mn} = 5^{mn} p^n \end{aligned}$$

Recall $h : R^m \to R$ where $|R| = p^n$ and consider its restriction to B_1. By a simple counting argument there must exist a pre-image set $S \subset B_1^+$ with more than $|B_1^+| / |R| = 5^{mn}$ elements. Choose any $\mathbf{s} \in S$ and consider $S - \mathbf{s}$. It is easy to see that this is a subset of $\ker(h)$ as well as of B_1. Thus we have $S - \mathbf{s} \subset K$ with

$$|S - \mathbf{s}| > 5^{mn} \quad .$$

□

Lemma 6.4. *Let $(\mathbf{x}, \mathbf{y})$ denote a (fixed) private key and $c \in D_c$ a document. Define l as the smallest natural number such that $(\mathbf{x}, \mathbf{y})$ is contained in the set $B_l \times B_{l\phi n}$. Assume further that $l < m^2$, i.e. that the key lies not in the outermost layer. Then the probability for a randomly generated key $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \leftarrow \mathtt{G}'$ to fall on $(\mathbf{x}, \mathbf{y})$, even under the condition that it does hash to the same public key and produce the same signature for c, is negligible.*

$$\mathrm{P}\left((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y}) \mid h(\tilde{\mathbf{x}}) = h(\mathbf{x}) \ \wedge \ h(\tilde{\mathbf{y}}) = h(\mathbf{y}) \ \wedge \ \tilde{\mathbf{x}}c + \tilde{\mathbf{y}} = \mathbf{x}c + \mathbf{y}\right) = O(n^{-k})$$

Proof. For convenience, abbreviate the condition

$$h(\tilde{\mathbf{x}}) = h(\mathbf{x}) \ \wedge \ h(\tilde{\mathbf{y}}) = h(\mathbf{y}) \ \wedge \ \tilde{\mathbf{x}}c + \tilde{\mathbf{y}} = \mathbf{x}c + \mathbf{y}$$

as an event $\mathfrak{E}$. Observe that $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y})$ implies $\mathfrak{E}$. Therefore

$$\mathrm{P}\left((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y}) \mid \mathfrak{E}\right) = \frac{\mathrm{P}\left((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y}) \ \wedge \ \mathfrak{E}\right)}{\mathrm{P}\left(\mathfrak{E}\right)} = \frac{\mathrm{P}\left((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y})\right)}{\mathrm{P}\left(\mathfrak{E}\right)} \quad .$$

We will determine $\mathrm{P}\left((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y})\right)$ and establish a lower bound on $\mathrm{P}\left(\mathfrak{E}\right)$.

Recall that $(\mathbf{x}, \mathbf{y})$ is contained in the sets at level l and above. The probability to pick this particular key is thus found as the total probability over picking it from any of these levels.

$$\mathrm{P}\left((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y})\right) = \sum_{i=l}^{m^2} \left(\frac{1}{2^i} \cdot \frac{1}{N_i} \right)$$

To bound the probability of $\mathfrak{E}$, recall the set

$$K = \{\mathbf{v} \in R^m \mid \|\mathbf{v}\| \leqslant r \ \wedge \ h(\mathbf{v}) = 0\} \quad .$$

Consider any

$$\mathbf{x}' = \mathbf{x} + \mathbf{v} \quad \in \quad \mathbf{x} + K$$

and note that it satisfies $h(\mathbf{x}') = h(\mathbf{x})$ and is contained in B_{l+1}. There is a corresponding

$$\mathbf{y}' := \mathbf{x}c + \mathbf{y} - \mathbf{v}c$$

which satisfies $h(\mathbf{y}') = h(\mathbf{y})$, is contained in $B_{(l+1)\phi n}$, and yields

$$\mathbf{x}'c + \mathbf{y}' = \mathbf{x}c + \mathbf{y} \quad .$$

Thus every element of K corresponds to a key that is equivalent to $(\mathbf{x}, \mathbf{y})$ in the sense of $\mathfrak{E}$. Note that these keys may lie at level $l + 1$. Therefore

$$\mathrm{P}(\mathfrak{E}) \geqslant \sum_{i=l+1}^{m^2} \left(\frac{1}{2^i} \cdot \frac{|K|}{N_i} \right) = |K| \cdot \tilde{q}$$

where

$$\tilde{q} := \sum_{i=l+1}^{m^2} \frac{1}{2^i \cdot N_i} \quad .$$

We have

$$\mathrm{P}((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y})) = \tilde{q} + \frac{1}{2^l \cdot N_l}$$

and thus obtain

$$\frac{\mathrm{P}((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y}))}{\mathrm{P}(\mathfrak{E})} \leqslant \frac{1}{|K|} \cdot \left(1 + \frac{1}{2^l \cdot N_l \cdot \tilde{q}} \right) \quad .$$

Observe that, using lemma 6.2,

$$2^l \cdot N_l \cdot \tilde{q} \geqslant \frac{N_l}{2 \cdot N_{l+1}} > \frac{1}{2 \cdot 4^{mn}} \quad .$$

The proposition now follows with lemma 6.3:

$$\frac{\mathrm{P}((\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (\mathbf{x}, \mathbf{y}))}{\mathrm{P}(\mathfrak{E})} < \frac{1 + 2 \cdot 4^{mn}}{|K|} < \frac{1 + 2 \cdot 4^{mn}}{5^{mn}} \qquad \square$$

Theorem 6.5 (security)**.** *An effective attack against the one-time signature scheme with the given parameters is at least as hard as solving $Col(h, D)$.*

Proof. We may assume without loss of generality that the generated private key lies not in the outermost layer: A key from layer m^2 can only be chosen if the key generator picked $l = m^2$. That case occurs with the negligible probability

$$\frac{1}{2^{m^2}} = \frac{1}{2^{\log n \cdot \log n}} = \frac{1}{n^{\log n}}$$

and a negligible success probability in other cases would contradict the assumption of an effective attacker.

Apply theorem 3.2; its conditions are satisfied by corollary 5.8 and lemma 6.4. Towards the former, we must still show that

$$2\phi n \cdot \|\mathbf{x}\| \cdot \|c_1 - c_2\| < p$$

for any $\mathbf{x} \in B_{m^2}$ and $c_1, c_2 \in D_c$. Recall that the parameter p was chosen such that

$$p \geqslant \left(40 \cdot \phi^3 \cdot n^{2.5} \cdot m^4\right)^{\frac{m}{m-1}} > \left(20 \cdot \phi n \cdot m^2\right)^{\frac{m}{m-1}} \quad .$$

It follows that

$$p^{1-\frac{1}{m}} > 20 \cdot \phi n \cdot m^2$$

and further

$$\begin{aligned} p &> 20 \cdot \phi n \cdot p^{1/m} \cdot m^2 \\ &= 2\phi n \cdot 5p^{1/m} m^2 \cdot 2 \\ &> 2\phi n \cdot rm^2 \cdot 2 \quad . \end{aligned}$$

Thus, recalling that $\|\mathbf{x}\| \leqslant rm^2$ and $\|c_i\| \leqslant 1$, we have

$$p > 2\phi n \cdot \|\mathbf{x}\| \cdot \|c_1 - c_2\| \quad ,$$

completing the proof. □

Theorem 6.6. *With the given parameters, the function h is collision-resistant on D if $f\text{-}SVP_\gamma^\infty$ is hard for $\gamma = 16\phi dm^3 n = 32\phi^2 r m^5 n^2 = \tilde{O}(n^2)$.*

Proof. Apply theorem 5.6; we must show that $p \geqslant 4\phi^2 \cdot dmn^{1.5} \log n$ and $m > \frac{\log p}{\log 2d}$. Towards the first condition, observe that our choice of

$$p \geqslant \left(40 \cdot \phi^3 \cdot n^{2.5} \cdot m^4\right)^{\frac{m}{m-1}}$$

implies

$$p \geqslant 8\phi^3 \cdot n^{2.5} m^4 \cdot 5p^{1/m} \quad .$$

With the definitions of $r = \lfloor 5p^{1/m} \rfloor$ and $d = 2\phi n r m^2$ we obtain

$$\begin{aligned} p &\geqslant 4\phi^2 \cdot 2\phi n r m^2 \cdot m n^{1.5} \log n \\ &= 4\phi^2 \cdot dmn^{1.5} \log n \end{aligned}$$

as required. For the second condition, note that d contains the term r, so $d > p^{1/m}$ and therefore

$$\begin{aligned} (2d)^m &> p \\ \Rightarrow m \cdot \log 2d &> \log p \quad , \end{aligned}$$

completing the proof. □

Corollary 6.7. *The lattice-based one-time signature scheme is secure if f-SVP_γ^∞ is hard for $\gamma = 16\phi dm^3 n = 32\phi^2 rm^5 n^2 = \tilde{O}(n^2)$.* □

7 Lattice-Based Identification Scheme

To instantiate the identification scheme of chapter 4 with the lattice hash function, we must again supply a key generation (and commitment) algorithm, choose appropriate parameters, and prove any preconditions for the security reduction.

Set $m := \log n$ as before and define the sets

$$D_{\mathbf{x}} := \{ \mathbf{x} \in R^m \mid \|\mathbf{x}\| \leqslant 1 \}$$
$$D_{\mathbf{y}} := \{ \mathbf{y} \in R^m \mid \|\mathbf{y}\| \leqslant mn^2 \}$$

from which, respectively, the private key $\mathbf{x}$ and secret $\mathbf{y}$ are picked uniformly at random. Figure 7.1 shows the instantiated scheme. Define further the set of challenges

$$D_c := \{ c \in R \mid \|c\| \leqslant 1 \}$$

and the target set of valid responses

$$D := \{ \mathbf{z} \in R^m \mid \|\mathbf{z}\| \leqslant mn^2 - \phi n \} \quad .$$

The bound on D is chosen such as to satisfy targetability.

Lemma 7.1. *The lattice-based identification scheme is targetable.*

Proof. Let $\mathbf{x} \in D_{\mathbf{x}}$, $\mathbf{z} \in D$, and $c \in D_c$. From the definitions we have

$$\|\mathbf{z} - \mathbf{x}c\| \leqslant \|\mathbf{z}\| + \phi n \cdot \|\mathbf{x}\| \cdot \|c\| \leqslant mn^2 - \phi n + \phi n = mn^2$$

and therefore $\|\mathbf{z} - \mathbf{x}c\| \in D_{\mathbf{y}}$ as required. □

Corollary 7.2. *The lattice-based identification scheme is witness-independent.* □

Now choose the prime p such that

$$4\phi^2 \cdot m^3 n^{3.5} \leqslant p < (2mn^2)^m \quad .$$

Note that the upper bound $(2mn^2)^m$ grows faster than the lower bound, so this choice is possible for sufficiently large n. Both bounds derive directly from theorem 5.6, providing the collision resistance of h.

Theorem 7.3. *With the above parameters, the function h is collision-resistant on $D_{\mathbf{y}}$ if $f\text{-}SVP_\gamma^\infty$ is hard for $\gamma = 16\phi m^4 n^3$.*

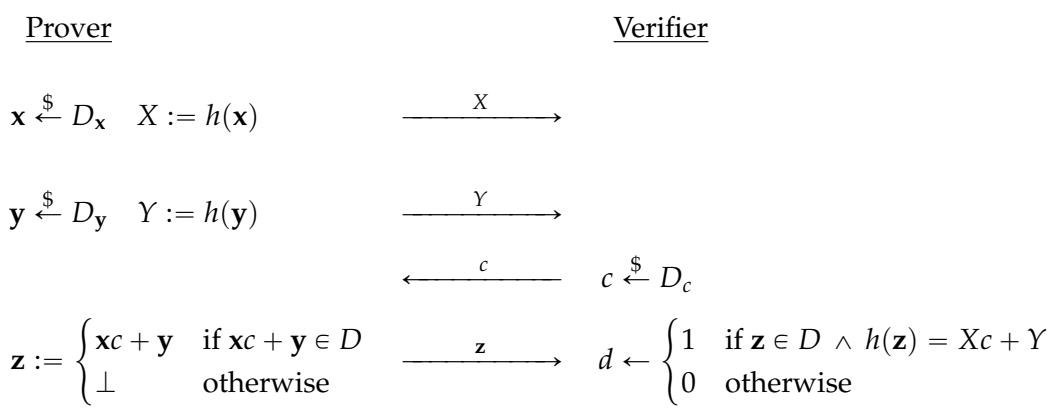

Figure 7.1: Lattice-based identification scheme

Proof. With

$$p \geqslant 4\phi^2 \cdot m^3 n^{3.5} = 4\phi^2 \cdot mn^2 \cdot mn^{1.5} \log n$$

where mn^2 is the radius of $D_{\mathbf{y}}$, and

$$\begin{aligned} & p < (2mn^2)^m \\ \Rightarrow \ & \log p < m \cdot \log(2mn^2) \quad , \end{aligned}$$

the conditions are satisfied to apply theorem 5.6. □

Given targetability and collision-resistance, the proofs of security and statistical completeness with multiple parallel executions follow by straight-forward application of the general theorems of chapter 4.

Theorem 7.4. *A successful attack on the lattice-based identification scheme can be used to construct a solution to $Col(h, D_{\mathbf{y}})$.*

Proof. Apply theorem 4.3. As shown above, the scheme is targetable by constrution. Prove the remaining conditions: Let $\mathbf{x} \in D_{\mathbf{x}}$ be any private key.

1. For any $c_1, c_2 \in D_c$ with $c_1 \neq c_2$ we have $\|c_1 - c_2\| \leqslant 2$ and thus

   $$2\phi n \cdot \|\mathbf{x}\| \cdot \|c\| \leqslant 4\phi n < p \quad .$$

   Therefore by corollary 5.8 the value of $\mathbf{x}$ is uniquely determined by $\mathbf{x}(c_1 - c_2)$.

2. To show that there must with overwhelming probability exist another key $\mathbf{x}' \in D_{\mathbf{x}}$ with $h(\mathbf{x}') = h(\mathbf{x})$ consider $h : R^m \to R$. By the pigeon-hole principle there can be at most $|R| = p^n$ values that have a unique pre-image in $D_{\mathbf{x}}$. Note that $\mathbf{x}$ is chosen uniformly at random, so the probability for a second pre-image $\mathbf{x}'$ to exist is at least

   $$1 - \frac{p^n}{|D_{\mathbf{x}}|} \ = \ 1 - \frac{p^n}{3^{mn}} \ > \ 1 - \frac{(2mn^2)^m}{3^{mn}} \quad .$$

To verify that the term

$$\frac{(2mn^2)^m}{3^{mn}} = \frac{nm^m n^{2m}}{3^{mn}}$$

is negligible, observe that $3^{mn} > 2^{mn} = n^n$ and $m^m = n^{\log\log n}$; therefore we have

$$\frac{(2mn^2)^m}{3^{mn}} < \frac{n^{1+2\log n+\log\log n}}{n^n} = n^{-n+1+2\log n+\log\log n} = O(n^{-k})$$

for any $k \in \mathbb{N}$.

3. The private key $\mathbf{x}$ is chosen uniformly at random.

□

Corollary 7.5. *The lattice-based identification scheme is secure in the active attack model if f-SVP_γ^∞ is hard for $\gamma = 16\phi m^4 n^3 = \tilde{O}(n^3)$.* □

Theorem 7.6. *The probability that the lattice-based identification scheme aborts is bounded by $1 - e^{-\phi}$ as n tends to infinitiy. That is, for any constant δ such that $\delta < 1$ and $\delta > 1 - e^{-\phi}$ we have*

$$\mathrm{P}\left(\mathbf{x}c + \mathbf{y} \notin D\right) < \delta$$

for sufficiently large n.

Proof. The elements of $D_{\mathbf{y}}$ consist of m polynomials of degree n with coefficients no larger in absolute value than mn^2. Thus we obtain

$$|D_{\mathbf{y}}| = (2mn^2 + 1)^{mn}$$

and similarly

$$|D| = (2(mn^2 - \phi n) + 1)^{mn} \quad .$$

By theorem 4.1, the lattice-based scheme succeeds with probability

$$q = \left(\frac{2(mn^2 - \phi n) + 1}{2mn^2 + 1}\right)^{mn} \quad .$$

Note that

$$\frac{2(mn^2 - \phi n) + 1}{2mn^2 + 1} = 1 - \frac{2\phi n}{2mn^2 + 1} > 1 - \frac{\phi}{mn}$$

and therefore q is greater than

$$\left(1 - \frac{\phi}{mn}\right)^{mn}$$

which approaches $e^{-\phi}$ from below as n (and thus mn) tends to infinity. □

Corollary 7.7. *A lattice-based scheme constructed as in section 4.3 with $t(n)$ parallel rounds where $t(n) = \omega(\log n)$ and $t(n) = O(n^k)$ is statistically complete, witness-independent, and secure against active attack.*

Proof. Apply theorems 4.7, 4.8, and 4.9 □

8 Conclusion

We have formulated two of Lyubashevsky's constructions in the abstract setting of a module and subsequently instantiated them with a hash function that is restricted collision-resistant if approximate f-SVP is hard in the worst case.

Chapter 3 presented the abstract one-time signature scheme after motivating its form. Restricted collision resistance was defined here, accomodated by allowing a partial signing function. The security proof relied on two assumptions, namely a cancellation property in the ring and an information-hiding property of the key generation. Chapter 4 then showed how a canonical identification scheme derives from the one-time signature. Witness-independence was shown assuming a form of closure under restricted sets (targetability). The proof of security again relied on cancellation. The proofs were transferred to parallel protocol variants that recover (statistical) completeness in the case that restricted collision-resistance causes failures.

Chapter 5 defined lattices and the shortest vector problem before introducing Micciancio's lattice hash function and citing the reduction from Col to SVP. It also provided the cancellation lemma needed for the security proofs. Chapter 6 showed the corresponding instantiation of the one-time signature scheme using a tailored key distribution that favors smaller values. A lower bound on the order of $\tilde{O}(n^{2.5})$ was derived for the integer modulus p and security follows if f-SVP is hard for an approximation factor also on the order of $\tilde{O}(n^{2.5})$. Chapter 7 instantiated the identification scheme using a simple uniform key distribution and a lower bound for p on the order of $\tilde{O}(n^{3.5})$. Security follows if f-SVP is hard for $\gamma = \tilde{O}(n^3)$. The scheme becomes statistically complete with $\omega(\log n)$ rounds. Both schemes use a dimension of $m = \log n$ which results in key sizes of $\tilde{O}(n)$ bits.

Further Directions

A natural continuation of this work would be to extend the abstract treatment to existing schemes that build on those presented here. Lyubashevsky [Lyu09], for instance, adapts the well-known Fiat-Shamir paradigm [FS87, AABN02] to construct a signature scheme from the possibly aborting lattice identification scheme. Rückert [Rüc10, Rüc11] extends the transformation to include a blinding factor that yields a blind signature scheme.

On another abstract front, a sore point in this work is that it might appear intuitive for the construction of an identification scheme from a (one-time) signature to be secure if the signature is secure. The reduction, however, does not work because of differences in the security notions. It could be interesting if a further investigation of the connection between one-time/split-key/two-tier signatures and identification schemes would make

this clearer.

On the lattice side of things, the hardness of restricted forms of SVP such as Ideal-SVP and f-SVP is an area of active research. It has been shown that algorithms for finding short vectors can indeed be improved by a linear factor in ideal lattices [Sch13]. While not affecting the asymptotic complexity, this speedup could be significant in practice. It is also important to remember that all our reductions are to *approximate* SVP. There is now a public contest to find short vectors in ideal lattices [PS13]. In this vein, it is important to look for ways to tighten the approximation parameters. One area that seems "wasteful" in this regard are the arguments concerning $\ker(h)$ that are based on very general counting and pigeon-hole arguments (cf. lemma 6.3 and theorem 7.4). On the whole of R^m, due to the linearity of h, the number of elements that map to 0 must be exactly $|R^m| / |R|$. One expects these pre-images to be spread evenly and so, for instance, the set $B_1 \cap \ker(h)$ that is the subject of lemma 6.3 should likewise contain about $|B_1| / |R|$ elements, a much better result than the general argument is able to give.

Bibliography

[AABN02] Michel Abdalla, Jee Hea An, Mihir Bellare, and Chanathip Namprempre. From identification to signatures via the fiat-shamir transform: Minimizing assumptions for security and forward-security. In Lars R. Knudsen, editor, *Advances in Cryptology — EUROCRYPT 2002*, volume 2332 of *Lecture Notes in Computer Science*, pages 418–433. Springer, Berlin, 2002.

[Ajt96] Miklós Ajtai. Generating hard instances of lattice problems. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing*, pages 99–108. ACM, New York, 1996.

[BS07] Mihir Bellare and Sarah Shoup. Two-tier signatures, strongly unforgeable signatures, and fiat-shamir without random oracles. In Tatsuaki Okamoto and Xiaoyun Wang, editors, *Public Key Cryptography — PKC 2007*, volume 4450 of *Lecture Notes in Computer Science*, pages 201–216. Springer, Berlin, 2007.

[CS99] John H. Conway and Neil J.A. Sloane. *Sphere packings, lattices, and groups*. Number 290 in Grundlehren der mathematischen Wissenschaften. Springer, Berlin, 1999.

[DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976.

[FS87] Amos Fiat and Adi Shamir. How to prove yourself: practical solutions to identification and signature problems. In Andrew M. Odlyzko, editor, *Advances in cryptology — CRYPTO '86: Proceedings*, volume 263 of *Lecture Notes in Computer Science*, pages 186–194. Springer, Berlin, 1987.

[FS90] Uriel Feige and Adi Shamir. Witness indistinguishable and witness hiding protocols. In Harriet Ortiz, editor, *Proceedings of the Twenty-second Annual ACM Symposium on Theory of Computing*, pages 416–426. ACM, New York, 1990.

[Gol01] Oded Goldreich. *Foundations of Cryptography: Volume 1, Basic Tools*. Cambridge University Press, New York, 2001.

[Gol04] Oded Goldreich. *Foundations of Cryptography: Volume 2, Basic Applications*. Cambridge University Press, New York, 2004.

[Knu76] Donald E. Knuth. Big omicron and big omega and big theta. *ACM SIGACT News*, 8(2):18–24, 1976.

[Lam79] Leslie Lamport. Constructing digital signatures from a one-way function. Technical report, SRI International Computer Science Laboratory, October 1979.

[LM06] Vadim Lyubashevsky and Daniele Micciancio. Generalized compact knapsacks are collision resistant. In Ingo Wegener, Vladimiro Sassone, and Bart Preneel, editors, *Proceedings of the 33rd international colloquium on automata, languages and programming — ICALP 2006*, volume 4052 of *Lecture Notes in Computer Science*, pages 144–155. Springer, Berlin, 2006.

[LM08] Vadim Lyubashevsky and Daniele Micciancio. Asymptotically efficient lattice-based digital signatures. In Ran Canetti, editor, *Theory of Cryptography: Fifth Theory of Cryptography Conference, TCC 2008*, volume 4948 of *Lecture Notes in Computer Science*, pages 37–54. Springer, Berlin, 2008.

[Lyu08a] Vadim Lyubashevsky. Lattice-based identification schemes secure under active attacks. In Ronald Cramer, editor, *Public Key Cryptography — PKC 2008*, volume 4939 of *Lecture Notes in Computer Science*, pages 162–179. Springer, Berlin, 2008.

[Lyu08b] Vadim Lyubashevsky. *Towards Practical Lattice-based Cryptography*. PhD thesis, University of California, San Diego, 2008.

[Lyu09] Vadim Lyubashevsky. Fiat-Shamir with aborts: Applications to lattice and factoring-based signatures. In *Advances in Cryptology — ASIACRYPT 2009*, volume 5912 of *Lecture Notes in Computer Science*, pages 598–616. Springer, Berlin, 2009.

[Mic07] Daniele Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. *Computational Complexity*, 16(4):365–411, 2007.

[NV09] Phong Q. Nguyen and Brigitte Vallée, editors. *The LLL Algorithm: Survey and Applications*, chapter Cryptographic functions from worst-case complexity assumptions, pages 427–452. Springer, 2009.

[PR06] Chris Peikert and Alon Rosen. Efficient collision-resistant hashing from worst-case assumptions on cyclic lattices. In *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006*, volume 3876 of *Lecture Notes in Computer Science*, pages 145–166. Springer, Berlin, 2006.

[PS13] Thomas Plantard and Michael Schneider. Creating a challenge for ideal lattices. Cryptology ePrint Archive, Report 2013/039, 2013.

[Rüc10] Markus Rückert. Lattice-based blind signatures. In Masayuki Abe, editor, *Advances in Cryptology — ASIACRYPT 2010*, volume 6477 of *Lecture Notes in Computer Science*, pages 413–430. Springer, Berlin, 2010.

[Rüc11] Markus Rückert. *Lattice-based Signature Schemes with Additional Features*. PhD thesis, TU Darmstadt, 2011.

[Sch13] Michael Schneider. Sieving for shortest vectors in ideal lattices. In Amr Youssef, Abderrahmane Nitaj, and Aboul-Ella Hassanien, editors, *Progress in Cryptology – AFRICACRYPT 2013*, volume 7918 of *Lecture Notes in Computer Science*, pages 375–391. Springer, Berlin, 2013.