Efficient Authentication Protocols from the Restricted Syndrome Decoding Problem

Thomas Johansson, Mustafa Khairallah, Vu Nguyen
Lund University

***Abstract*—In this paper, we introduce an oracle version of the Restricted Syndrome Decoding Problem (RSDP) and propose novel authentication protocols based on the hardness of this problem. They follow the basic structure of the HB-family of authentication protocols and later improvements but demonstrate several advantages.**

An appropriate choice of multiplicative subgroup and ring structure gives rise to a very efficient hardware implementation compared to other *Learning Parity with Noise* based approaches. In addition, the new protocols also have lower key size, lower communication costs, and potentially better completeness/soundness compared to learning-based alternatives. This is appealing in the context of low-cost, low-powered authenticating devices such as radio frequency identification (RFID) systems. Lastly, we show that with additional assumptions, RSDP can be used to instantiate a Man-in-the-Middle secured authentication protocol.

1. Introduction

Authentication serves as one of the foundational pillars of cryptography, playing a crucial role in securing communication and data integrity. As technology advances, particularly with the widespread adoption of low-cost RFID tags, the demand for lightweight yet robust authentication protocols has intensified. Due to the efficient nature of using only simple operations, *Learning parity with Noise* (LPN) emerged as a prime candidate for constructing lightweight identification protocol. The LPN problem does not only stand out under the "practical" but also from the theoretical viewpoint. It is closely related to the well-studied *decoding a random linear code* problem or the equivalent *syndrome decoding* problem, and they are both strongly believed to resist known quantum attacks.

However, traditional approaches often struggled to balance efficiency and security, prompting an ongoing search for solutions catering to resource-constrained environments without compromising security. For example, early attempts at employing LPN, e.g., [1], [2], [3], [4], [5], [6], were met with security issues when examined under the scopes of different, more advanced adversarial models, such as in [7], [8], [9]. Numerous attempts to improve and explore other (yet still related to the LPN-based) directions have been proposed in recent years [10], [11], [12]. They appear to have achieved some level of security while still promising efficiency.

Hopper and Blum [1] laid the foundation for LPN-based authentication with their strikingly simple 2-round design, called HB. It provides provable security under *passive attack*, in which an adversary can only eavesdrop on the communications. It was noted by Juels and Weis [2] that an active attacker (i.e., with query power) could easily break HB, and they proposed an augmented 3-round version called HB+. HB+ was shown to be susceptible to attacks in the Man-in-the-Middle attack model (MitM) [7]. Gilbert et al. [6] proposed HB# to resist the attack from [7]. To make the proposal suitable for low-cost hardware, the authors also proposed using $\mathbf{X}$ as a *Toepliz* matrix, which implies a slightly different LPN hardness assumption. Unfortunately, in a more general model of MitM adversary, HB# was shown to be not sufficient [9].

LPN-based MitM-secured authentication was finally achieved with a series of breakthrough works: Kiltz et al. [11] proposed a variant of LPN called *Subset-LPN* and built efficient MACs based on this problem. Notably, they also achieve ϵ security (compared to $\sqrt{\epsilon}$ for other LPN-based protocols, excluding HB) by avoiding *rewinding* in their security proof. One of the drawbacks is a large key size that can be prohibitive in some cases. While HB$^{\#}$ can circumvent this problem by using "structural" LPN (e.g., employing a Toepliz matrix) [6], it is unfortunately not the case for [11]. Li et al. in [10] showed an interesting design called LCMQ without constructing MACs by applying a chosen ciphertext secure encryption on top of the LPN samples. To reduce communication and the computation burden on the low-cost tag, the encryption scheme uses the so-called P2-circulant matrix. By masking the LPN samples, they proposed aggressive parameters for their claim security, thus demonstrating the efficiency of the design. However, Nguyen et al. [13] have recently devised a key-recovery attack on LCMQ, which could significantly compromise the performance by forcing bigger key sizes.

Further advancement came when Lyubashevsky and Masny showed how to generically build a MitM-secured authentication from any (randomized) weak pseudo-random functions (wPRFs), which only has to fulfill a few (reasonable) properties [12]. Therefore, LPN can be seen as an instantiation of this design. Other well-studied LPN-type assumptions were employed to achieve manageable key sizes and efficiency required for low-cost environments, such as *Ring-LPN* [14] and *Toepliz-LPN* [15]. Despite the tightness of the security proof being "only" $\sqrt{\epsilon}$, their proposal remains impressive as it is more efficient than MACs construction like [11]. Moreover, it is unclear how the gap in the security proof affects instances for practical uses.

Contributions. Similarly to LPN, the *Restricted Syndrome Decoding Problem* (RSDP) is another NP-hard variant of SDP. It was first proposed by Baldi et al. in [16] as a new research direction toward efficient code-based zero-knowledge identification. It has recently been featured in one of the post-quantum signature candidates, CROSS [17], in the ongoing NIST additional call for signatures.

In this work, we introduce an Oracle version of the RSDP and propose novel authentication protocols based on the hardness of this problem. We first explore the idea of building a highly efficient RSDP-based authentication protocol that is secure in the *active attack* model. By selecting the secret keys from a suitable restricted set, multiplication in the finite field can be as simple as a cyclic shift, which can be performed very efficiently even with low-cost RFID tags. Our designs also have lower key size, lower communication costs, and potentially better completeness/soundness compared to HB-like protocols based on LPN or similar problems. On the theoretical side, restricting the secret keys in such a fashion does not compromise the hardness of RSDP. Moreover, relying on the best cryptanalytic tools for RSDP [17], we show, for example, that the key size for 128-bit security can be as low as 396-bit compared to 768 for LPN-based constructions. We further propose another stronger design that is proven to be secure in the MitM model. The cost is larger keys and more costly implementation, but still, the design compares favorably with other protocols secure in the MitM model.

Organization. In Section 2, we give the basics on authentication protocols, coding theory, LPN, and RSDP. We introduce some new definitions in relation to RSDP. In Section 3, we give the basic RSDP authentication protocol with security against active attacks. The proof of security is given in the same section. Section 4 is devoted to parameter instantiation and performance evaluation, including some comparisons based on FPGA and ASIC implementations. Section 5 then gives an extended protocol that is secure in the MitM and includes proof of this fact.

2. Preliminaries

Throughout the paper, we use the notation

- $a, \mathbf{a}, \mathbf{A}$ for single elements, vectors, and matrices, respectively.
- $\langle \cdot, \cdot \rangle$ the inner products of two vectors.
- $\mathbf{I}_n$ the identity matrix of size $n \times n$.
- $\mathbb{F}_p$ the finite field of order p, where p is a prime and $\mathbb{F}_p^n$ is the corresponding vector space of dimension n.
- $a \xleftarrow{\$} A$ an element drawn uniformly at random from some set A.

2.1. Authentication Protocol

We are interested in an interactive authentication protocol where we define two entities: a *Tag* (a.k.a Prover, denoted by $\mathcal{T}$) and a *Reader* (a.k.a Verifier, denoted by $\mathcal{R}$). Together, they share secret keys, which have been communicated via a secure channel before the authentication begins. Moreover, they are also parameterized by other (public) values, such as the length of the secret key, the domain, and so on. They interact on an insecure communication channel.

After the interaction, $\mathcal{R}$ either outputs accept or reject. An authentication protocol is said to have a completeness error P_c if $\mathcal{R}$ rejects a legitimate $\mathcal{T}$ with probability at most P_c. Conversely, the protocol is said to have a soundness error P_s if $\mathcal{R}$ accepts a random response from $\mathcal{T}$ with probability at most P_s.

The security of an authentication protocol depends on the adversarial power. We consider the most common adversary models: *passive attack*, *active attack*, and *Man-in-the-middle* attack. For all adversarial models, an attacker $\mathcal{A}$ operates in two phases described in the following.

Passive attack. A passive attacker $\mathcal{A}$ does not have query power: it can only observe interactions between $\mathcal{T}$ and $\mathcal{R}$ for a (polynomial) number of times in Phase 1. Then $\mathcal{A}$ tries to impersonate the $\mathcal{T}$ in Phase 2. In particular, in this phase, $\mathcal{A}$ only has one chance to convince $\mathcal{R}$. An authentication protocol is secure against such an attack if $\mathcal{R}$ outputs accept to such $\mathcal{A}$ with negligible probability.

Active attack. This attack mode is often called the detection-based (DET) model in previous literature. Here, $\mathcal{A}$ is given more power in Phase 1: $\mathcal{A}$ is allowed to interact with an honest $\mathcal{T}$ (e.g., sending "challenges" of its choosing) and observe the responses and learn about the authentication output. Similarly, in Phase 2, $\mathcal{A}$ interacts with $\mathcal{R}$, attempting to pass as an honest $\mathcal{T}$.

Man-in-the-Middle attack. In this scenario, $\mathcal{A}$ is given the power to manipulate communication between honest $\mathcal{T}$ and $\mathcal{R}$ in Phase 1. For example, it can modify messages in both directions and specifically observe the authentication output, i.e., whether $\mathcal{R}$ accepts or rejects a given transcript. The behavior of $\mathcal{A}$ in Phase 2 and the security of an authentication protocol in this model is defined similarly as in the active model.

An authentication protocol is said to be (t, Q, ϵ)-secure in the X-model (X is either passive, active, or MitM) if for all X-adversaries $\mathcal{A}$, running in time t, making Q queries with $\mathcal{T}$, the probability of $\mathcal{R}$ outputs accept in Phase 2 is at most ϵ.

2.2. Coding theory

Let p be a prime number and $\mathbb{F}_p$ be a finite field. A $[n, k]$-linear code $\mathcal{C}$ over $\mathbb{F}_p$ ($k \leq n$) is a vector subspace of dimension k in $\mathbb{F}_p^n$. A full-rank matrix $\mathbf{G} \in \mathbb{F}_p^{k \times n}$ is said to be the generator of $\mathcal{C}$ if $\mathcal{C} = \{\mathbf{uG} : \mathbf{u} \in \mathbb{F}_p^k\}$, i.e., $\mathbf{G}$ is a basis of $\mathcal{C}$. Let $\mathbf{H} \in \mathbb{F}_p^{(n-k) \times n}$ be a matrix such that $\mathbf{GH}^\intercal = \mathbf{0}$. Then $\mathbf{H}$ is called the *parity-check* matrix of $\mathcal{C}$ and for $\mathbf{y} \in \mathbb{F}_p^n$, $\mathbf{s} = \mathbf{yH}^\intercal$ is called its *syndrome* (w.r.t $\mathbf{H}$). As $\mathbf{G}$ is full-rank, it is sometimes convenient to assume $\mathbf{G}$ to be in its *systematic form*, that is $\mathbf{G} = \begin{pmatrix}\mathbf{I}_k & \mathbf{A}\end{pmatrix}$ for some $\mathbf{A} \in \mathbb{F}_p^{k \times (n-k)}$. We can then readily compute the parity-check matrix, also in systematic form, as $\mathbf{H} = \begin{pmatrix}-\mathbf{A}^\intercal & \mathbf{I}_{n-k}\end{pmatrix}$. The Hamming weight of a vector $\mathbf{x} \in \mathbb{F}_p^n$, denoted by $\omega_H(\mathbf{x})$, is defined as the number of its non-zero coordinates.

Problem 1 (Syndrome Decoding Problem (SDP)). Given $\mathbf{H} \in \mathbb{F}_p^{(n-k) \times n}$, a syndrome $\mathbf{s} \in \mathbb{F}_p^{n-k}$, and a positive integer $t \leq n$. Find $\mathbf{e}$ (if any) where $\omega_H(\mathbf{e}) \leq t$ such that $\mathbf{e}\mathbf{H}^\intercal = \mathbf{s}$.

It is a well-studied NP-complete [18], highlighted by the fact that it has been a crucial building block in many cryptosystems, such as zero-knowledge proof [19], [20], signatures [21], hash functions [22], [23], stream ciphers [24], and so on. In particular, for post-quantum cryptography, it has been popular and appeared in many public key cryptosystems (such as McEliece [25], BIKE [26], HQC [27]), as well as the recent NIST call for signatures (CROSS [17], SDitH [28], Wave [29], and so on).

One also finds many variants of SDP, such as the *Restricted Decoding Problem*, *Regular Decoding Problem*, *Permuted Kernel Problem*, ..., or the SDP defined with different metrics (e.g., rank metric, Lee metric), that all prove useful in constructing cryptographic primitives. This work involves employing the *Restricted Decoding Problem* (RSDP) in lightweight authentication protocols.

Let $\mathbb{F}_p$ be a finite field where p is a prime, and let $\mathbb{E} = \{g^i, i = 0, \dots, z\}$ be a multiplicative subgroup of order z, generated by some element $g \in \mathbb{F}_p$.

Problem 2 (Restricted Syndrome Decoding Problem). Given $\mathbf{H} \in \mathbb{F}_p^{(n-k)\times n}$, $\mathbf{s} \in \mathbb{F}_p^{n-k}$. Find $\mathbf{e} \in \mathbb{E}^n$ such that $\mathbf{e}\mathbf{H}^\intercal = \mathbf{s}$.

The RSDP was first introduced in [16] for $z = 2$. Moreover, the original RSDP asks for non-full weight $\mathbf{e}$, that is, entries of $\mathbf{e}$ are sampled from $\mathbb{E}_0 := \mathbb{E} \cup \{0\}$ and $\omega_H(\mathbf{e}) \leq t$ for some $t \in \mathbb{N}$. Notably, via a reduction to the classical SDP, they showed that this new problem is also NP-complete. As an application, the authors proposed a zero-knowledge identification scheme (adaptation by replacing SDP with RSDP), which is promising in terms of reduced public key size and communication cost. Recently, the versatility of RSDP has been extended by Baldi et al. [17], where the order z is no longer restricted to $z = 2$. In addition, CROSS - a digital signature scheme based on maximum-weight RSDP (as defined in Problem 2) was proposed.[1] Such a modification does not compromise the problem hardness [30], and contrarily to Hamming-weight RSDP, the uniqueness of the solution is still guaranteed.

For our application (in later sections), we rely on the security of RSDP in the "oracle form". We thus define the so-called RSDP Oracle, which will later be used in our construction.

Definition 1 (RSDP-Oracle). Let p be a prime number and $\mathbb{E} = \{g^i, i = 1, \dots, z\}$ be a multiplicative subgroup of order z in $\mathbb{F}_p$. Fix a secret $\mathbf{s} \in \mathbb{F}_p^k$. The RSDP Oracle, denoted by $\mathcal{O}_{\mathbf{s}}^{\mathsf{RSDP}}$, gives pairs of samples

$$\{\mathbf{a} \in \mathbb{F}_p^k, b = \langle \mathbf{a}, \mathbf{s} \rangle + e \bmod p\}$$

where $\mathbf{a} \xleftarrow{\$} \mathbb{F}_p^k$ and $e \xleftarrow{\$} \mathbb{E}$.

We write $\Lambda_k(\mathbf{s})$ for the distribution over $\mathbb{F}_p^k \times \mathbb{F}_p$, where the samples are obtained by querying $\mathcal{O}_{\mathbf{s}}^{\mathsf{RSDP}}$. The *decisional* RSDP, written as ORSDP_k, is to distinguish the $\Lambda_k(\mathbf{s})$ samples from the uniform distribution U_{k+1}. On the other hand, the *search* version asks for the recovery of the secret $\mathbf{s}$.

Definition 2. The ORSDP_k is said to be (t, Q, ϵ)-hard if for every algorithm D running in time t, making Q oracle queries

$$\left|\Pr[\mathbf{s} \xleftarrow{\$} \mathbb{F}_p^k : D^{\Lambda_k(\mathbf{s})} = 1] - \Pr[D^{U_{k+1}} = 1]\right| \leq \epsilon,$$

where we denote D^X by the algorithm D taking oracle input from a distribution X.

Next, we informally argue that the decisional RSDP problem is **as hard as** the search version via simple reduction. Indeed, suppose there exists some distinguisher D that can solve ORSDP_k with non-negligible probability ϵ. It can then be used to recover $\mathbf{s}$. Let $\{\mathbf{a}, b\}$ be a $\mathcal{O}_{\mathbf{s}}^{\mathsf{RSDP}}$ sample. The distinguisher D then picks $\bar{s}_1 \in \mathbb{F}_p$ as its guess for the first value of $\mathbf{s}$, and computes

$$\{\mathbf{a}' = (a_2, a_3, \dots, a_k), b' = b - a_1\bar{s}_1\}.$$

If the guess $\bar{s}_1$ is correct, then $\{\mathbf{a}', b'\}$ is precisely a sample from $\Lambda_{k-1}(\mathbf{s})$. By definition, D can successfully distinguish such samples after Q queries with probability at least ϵ. On the other hand, if the guess is incorrect, b' will be independent of $\mathbf{a}'$ (since a_1 is chosen uniformly at random). In such a case, D enjoys no advantage. Repeating the same procedure in the order of $1/\epsilon$ times, D eventually recovers $\mathbf{s}$.

One can see that trying to recover the secret $\mathbf{s}$ after some n queries amounts to finding a noisy codeword $\mathbf{b}$ in the code $\mathcal{C}$ generated by $\mathbf{a}_i, i = 1, \dots n$, where the noise consists of elements in $\mathbb{E}$. Equivalently, it suffices to find $\mathbf{e} \in \mathbb{E}^n$ such that $\mathbf{e}\mathbf{H}^\intercal = \mathbf{b}\mathbf{H}^\intercal = \mathbf{s}$. In conclusion, the hardness of RSDP implies the hardness of ORSDP_k, which again implies the "pseudo-randomness" of $\Lambda_k(\mathbf{s})$. In the next section, we show that it also extends to the case where the secret is not drawn uniformly at random.

When the number of oracle calls is unlimited and the field size in the problem is fixed, then the RSDP problem from Definition 1 can be solved in polynomial time. This follows from algebraic attacks in the style of Aurora-Ge [31] and can be launched whenever there are many values for the noise with probability 0. To this end, we introduce a slightly modified version of the RSDP-Oracle, called the δ-RSDP-Oracle.

Definition 3 (δ-RSDP-Oracle). Fix $0 < \delta < 1$. The δ-RSDP-Oracle, denoted by $\mathcal{O}_{\mathbf{s}}^{\delta-\mathsf{RSDP}}$, responds as an RSDP-Oracle with probability $1-\delta$ and with a random pair of samples $\{\mathbf{a} \in \mathbb{F}_p^k, b \in \mathbb{F}_p\}$ with probability δ.

To sum up, we consider two presumably difficult problems, the first one is the problem of distinguishing the RSDP-Oracle from random assuming a limited fixed number Q of queries, which corresponds to the RSDP problem. The second one is the same for the δ-RSDP-Oracle without limit on queries. In the sequel, we use the RSDP-Oracle, but the modifications needed to fit the use of δ-RSDP-Oracle are rather straightforward.

RSDP with a non-uniformly random secret. We briefly also discuss the hardness of RSDP when the secret $\mathbf{s}$ is not chosen randomly. In particular, the entries of the secret itself $\mathbf{s}$ can be drawn from the same distribution as the error, i.e., randomly from $\mathbb{E}$. The reason for wanting this specific choice is that it lowers the secret key size and can give rise to very efficient implementations. For

[1] A specialized version of RSDP, called RSDP(G) was also investigated in the same work.

instance, with a proper choice of p and $\mathbb{E}$, e.g., $p = 127$ and $\mathbb{E} = \{2^i, i = 0, \ldots, 6\}$, multiplication of $\alpha \in \mathbb{F}_p$ with elements in $\mathbb{E}$ amounts to performing cyclic shifting on (the binary representation) of α, which can be implemented very efficiently on hardware as well as in software. This is crucial to our goal of constructing a lightweight cryptosystem. In our proposed constructions, we will however use $p = 127$ and $\mathbb{E} = \{(-2)^i, i = 0, \ldots, 13\}$, which share the same nice implementation properties.

However, one needs to be careful whether such a particular form of the secret key compromises the security of RSDP. Our case resembles the situation in the very well-studied learning problem *Learning with Errors* (LWE) [32]. In their seminal work, Applebaum et al. [33] shed light on a crucial observation: for an arbitrary noise distribution, the Learning with Error problem, where the secret's values are sampled from the same distribution as the noise, is **as hard as** the case where the secret is taken uniformly at random. Let us denote ORSDP_k^* by the problem of distinguishing the samples of a $\mathcal{O}_{\mathbf{s}}^{\mathsf{RSDP}}$ where values of $\mathbf{s}$ are drawn according to the noise distribution. By applying the same standard reduction as in [33], we directly achieve that ORSDP_k^* is as hard as ORSDP_k.

2.3. RSDP-solving Algorithms

With the introduction of RSDP, adaptations of SDP-solving algorithms have been proposed [16], [17], [34], [35] in the new setting. In particular, we briefly investigate the Stern/Dumer variant, the BJMM variant, and the algebraic approach, which we will use to derive security parameters for our construction.

Stern/Dumer algorithm. The following is a straightforward adaptation of the Stern/Dumer algorithm [36], originally proposed for the classical SDP. We recall the SDP instance given by $(\mathbf{H}, \mathbf{s}, \mathbb{E})$, where $\mathbf{H} \in \mathbb{F}_p^{(n-k)\times n}$ and $\mathbb{E}$ is a multiplicative subgroup of order z in $\mathbb{F}_p$. Our task is to find $\mathbf{e}\mathbf{H}^\intercal = \mathbf{s}$, where $\mathbf{e} \in \mathbb{E}^n$. To ease the notation, we omit the transposition notation from now on. First, one applies to the parity-check matrix a partial Gaussian Elimination $\mathbf{H} \leftarrow \mathbf{UHP}$, with an invertible matrix $\mathbf{U}$ and some permutation $\mathbf{P}$, resulting in

$$(\mathbf{e}_1, \mathbf{e}_2)\begin{pmatrix}\mathbf{I}_{n-k-\ell} & \mathbf{H}_1 \\ \mathbf{0} & \mathbf{H}_2\end{pmatrix} = (\mathbf{s}_1, \mathbf{s}_2),$$

where $\mathbf{e}\mathbf{P}^{-1} = (\mathbf{e}_1, \mathbf{e}_2) \in \mathbb{F}_p^{k+\ell} \times \mathbb{F}_p^{k+\ell}$ (with an additional later-to-be-optimized parameter ℓ). The equation gives rise to the identities

$$\mathbf{e}_1 + \mathbf{e}_2\mathbf{H}_1 = \mathbf{s}_1 \tag{1}$$

$$\mathbf{e}_2\mathbf{H}_2 = \mathbf{s}_2 \tag{2}$$

The next step of the Stern/Dumer approach is to enumerate $\mathbf{e}_2 \in \mathbb{E}^{k+\ell}$ to have candidates for (2). Equation (1) is then used to check if the candidates are the solution(s), that is, $\mathbf{s}_1 - \mathbf{e}\mathbf{H}_1$ is also a $\mathbb{E}^{k+\ell}$ vector. The enumeration step can be done via a Meet-in-the-Middle approach. In particular, we further parse $\mathbf{e}_2 = (\mathbf{x}_1, \mathbf{x}_2)$ where $\mathbf{x}_i \in \mathbb{E}^{(k+\ell)/2}$ and $\mathbf{H}_2 = \begin{pmatrix}\mathbf{H}_{21} \\ \mathbf{H}_{22}\end{pmatrix}$. We then construct two lists

$$\mathcal{L}_1 = \{(\mathbf{x}_1, \mathbf{x}_1\mathbf{H}_{21}), \mathbf{x}_1 \in \mathbb{E}^{(k+\ell)/2}\},$$

$$\mathcal{L}_2 = \{(\mathbf{x}_2, \mathbf{s}_2 - \mathbf{x}_2\mathbf{H}_{22}), \mathbf{x}_2 \in \mathbb{E}^{(k+\ell)/2}\},$$

and find collisions between them. For simplicity, we can assume $|\mathcal{L}_1| = |\mathcal{L}_2| = z^{(k+\ell)/2}$, and on average, there are $|\mathcal{L}_1|^2 \cdot p^{-\ell}$ collisions between them. The complexity of this step consists of building the lists and checking for solutions. In particular,

$$C_{\mathsf{Stern/Dumer}} = \frac{C_1 + C_2 + C_{\text{coll}}}{(\mathsf{number.of.solutions})} \times (\mathsf{memory.access.cost}), \tag{3}$$

where

$$C_1 = C_2 = |\mathcal{L}_1| \cdot \left(\frac{k+\ell}{2} \cdot \log(z) + \ell \cdot \log(p)\right),$$

$$C_{\text{coll}} = |\mathcal{L}_1|^2 \cdot p^{-\ell} \cdot (k+\ell) \cdot \log(p).$$

The number of solutions is given by

$$1 + |\{\mathbf{e} \in \mathbb{E}^n : \mathbf{e}\mathbf{H} = \mathbf{s}\}| = 1 + z^n p^{k-n}.$$

BJMM algorithm. Similar to the BJMM algorithm [37] in classical SDP, we can also adapt the *representation technique* in the enumeration step. This has been investigated in [34] and [17]. Interestingly, the performance of such adaptations depends on the additive structure of the restricted set $\mathbb{E}$. We will briefly explain the idea and skip the details of the approach. For a more thorough understanding, we refer the readers to [17].

Simplistically speaking, the representation technique aims to construct $\mathbf{e}_2 = \mathbf{e}_1^{(1)} + \mathbf{e}_2^{(1)}$ that solves equation (2). Since $\mathbb{E}$ is only a multiplicative subgroup, it is usually not closed under addition. Therefore, we have to sample $\mathbf{e}_i^{(1)}$ from some larger search domain defined as $\mathbb{E} \cup \mathbb{D} \cup \{0\}$ for some carefully chosen $\mathbb{D}$. The cost of this approach depends on the so-called "linearity" of $\mathbb{E}$ and $\mathbb{D}$. In particular, for a random element $a \in \mathbb{E}$, we define the two quantities

$$\ell_{\mathbb{E}}(a) = |[b \in \mathbb{E} : \exists c \in \mathbb{E}, b + c = a]|,$$

$$\ell_{\mathbb{D}}(a) = |[b \in \mathbb{E} : \exists c \in \mathbb{D}, b + c = a]|.$$

These values will determine how many representations r for an element during the enumerating process; hence, it implies how much we can save (by only enumerating a $1/r$-fraction). For the choice of p and $\mathbb{E}$ in CROSS, the above quantities are independent of a. Similar to the SDP case, one can repeat this step for $\mathbf{e}^{(1)} = \mathbf{e}_1^{(2)} + \mathbf{e}_2^{(2)}$ and so on, increasing the depth of the tree. However, for a fixed code rate of k/n and in the full-weight RSDP setting (as in CROSS parameters [17]), this approach does not seem to yield improvements over the Stern/Dumer algorithm.

A clever alternative was proposed by the same authors: one can solve for a *shifted* instance of RSDP as $(\mathbf{e} + \mathbf{x})\mathbf{H} = \mathbf{s} + \mathbf{s}_{\mathbf{x}}$, where $\mathbf{s}_{\mathbf{x}}$ is the syndrome of a well-chosen $\mathbf{x}$ (e.g., $\mathbf{x} \in \mathbb{E}^n$). Such a transformation introduces 0 to the shifted error, from which BJMM might benefit. In particular, it was found that for the CROSS parameter, shifted-BJMM yields smaller spaces *shifted* $\mathbb{E}_x$ and $\mathbb{D}_x$ (while having the same linearity), which slightly outperforms Stern/Dumer.

However, as we will see in our setting and parameters selection, applying shifted-BJMM is very tricky. For example, if $\mathbb{E} := \{\pm 2^i\}$, then the linearity of shifted $\mathbb{E}$ and $\mathbb{D}$ are not constant. Nevertheless, we can conservatively lower bound the cost of shifted-BJMM by testing with different possible $\ell_{\mathbb{E}}$ and $\ell_{\mathbb{D}}$.

Both Stern/Dumer and BJMM enumerating approach requires very high memory usage (e.g., 2^{141} and 2^{116}, respectively, for NIST Category 1 parameters [17]). Therefore, it is reasonable to take into account certain memory cost models. Similar to the CROSS authors, we use log model to estimate the cost of memory access. For example, in Equation (3), the complexity is multiplied with $\log(\max(|\mathcal{L}_1|, |\mathcal{L}_2|))$.

Algebraic approach. The basic Aurora-Ge approach [31] involves forming nonlinear equations for each sample of the form

$$\prod_{e\in\mathbb{E}}(\langle \mathbf{a}, \mathbf{s}\rangle + e - b) = 0,$$

which will be of degree z. Then, the system of equations is transformed into a linear one with around $\sum_{i=1}^{z}\binom{k}{i}$ unknowns (each is a monomial up to degree z of variables s_i in $\mathbf{s}$). Assuming in the order of $\binom{k}{z}$ oracle calls and $z < k/2$, one can then solve for the unknowns by Gaussian elimination in time around $\binom{k}{z}^3$. Better algorithms can lower the exponent.

More advanced algebraic attacks have been recently studied as an alternative to the traditional ISD algorithms, particularly for specific variants of SDP such as the Regular SDP [38]. In [35], the authors investigated the prospect of applying algebraic attacks to various other variants, one of which is RSDP. Within this framework, SDP is modeled as solving a polynomial system. A general approach is to compute the Gröbner basis for the system, whose complexity depends on estimating the *degree of regularity* of said system. Algebraic approaches based on Gröbner basis techniques improve performance, but the improvement is often difficult to estimate.

2.4. The HB-family of authentication protocols

The HB-family authentication protocols were pioneered by Hopper and Blum [1] to achieve a simple yet secure protocol based on the hardness of the learning problem. Since our design takes inspiration from HB-family protocols, especially the HB and the HB+ protocols, we briefly revisit the LPN problem, which gives rise to the aforementioned constructions.

Problem 3 (Learning Parity with Noise (LPN) Problem). Let Ber_τ be the Bernoulli distribution over $\mathbb{Z}_2$ with bias $\tau \in \left(0, \frac{1}{2}\right)$. That is, a random variable $x \leftarrow \mathsf{Ber}_\tau$ if $\Pr[x = 1] = \tau$. Given a secret $\mathbf{s} \in \mathbb{Z}_2^k$, the LPN Oracle $\mathcal{O}_{\mathbf{s},\tau}^{LPN}$ returns pairs in $\mathbb{Z}_2^k \times \mathbb{Z}_2$ of the form

$$\{\mathbf{a} \xleftarrow{\$} \mathbb{Z}_2^k, z = \langle \mathbf{a}, \mathbf{s}\rangle \oplus e\},$$

where $e \leftarrow \mathsf{Ber}_\tau$. The (decisional) $\mathsf{LPN}_{\tau,k}$ problem is defined as distinguishing the samples obtained from the above $\mathcal{O}_{\mathbf{s},\tau}^{LPN}$ from the uniform distribution.

The HB protocol. We recall the HB protocol proposed by Hopper and Blum [1]. The Tag and the Reader share a binary length-k secret $\mathbf{s}$. One round of the protocol is illustrated in Figure 1. It is an simple 2-round interaction between $\mathcal{T}$ and $\mathcal{R}$. First, a "challenge" $\mathbf{a}$ is send by $\mathcal{R}$ and $\mathcal{T}$ responds with $z = \langle \mathbf{a}, \mathbf{s}\rangle \oplus e$ where $e \leftarrow \mathsf{Ber}_\tau$. Finally, $\mathcal{R}$ verifies where $z = \langle \mathbf{a}, \mathbf{s}\rangle$. On average, the check passes with probability $(1-\tau)$, and a random response from an illegitimate tag is accepted with probability $\frac{1}{2}$. Therefore, to raise the confidence that $\mathcal{T}$ has the secret key, one repeats this protocol for n times. Therefore, $\mathcal{R}$ outputs accept if there are at most $\tau \cdot n$ failed checks[2].

The HB+ protocol. It was later shown by Juels and Weis [2] that HB is only secure against a passive attacker. However, against a (more realistic) active adversary, the protocol is easily compromised. In particular, such an attacker $\mathcal{A}$ in Phase 1 can repeatedly challenge $\mathcal{T}_{\mathbf{s}}$ with the same $\mathbf{a}$. Since e is sampled according to Ber_τ, $\mathcal{A}$ can perform majority voting (after enough queries) to reveal the noise-free value $\langle \mathbf{a}, \mathbf{s}\rangle$. Repeating this process with linearly independent challenges $\mathbf{a}_i$, $\mathcal{A}$ can recover $\mathbf{s}$ by Gaussian Elimination. Hence, HB+ was proposed by Juels and Weis as an augmented version of HB. They turned HB into a 3-round interaction by requiring $\mathcal{T}$ to sends a "blinding" factor $\mathbf{b} \xleftarrow{\$} \mathbb{Z}_2^k$ before $\mathcal{R}$'s challenge. Moreover, there are now two secret keys $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^k$. One round of the HB+ protocol is shown in Figure 2. Similarly, $\mathcal{R}$ outputs accept after n rounds if the number of failed checks is, at most, up to a certain threshold.

3. An authentication protocol based on RSDP

In this section, we propose a simple novel authentication protocol that is secure in the active model. In essence, it is the HB+ protocol, but the LPN problem has been replaced by the RSDP problem. In particular, we deploy the full-weight version of RSDP that has also been used recently in CROSS [17].

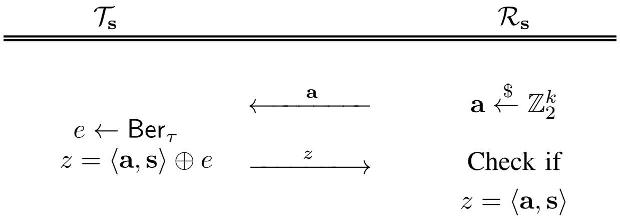
Public parameters. The following are public parameters where n, k, p and z depends on the security parameter λ.

- $\mathbb{F}_p$: integers modulo prime p.
- $\mathbb{E} = \{g^i, i = 0, \ldots, z-1\}$: the multiplicative subgroup of order z in $\mathbb{Z}_p$.
- $k \in \mathbb{N}$: length of the secret keys.
- $n \in \mathbb{N}$: number of rounds in the authentication protocol.

Secret keys. The Tag and the Reader share two secret keys $\mathbf{x}, \mathbf{y} \in \mathbb{E}^k$. Figure 3 describes one round in the authentication protocol with $\mathcal{T}_{\mathbf{x},\mathbf{y}}$ and $\mathcal{R}_{\mathbf{x},\mathbf{y}}$. Similar to the HB+ protocol, it consists of three interactions between the Tag and the Reader. First, $\mathcal{T}_{\mathbf{x},\mathbf{y}}$ sends a "blinding" factor $\mathbf{b}$ and $\mathcal{R}_{\mathbf{x},\mathbf{y}}$ responds with a challenge $\mathbf{a}$. Then, $\mathcal{T}_{\mathbf{x},\mathbf{y}}$ samples an element $e \in \mathbb{E}$ at uniform random and compute $u = \langle \mathbf{a}, \mathbf{x}\rangle + \langle \mathbf{b}, \mathbf{y}\rangle + e \bmod p$, which is then relayed to $\mathcal{R}_{\mathbf{x},\mathbf{y}}$. The Reader performs a check if $u - (\langle \mathbf{a}, \mathbf{x}\rangle + \langle \mathbf{b}, \mathbf{y}\rangle) \in \mathbb{E}$. The round is then repeated n times to raise the confidence in the Tag. The Reader accepts if and only if all n checks were fulfilled.

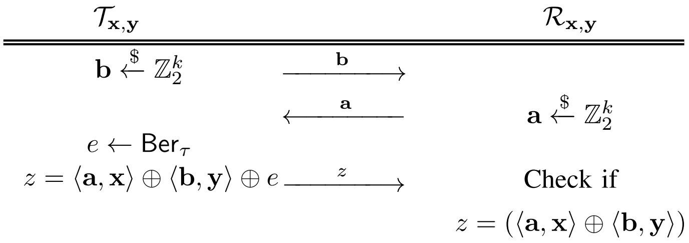
Remark 1. An active Adversary can repeatedly query with a challenge $\mathbf{a}$, e.g., $\mathbf{a} = \mathbf{0}$, trying to solve recover $\mathbf{y}$ as

[2] In [2], a higher threshold τ' can be chosen to minimize the number of rounds needed to obtain $P_c \leq 2^{-40}$ and $P_s \leq 2^{-80}$.



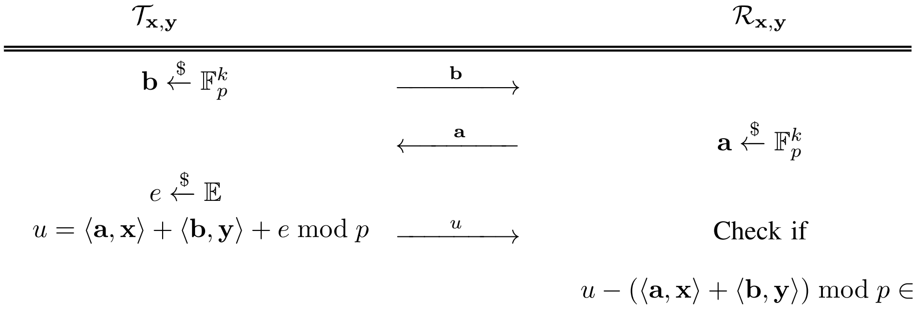
$\mathcal{T}_{\mathbf{s}}$
$\mathcal{R}_{\mathbf{s}}$
$\mathbf{a}$
$\mathbf{a} \overset{\$}{\leftarrow} \mathbb{Z}_2^k$
$e \leftarrow \mathsf{Ber}_\tau$
$z = \langle \mathbf{a}, \mathbf{s} \rangle \oplus e$
z
Check if
$z = \langle \mathbf{a}, \mathbf{s} \rangle$

Figure 1: One round of the HB authentication protocol.



$\mathcal{T}_{\mathbf{x},\mathbf{y}}$
$\mathcal{R}_{\mathbf{x},\mathbf{y}}$
$\mathbf{b} \overset{\$}{\leftarrow} \mathbb{Z}_2^k$
$\mathbf{b}$
$\mathbf{a}$
$\mathbf{a} \overset{\$}{\leftarrow} \mathbb{Z}_2^k$
$e \leftarrow \mathsf{Ber}_\tau$
$z = \langle \mathbf{a}, \mathbf{x} \rangle \oplus \langle \mathbf{b}, \mathbf{y} \rangle \oplus e$
z
Check if
$z = (\langle \mathbf{a}, \mathbf{x} \rangle \oplus \langle \mathbf{b}, \mathbf{y} \rangle)$

Figure 2: The HB+ authentication protocol.



$\mathcal{T}_{\mathbf{x},\mathbf{y}}$
$\mathcal{R}_{\mathbf{x},\mathbf{y}}$
$\mathbf{b} \overset{\$}{\leftarrow} \mathbb{F}_p^k$
$\mathbf{b}$
$\mathbf{a}$
$\mathbf{a} \overset{\$}{\leftarrow} \mathbb{F}_p^k$
$e \overset{\$}{\leftarrow} \mathbb{E}$
$u = \langle \mathbf{a}, \mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{y} \rangle + e \bmod p$
u
Check if
$u - (\langle \mathbf{a}, \mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{y} \rangle) \bmod p \in \mathbb{E}$

Figure 3: One round of the RSDP-based HB+ authentication protocol. For simplicity, we let $k_{\mathbf{x}} = k_{\mathbf{y}} = k$.

the other key $\mathbf{x}$ no longer contributes to the response z. Therefore, the protocol security relies on the length of $\mathbf{y}$, and we only need to set the length of $\mathbf{x}$ so that guessing is infeasible. For example, 80-bit security requires the length of $\mathbf{x}$ to be at least 22 as $z^{22} \approx 2^{83}$. The protocol can be instantiated with different key lengths, denoted by $k_{\mathbf{x}}, k_{\mathbf{y}}$. For simplicity, in Figure 3 and the proof, we use the same key length.

Completeness of the authentication protocol. Contrary to HB-family authentication protocol, a legitimate tag in our proposal will produce $\mathbf{u}$ that passes the check of $\mathcal{R}_{\mathbf{x},\mathbf{y}}$ with probability 1.

Soundness of the authentication protocol. Let u_i be a random response in round i. To be authenticated, all $u_i, i = 1, \ldots, n$ need to pass the Reader's checks. Without knowing the secret, such a single event happens with a probability

$$P_s = (\Pr[u_i - (\langle \mathbf{a}, \mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{y} \rangle) \bmod p \in \mathbb{E}])^n = \left(\frac{z}{p}\right)^n.$$

3.1. Security against an active adversary

In this section, we show a reduction of the RSDP problem to the authentication protocol in the active model. To formalize an active attacker, we introduce the following notation. Let $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ be the active adversary in Phase 1 (querying the tag) and Phase 2 (authenticate to the reader), respectively. In phase 1, $\mathcal{A}$ has access to $\mathcal{T}$ in at most Q authenticating executions, and the adversary's actions are characterized by the program $\mathcal{A}_1$. During each execution, it receives $\mathbf{b}_i, i = 1, \ldots, n$ from $\mathcal{T}$, and sends back $\mathbf{a}_i$, for $i = 1, \ldots, n$ respectively, as challenges. Then, $\mathcal{T}$ computes some z_i for $i = 1, \ldots, n$ as responses. In the end, $\mathcal{A}_1$ outputs some state σ that contains all information used for the next phase.

In the second phase, the attacking adversary impersonates the tag $\mathcal{T}$. Its complete action is described by the program $\mathcal{A}_2$, which takes the state σ as input. $\mathcal{A}_2$ sends some $\widehat{\mathbf{b}}_i$ (which is derived from σ), $\mathcal{R}$ then challenges with $\widehat{\mathbf{a}}_i$ and $\mathcal{A}_2$ provides the final $\widehat{z}_i$, attempting to pass the authentication protocol. After n rounds, $i = 1, \ldots, n$, the reader decides to accept or reject. The running time of $\mathcal{A}$, denoted by t, is determined by the maximum run time between $\mathcal{A}_1$ and $\mathcal{A}_2$.

We need to introduce further notation related to the error set $\mathbb{E}$. Let $\mathbb{D}$ denote the set

$$\mathbb{D} := \{e - e' | e, e' \in \mathbb{E}\}.$$

Furthermore, let $\alpha_{\mathbb{E}} := \frac{|\mathbb{D}|}{p}$.

Theorem 1. Assume that the RSDP-based HB+ protocol in Figure 3(with parameters p, k, n, and $\mathbb{E} = \{g^i, i =$

$0,\dots,z{-}1\}$) is not (t,Q,ϵ)-secure, that is, there exists an active adversary $\mathcal{A}$, running in time t, interacting with the tag in at most Q executions, and achieving success probability at least ϵ in Phase 2. Then there exists a distinguisher D running in time $\mathcal{O}(t/\epsilon)$, making $Q\cdot n$ queries to an RSDP oracle $\mathcal{O}^{\mathsf{RSDP}}_{\mathbf{s}}$, for which

$$\left|\Pr\left[\mathbf{s}\leftarrow\mathbb{E}^k : D^{\Lambda_k(\mathbf{s})}=1\right]-\Pr\left[D^{U_{k+1}}=1\right]\right|\geq 1-\epsilon',$$

with

$$\epsilon'\approx\alpha^n_{\mathbb{E}}\cdot c^2/\epsilon^2+(c+1)\cdot\exp(-c),$$

for some small constant c.

Proof: Let $(\mathbf{b}_i,u'_i)\in\mathbb{Z}_p^k\times\mathbb{Z}_p$ be ordered pairs that are from an unknown distribution, which is either $\Lambda_k(\mathbf{s})$ or U_{k+1}. Assume that D has access to the program description $\mathcal{A}=(\mathcal{A}_1,\mathcal{A}_2)$ of an adversary $\mathcal{A}$ that can successfully attack the protocol described in Figure 3, in time t and after Q queries, with advantage ϵ. We now show that D can use $\mathcal{A}$ to distinguish whether its samples are from an RSDP oracle or drawn uniformly at random.

The following steps take inspiration from the reduction for the HB+ protocol in [39]. The algorithm D first picks a random $\mathbf{y}\overset{\$}{\leftarrow}\mathbb{E}^k$. The idea is that D simulates a transcript as if, in the query phase, $\mathcal{A}$ is interacting with an honest tag $\mathcal{T}_{\mathbf{y},\mathbf{s}}$.

Phase 1. D runs $\mathcal{A}_1$. Recall that, in this phase $\mathcal{A}_1$ challenges the Tag with $\mathbf{a}_i$ after receiving some $\mathbf{b}_i$. First, D relays $\mathbf{b}_i$ to $\mathcal{A}_1$ as the blinding factor. For every challenge $\mathbf{a}_i$ made by $\mathcal{A}_1$, D computes and responds with

$$u_i=u'_i+\langle\mathbf{a}_i,\mathbf{y}\rangle \bmod p,$$

for $i=1,\dots,n$. After Q such executions, $\mathcal{A}_1$ outputs the state σ.

Phase 2. Next, D will use the $\mathcal{A}_2$ program. $\mathcal{A}_2(\sigma)$ starts by sending $\widehat{\mathbf{b}}_i$ to D, who then responds with $\widehat{\mathbf{a}}^1_i, i=1,\dots,n$. Then, $\mathcal{A}_2(\sigma)$ computes some responses u^1_i for $i=1,\dots,n$. The distinguisher D rewinds $\mathcal{A}_2(\sigma)$ but this time inputs a different $\widehat{\mathbf{a}}^2_i$ and observes the responses u^2_i for $i=1,\dots,n$.

The distinguisher D proceeds to rewind and run $\mathcal{A}_2$ for M times and inputs different random $\widehat{\mathbf{a}}^j_i, i=1,\dots,n$, and $j=1,\dots,M$ and obtains u^j_i for $i=1,\dots,n$ and $j=1,\dots,M$.

Finally, D runs through all $1\leq j_1<j_2\leq M$ and computes $\widehat{u}_i=u^{j_1}_i-u^{j_2}_i \bmod p$ for $i=1,\dots,n$ and $\widehat{u}'_i=\langle(\widehat{\mathbf{a}}^{j_1}_i-\widehat{\mathbf{a}}^{j_2}_i),\mathbf{y}\rangle$ for $i=1,\dots,n$. D checks if there exist two different such j values, j_1,j_2 for which we have that for all $i=1,\dots n$, $\exists e_i,e'_i\in\mathbb{E}$ such that $e_i-e'_i=\widehat{u}_i-\widehat{u}'_i$. If this is the case, D outputs 1 ($\Lambda_k(\mathbf{s})$), otherwise 0 (U_{k+1}).

- Case 1: Assume D gets samples from U_{k+1} in Phase 1. Then, the values $\widehat{u}_i-\widehat{u}'_i$ are uniformly and independently distributed. That is, D will make an error and output 1 iff $\widehat{U}_i=\widehat{u}_i-\widehat{u}'_i$ can be written as the subtraction of two elements in $\mathbb{E}$. This happens iff $\widehat{U}_i$ is in the set $\mathbb{D}$. A pair of j_1,j_2 then gives all n responses from a legit tag with probability $\alpha^n_{\mathbb{E}}:=\left(\frac{|\mathbb{D}|}{p}\right)^n$. In other words, running through all pairs gives an error probability of D as most as $\alpha_{\mathbb{E}}\cdot\binom{M}{2}$. In other word,
$$\Pr\left[D^{U_{k+1}}=1\right]=\alpha^n_{\mathbb{E}}\cdot\binom{M}{2}.$$
- Case 2: Assume D gets samples from $\Lambda_k(\mathbf{s})$. Then, in Phase 1, we simulated the transcript as if $\mathcal{A}$ interacted with $\mathcal{T}_{\mathbf{y},\mathbf{s}}$. Indeed,
$$u_i=u'_i+\langle\mathbf{a}_i,\mathbf{y}\rangle \bmod p=\langle\mathbf{b}_i,\mathbf{s}\rangle+\langle\mathbf{a}_i,\mathbf{y}\rangle+e_i.$$
By definition, $\mathcal{A}_2$ produces a correct response with probability at least ϵ. We now show that for a well-chosen M, at least a pair of correct responses exists with high probability. $\mathcal{A}$ fails to produce any correct pair of responses with probability
$$M\cdot\epsilon\cdot(1-\epsilon)^{M-1}+(1-\epsilon)^M\approx$$
$$M\cdot\epsilon\cdot\exp(-\epsilon\cdot(M-1))+\exp(-\epsilon\cdot M).$$
That is,
$$\Pr\left[\mathbf{s}\leftarrow\mathbb{E}^k : D^{\Lambda_k(\mathbf{s})}=1\right]=$$
$$1-M\cdot\epsilon\cdot\exp(-\epsilon\cdot(M-1))+\exp(-\epsilon\cdot M).$$
We want the probability of D being successful to be close to 1; hence, one chooses $M=c\cdot\frac{1}{\epsilon}$, for some constant $c>1$. This proves the theorem. Hence, if c is big enough, D can tell if the samples are from $\Lambda_k(\mathbf{s})$ with probability very close to 1.
Example 1. Let $\epsilon=2^{-20}$, we only need $M=\frac{5}{\epsilon}$ to have probability 0.96 to have at least one correct pair.

□

3.2. A MitM Attack Strategy

Here, we point out that the proposed protocol from Figure 3 is not secure against a MitM attacker. The attack depends on the additive structure of the restricted set $\mathbb{E}$. Consider an adversary that observes one sample in the first round of the protocol as before:

$$u=\langle\mathbf{a},\mathbf{x}\rangle+\langle\mathbf{b},\mathbf{y}\rangle+e \bmod p$$

and tries to guess $e\in\mathbb{E}$. The adversary $\mathcal{A}$ first picks an element $e'\in\mathbb{E}$ and assumes that $e=e'$. Then pick another element $e''\neq e'$, such that $e''\in\mathbb{E}$. Then substitute u with u', which is calculated as

$$u'=\langle\mathbf{a},\mathbf{x}\rangle+\langle\mathbf{b},\mathbf{y}\rangle+e-e'+e'' \bmod p$$

in this first round. In all other rounds $i=2,\dots,n$, the adversary forwards the responses u without modifications.

Assume u' passes the verification check, which can only occur when $e-e'+e''\in\mathbb{E}$, and depending on the choice of $\mathbb{E}$, $\mathcal{A}$ can obtain some information regarding e or even correctly guess e.

Example 2. Let $p=127$ and $\mathbb{E}=\{1,2,4,8,16,32,64\}$. It is then easily checked that if $e'\neq e''\in\mathbb{E}$ and $e-e'+e''\in\mathbb{E}$ then $e=e'$. Equivalently, when u' passes the check, $\mathcal{A}$ correctly guess $e=e'$. Knowing e, the adversary can compute $u-e$ and conclude that

$$u-e=\langle\mathbf{a},\mathbf{x}\rangle+\langle\mathbf{b},\mathbf{y}\rangle \bmod p.$$

This gives a linear equation in the unknown key variables $\mathbf{x}, \mathbf{y}$. The Reader accepts the corrupted response with probability $1/|\mathbb{E}|$ (for each authentication). Thus, it takes the attacker on average $|\mathbb{E}|$ attempts to reveal a noise-free value $\langle \mathbf{a}, \mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{y} \rangle$, or at most $(n+k)|\mathbb{E}|$ attempts to reveal the entire secret key by solving a system of linear equations through Gaussian elimination.

Example 3. For our later choice of $\mathbb{E}$, $p = 127$ and $\mathbb{E} = \{(-2)^i, i = 0, \ldots, 13\}$. Guessing e is more challenging as there are many e' and e'' that yield a seemingly valid u'. However, by definition, the design is not MitM-secure as it means that $\mathcal{A}$ can forge a response $\mathbf{u}' \neq \mathbf{u}$ (with non-negligible probability) that will be authenticated by $\mathcal{R}$.
Going further, $\mathcal{A}$ can also perform a key recovery attack. In contrast to Example 2, for fixed e' and e'', there are more than one possible values for e (that makes u' pass). As an example, $(e', e'') = (1, 8)$ yields $e = 1$ or $e = 19$, which allows $\mathcal{A}$ to form a quadratic equation as $(e-1)(e-8) = 0 \bmod p$. By changing different (e', e'') and going through n responses coming from the Tag, $\mathcal{A}$ obtains more information (and equations) about each error position, which could be used to speed up combinatoric solvers (or algebraic).

4. Parameters for the proposed protocol

W propose parameters for 3 security levels, namely $80, 112$, and 128 bits. The parameters are selected based on the performance of the various RSDP-solving algorithms.

As in Remark 1, an active attacker using RSDP solvers to recover the secret key(s) amounts to solving an RSDP instance with parameter n and secret length $k_{\mathbf{y}}$. When $n < k_{\mathbf{y}}$, there are exponentially many solutions. However, one has to find the exact $e_1, \ldots, e_n$ produced by $\mathcal{T}$ to recover $\mathbf{x}$. The attacker can indeed observe many authentications to get more than n samples to guarantee that the solution from the solver is indeed correct. However, once the average number of solutions is 1, Stern/Dumer and BJMM do not seem to take advantage of the extra samples.

We stress that the performance of Stern/Dumer or BJMM takes into account the cost of accessing huge memory under the log model. Indeed, such algorithms require such high memory usage that can not be disregarded. For example, Shifted-BJMM needs 2^{70} bits of memory 80-bit security parameter.

All proposed instances use $p = 127$ and $\mathbb{E} = \{(-2)^i, i = 0, \ldots 13\}$ as the multiplicative subgroup in $\mathbb{F}_p$, so $z = 14$.

We present the cryptanalysis on the 80-bit security parameters with $\mathbb{E} = \{\pm 1, \pm 2, \ldots, \pm 64\}$ Interestingly, the biggest threat to this parameter set is the algebraic attack. Since the algebraic approach cost can be lower-bounded by roughly $\binom{k}{z}^3$, it prevents $k_{\mathbf{y}}$ from being too small. Since there are better methods compared to Gaussian elimination that can lower the exponent, we conservatively choose $k_{\mathbf{y}} = 34$ as $\binom{34}{14}^3 \approx 2^{91}$. For a d-bit security level, we require the soundness error $\left(\frac{14}{127}\right)^n \leq 2^{-d}$. For example, $n = 26$ for 80-bit security level.

TABLE 1: Parameters recommendation for the RSDP HB+ protocol with n as the number of authentication steps, $k_{\mathbf{x}}$, $k_{\mathbf{y}}$ as the length of two secret keys, and p as the field size. The restricted set is set to be $\mathbb{E} := \{(-2)^i, i = 0, \ldots z-1\}$.

Security Level	80	112	128
(p, z)	(127,14)	(127,14)	(127,14)
$(k_{\mathbf{x}}, k_{\mathbf{y}})$	(22,34)	(30,54)	(34,70)
n	26	36	41

Now we consider the combinatorial approach, such as Stern and BJMM, assuming an adversary using such solvers is allowed multiple interactions (to guarantee the correctness of the found solution). In that case, it does not seem to yield improvements as the size of $\mathbb{E}$ is quite big, thus significantly increasing enumerating efforts. For example, the Stern(BJMM) in Section 2.3 requires 2^{105} (2^{110}, resp.) operations and 2^{94} (2^{98}, resp.) bits in memory. In contrast to the CROSS parameter, shifted-BJMM does not improve over Stern for our choice of $\mathbb{E}$. In addition, for higher security-level parameters, the gap between algebraic and combinatorial solvers gets significantly wider.

4.1. Performance

In this section, we discuss the proposed protocol with different benchmarks such as computation, communication, and key size(s).

Key size. The protocol employs two keys of lengths $k_{\mathbf{x}}, k_{\mathbf{y}}$ each, of which entries are drawn from the restricted set $\mathbb{E}$; therefore, we need roughly $(k_{\mathbf{x}} + k_{\mathbf{y}}) \cdot \log(z)$ bits to store the keys.

Communication. One authentication consists of n 3-round authenticating step. In each step, the Tag and the Reader exchange $(\mathbf{a}, \mathbf{b}, u) \in \mathbb{F}_p^k \times \mathbb{F}_p^k \times \mathbb{F}_p$, which incurs $n \times\ (2k+1) \times \log(p)$ bits in communication.

Example 4. 80-bit security parameters of LPN-based HB+ are roughly $(512, 1/8)$ where 512 is the length of the key and $1/8$ is the noise rate. To achieve good completeness and soundness probability (2^{-40} and 2^{-80}, respectively), one needs to repeat an authentication step by roughly $n = 441$ times.

TABLE 2: Key size of (in bits) different protocols.

Security Level	80	112	128
RSA	1024	2048	3072
DSA	1024	2048	3072
LPN-based	512	768	768
RSDP HB+	217	320	396

Table 2 shows the key sizes of our protocol in comparison with some traditional cryptographic primitives and LPN-based protocols. Note that there are several options for LPN-based protocols to achieve one security level. For instance, one can select $(k = 512, \tau = 0.49)$ for 112-bit security. However, such a high noise level will be detrimental to both the completeness and soundness of the scheme, thus implying a very high number of authentication steps, i.e., high communication cost. Therefore, for LPN-based protocols, we pick values of k that correspond to “reasonable” noise rates (typically ≤ 0.3).[3]

4.2. Hardware Implementation

The most critical task that affects performance in Figure 3 is computing

$$u = \langle \mathbf{a}, \mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{y} \rangle + e \bmod p$$

This operation is implemented using hardware in the tag and needs to be as lightweight as possible. Our choice of the error set $\mathbb{E}$, and the choice of p ensures that the operation remains cheap. We implemented two strategies for both protocols based on RSDP and LPN, with parameters targeting 80-bit security. The implementations focus on the inner product operation, which is the most costly operation implemented on the tag hardware. In both strategies, the cost is dominated by the circuits needed to store the key and the challenge. In the case of LPN, both variables consist of 512 bits, while in the case of RSDP, the key consists of 136 bits, while the challenge consists of 238 bits. In the case of LPN, one vector coordinate consists of 1 bit, while in the case of RSDP, one challenge vector coordinate consists of 7 bits, and 1 key vector coordinate consists of 4 bits in a signed integer format. Each multiplication, in the case of RSDP, consists of two steps:

1) Multiplication of the challenge coordinate by the unsigned value of the key coordinate $\bmod 127$, which boils down to a rotation based on the key values.
2) Multiplication by either 1 or $-1 \bmod 127$, which boils down to a conditional bitwise negation.

In strategy A, we implement the operation serially, as depicted in Figure 4. Every clock cycle, we perform one multiplication operation followed by an accumulation. The inner product operation takes several clock cycles. In strategy B, we implement the inner product operation in one clock cycle, and add all the products using an addition tree. Strategy A targets the minimum possible area (storage with a minimal combinational circuit), while strategy B targets the highest possible speed (1 clock cycle per inner product). We synthesized the circuits for both the Artix 7 FPGA using Xilinx Vivado and FDSOI 28nm using Synopsys Design Compiler, and the results are provided in Tables 3 and 4, respectively. We observe that on FPGA, strategy B works better for LPN than strategy A, but the RSDP-based protocol is smaller and faster for both strategies. Similar trends emerge in ASIC implementations, the RSDP circuit being 60% and 30% smaller for strategies A and B, respectively. We note that LPN favors strategy B over strategy A due to its very lightweight combinational logic compared to the storage cost, while RDSP offers a somewhat linear trade-off due to its more involved combinational/arithmetic logic. However, it still experiences a significant gain in all cases.

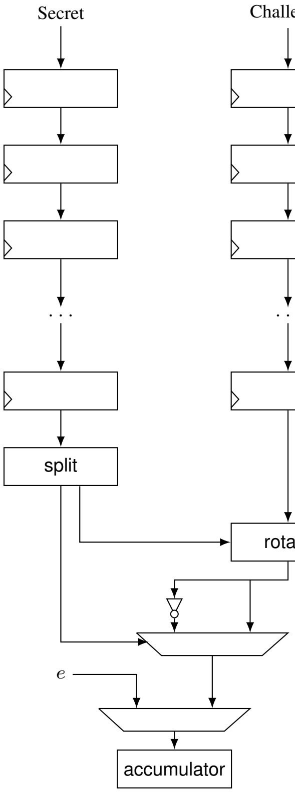


Figure 4: Serialized implementation of the inner product operation (strategy A).

TABLE 3: FPGA Resource Utilization for 80-bit security using Xilinx Artix 7 and Xilinx Vivado.

Design	LUTs	FFs	Cycles
LPN-based	1028	1025	512
	1032	1025	64
	780	1025	1
RSDP-based	399	381	34
	421	381	18
	678	266	1

TABLE 4: ASIC area for 80-bit security using FDSOI 28nm.

Design	GEs	Cycles
LPN-based	12243.75	512
	12262.5	64
	12985	1
RSDP-based	4678	34
	4797.18	18
	8614.69	1

5. A MitM-secured proposal based on RSDP

In this section, we present an extended protocol that provides security also in the MitM model. We adopt the approach of Lyubashevsky et al. [12] to achieve this, with the addition of a few additional assumptions.

Definition 4. A family function $\mathcal{F} : \mathbb{D} \to \mathbb{F}$ is said to be a *weak pseudorandom functions* (wPRFs) if for $f \xleftarrow{\$} \mathcal{F}$, it is computationally infeasible to distinguish input-output pairs $(x_i, f(x_i))$, where x_i are chosen

[3] Table in Section 5.2 in [40] can serve as a rough guideline for LPN parameters.

uniformly at random from $\mathbb{D}$, from random pairs $(x_i, y_i) \in (\mathbb{D}, \mathbb{F})$.

The condition of wPRFs can be further relaxed by a family of functions $\mathcal{F}_\chi$, provided the output of $f \in \mathcal{F}$ is indistinguishable after perturbation of *noise* characterized by a distribution χ. Such a family is called *randomized wPRFs*.

Definition 5. $\mathcal{H} : \mathbb{D} \to \mathbb{F}$ is a pairwise independent function family if

$$\Pr_{h \xleftarrow{\$} \mathcal{H}} [h(x_1) = y_1 \wedge h(x_2) = y_2] = 1/|\mathbb{F}|^2,$$

for all $x_1 \neq x_2$, $y_1 \neq y_2$.

In brevity, the design in [12] relies on a (randomized) wPRFs $\mathcal{F}_\chi : \mathbb{D} \to \mathbb{F}$, a pairwise independent function family $\mathcal{H} : \mathbb{D} \to \mathbb{F}$, where $\mathbb{F}$ is a finite field. Importantly, a weight function $|\cdot|$ (defined on $\mathbb{F}$), the field $\mathbb{F}$, and χ have to satisfy certain (reasonable) properties. In essence, together, they have to provide good completeness (close to 1) and soundness probability. A natural instantiation for LPN is $\mathbb{F} = \mathbb{Z}_2[x]/\langle g(x) \rangle$, for some irreducible polynomial $g(x)$, and χ as the Bernoulli distribution Ber_η^n, and the Hamming weight function. Despite such requirements not applying to our situation, we have seen in Section 3 that RSDP naturally yields good completeness and soundness probability. Hence, we propose a protocol that achieves MitM security in Figure 5.

The following instantiation of the protocol in Figure 5 can be seen as an RSDP version of the protocol in [12]. Note that the protocol now runs only a single round as the final response $\mathbf{u}$ from the Tag is a vector (or, equivalently, an element in a finite field).

Public parameters. The following are public parameters where k, n, p and z depends on the security parameter λ.

- $\mathbb{Z}_p$: integers modulo p.
- $\mathbb{E}$: a multiplicative subgroup of order z in $\mathbb{Z}_p$.
- $\mathbb{F} = \mathbb{Z}_p[x]/\langle g(x) \rangle$ for some irreducible polynomial $g(x)$ over $\mathbb{Z}_p$ of degree n. Multiplications, denoted by $\cdot$ in $\mathbb{F}$, are seen as polynomial multiplications. Operations between a vector and an element in $\mathbb{F}$ are assumed to be done after converting the vector to a corresponding element. For instance, a vector can be seen as coefficients of a polynomial in $\mathbb{F}$, and vice versa.

Secret keys. The Tag and the Reader both share secret keys $\mathbf{X} \in \mathbb{E}^{n \times k}, h \in \mathcal{H}$, and $s \in \mathbb{F}$. The function h is defined as $h(\mathbf{x}) = h_1 \cdot x + h_2$, for $h_i \in \mathbb{F}$ and x is the corresponding polynomial to vector $\mathbf{x}$ in $\mathbb{F}$.

Hardness assumptions. Having $\mathbf{X}$ as a secret matrix is equivalent to having many RSDP instances with different secrets $\mathbf{x}_i$ (columns of $\mathbf{X}$). However, as we have discussed, RSDP is not hard when an adversary $\mathcal{A}$ has access to an unlimited number of queries. Therefore, for the reduction proof of this design, we have to assume that $\mathcal{A}$ is allowed up to Q interactions with the Tag, where Q is a fixed value. Storing a matrix $\mathbf{X} \in \mathbb{E}^{k \times n}$ could pose a practical challenge for a low-cost RFID Tag. Therefore, similar to previous work, one can consider a version where $\mathbf{X}$ is a Toeplitz matrix.

5.1. Security reduction of the MitM proposal

We now want to prove security for the protocol in Figure 5. We are inspired by the methodology from [12]. Their work considers a slightly different MitM adversary where it does not operate in two phases. Instead, a MitM attacker interacts with $\mathcal{T}$ and $\mathcal{R}$ for Q times and modifies the communications $(\mathbf{b}, a, u) \to (\mathbf{b} + \mathbf{b}', a + a', u + u')$. The attacker wins the game if one out of Q interaction, a non-trivial change in the communication, i.e., $(\mathbf{b}', a', u')$ are not simultaneously 0, yields an accept from $\mathcal{R}$. However, if the attacker wins in this scenario, it also wins in the 2-stage model.

In this reduction, we assume the existence of an adversary $\mathcal{A}$ that breaks the design after Q authentication queries with probability ϵ. It means that we assume that $\mathcal{A}$ makes $Q - 1$ unsuccessful attempts followed by a Q-th query, where the Reader output accept for $(\mathbf{b}'_i, a'_i, u'_i) \neq (\mathbf{0}, 0, 0)$ with probability ϵ. Obviously, this implies that an adversary $\mathcal{A}$ that is allowed to win in any query has success probability at most ϵQ.

Theorem 2. Assume that the protocol in Figure 5 is not (t, Q, ϵ)-secured, that is, there exists an active adversary $\mathcal{A}$, running in time t, interacting with the Tag and Reader in at most Q executions, and can produce $(\mathbf{b}', a', u') \neq (\mathbf{0}, 0, 0)$ that is authenticated with probability at least ϵ. Then there exists an algorithm D running in time $\mathcal{O}(t)$, making $Q \cdot n$ queries to an RSDP oracle (with secret as a matrix), and

$$\left|\Pr\left[\mathbf{X} \leftarrow \mathbb{E}^{k \times n} : D^{\Lambda_k(\mathbf{X})} = 1\right] - \Pr\left[D^{U_{k+1}} = 1\right]\right| \geq \left(\epsilon - 1/p^k\right)^2 - \alpha_{\mathbb{E}}^n.$$

Proof: Let $\mathcal{A}$ be a MitM attacker as described above. In each authentication query, $\mathcal{A}$ changes $\mathbf{b} \to \mathbf{b} + \mathbf{b}'$, $a \to a + a'$, $u \to u + u'$.

We now describe a distinguisher for the (vectorized) RSDP oracle using $\mathcal{A}$ as a part. The Challenger sends to the distinguisher pairs $(\mathbf{b}_i, \mathbf{y}_i)$ where $\mathbf{b}_i$ are uniformly random and $\mathbf{y}_i$ are either uniformly random or $\mathbf{y}_i = \mathbf{X}\mathbf{b}_i + \mathbf{e}_i$, for some random matrix $\mathbf{X}$. Now, the distinguisher simulates the Tag and the Reader for all queries. First, random values of the secrets h, s are chosen. For execution i, D receives $(\mathbf{b}_i, \mathbf{y}_i)$ from the challenger and chooses a random a_i. Then $(\mathbf{b}_i, a_i)$ is input to $\mathcal{A}$ who responds with $(\mathbf{b}'_i, a'_i)$. D computes u_i from its known values as

$$u_i = \mathbf{y}_i \cdot s + h(\mathbf{b}_i) \cdot (a_i + a'_i),$$

and gives to $\mathcal{A}$, who responds with u'_i. Finally, we simulate the response from the reader by simply rejecting if $(\mathbf{b}'_i, a'_i, u'_i) \neq (\mathbf{0}, 0, 0)$. This is the simulation before the winning query, and it is depicted in Figure 6.

We argue that this is a correct simulation of the protocol up to $Q - 1$ queries. If $\mathbf{y}_i = \mathbf{X}\mathbf{b}_i + \mathbf{e}_i$, then $\mathcal{T}$ and $\mathcal{R}$ behaves as if they have secrets $\mathbf{X}, s$ and h. Then, for $(\mathbf{b}'_i, a'_i, u'_i) = (\mathbf{0}, 0, 0)$, $\mathcal{R}$ outputs accept correctly. If $(\mathbf{b}'_i, a'_i, u'_i) \neq (\mathbf{0}, 0, 0)$, $\mathcal{R}$ outputs reject correctly since the winning query has not been reached.

Next, we will show how to use the Q-th winning query to output the answer to the Challenger correctly. The distinguished acts differently in the two cases: $\mathbf{b}' = \mathbf{0}$ and $\mathbf{b}' \neq \mathbf{0}$. From now on, we denote the winning query by $(\mathbf{b}', a', u') \neq (\mathbf{0}, 0, 0)$.

$\mathcal{T}_{\mathbf{X},h,s}$		$\mathcal{R}_{\mathbf{X},h,s}$
$\mathbf{b} \xleftarrow{\$} \mathbb{Z}_p^k$	$\xrightarrow{\mathbf{b}}$	
	$\xleftarrow{a}$	$a \xleftarrow{\$} \mathbb{F}$
$\mathbf{e} \xleftarrow{\$} \mathbb{E}^n$		
$u = (\mathbf{Xb} + \mathbf{e}) \cdot s + h(\mathbf{b}) \cdot a$	$\xrightarrow{u}$	Check if
		$(u - h(\mathbf{b}) \cdot a) \cdot s^{-1} - \mathbf{Xb} \in \mathbb{E}^n$

Figure 5: An RSDP adaptation to Lyubashevski et al. for MitM security. Here, we use the vector form of $(u-h(\mathbf{b})\cdot a)\cdot s^{-1}$ in the check.

$\mathcal{T}$ (sim.)	$\mathcal{A}$	$\mathcal{R}$ (sim.)
choose some random s, h		
$(\mathbf{b}_i, \mathbf{y}_i) \rightarrow$	$\xrightarrow{\mathbf{b}_i+\mathbf{b}'_i}$	
	$\xleftarrow{a_i+a'_i}$	$a_i \xleftarrow{\$} \mathbb{F}$
$u_i = \mathbf{y}_i \cdot s + h(\mathbf{b}_i) \cdot (a_i + a'_i)$	$\xrightarrow{u_i+u'_i}$	Check if
		$(\mathbf{b}'_i, a'_i, u'_i) \neq (\mathbf{0}, 0, 0)$, reject
		else, accept

Figure 6: Simulating $\mathcal{T}$ and $\mathcal{R}$ before the winning query.

$\mathcal{T}$ (sim.)	$\mathcal{A}$	$\mathcal{R}$ (sim.)
$(\mathbf{b}, \mathbf{y}) \rightarrow$	$\xrightarrow{\mathbf{b}}$	
	$\xleftarrow{a+a'}$	$a \xleftarrow{\$} \mathbb{F}$
$u = \mathbf{y} \cdot s + h(\mathbf{b}) \cdot (a + a')$	$\xrightarrow{u+u'}$	Check if
		$(u' + h(\mathbf{b}) \cdot a') \cdot s^{-1} \in \mathbb{D}^n$, not_random
		else, random

Figure 7: Response with the winning query when $\mathbf{b}' = \mathbf{0}$.

The case $\mathbf{b}' = \mathbf{0}$. The response to the challenger is as given in Figure 7.

- Case 1: Assume that the challenger sends RSDP pairs. Then, since (a', u') is the winning response, the following must be accepted by the Reader.

  $$(u + u' - h(\mathbf{b}) \cdot a) \cdot s^{-1} - \mathbf{Xb} \in \mathbb{E}^n.$$

  On the other hand, since the pairs $(\mathbf{b}, \mathbf{y})$ from the Challenger is an "RSDP" pair, we also have

  $$(u - h(\mathbf{b}) \cdot (a + a')) \cdot s^{-1} - \mathbf{Xb} \in \mathbb{E}^n.$$

  Subtracting the two equations, one obtains

  $$(u' + h(\mathbf{b}) \cdot a') \cdot s^{-1} \in \mathbb{D}^n,$$

  i.e., in this case, the response is always not_random.
- Case 2: Assume that the challenger sends a random pair. Then, the simulator responds not_random if and only if

  $$(u' + h(\mathbf{b}) \cdot a') \cdot s^{-1} \in \mathbb{D}^n$$

  with randomly chosen h and s. Moreover, previous unsuccessful queries do not leak information about these secrets, So $\mathcal{A}$ behaves as if u', a' are chosen before h and s. As $\mathbf{b}' = \mathbf{0}$, we have that u', a' cannot both be zero. This leads to

  $$\Pr[(u' + h(\mathbf{b}) \cdot a') \cdot s^{-1} \in \mathbb{D}^n] = \alpha_{\mathbb{E}}^n,$$

  for any choice of u', a'.
  In summary, if $\mathbf{b}' = \mathbf{0}$, then one has

  $$\left|\Pr\left[\mathbf{X} \leftarrow \mathbb{E}^{k \times n} : D^{\Lambda_k(\mathbf{X})} = 1\right] - \Pr\left[D^{U_{k+1}} = 1\right]\right| = \epsilon - \alpha_{\mathbb{E}}^n.$$

The case $\mathbf{b}' \neq \mathbf{0}$. The response to the challenger in this case is as given in Figure 8. Here we first run $\mathcal{A}$ in the winning query with random input $\mathbf{a_0}$, then we rewind $\mathcal{A}$ and run it again with another random input $\mathbf{a_1}$.

- Case 1: Assume that the challenger sends an RSDP pair. To detect if $\mathcal{A}$ responds correctly without knowing the secret $\mathbf{X}$, one needs to get rid of $\mathbf{X}(\mathbf{b}+\mathbf{b}')$. Therefore, we have to *rewind* $\mathcal{A}$ to produce $\mathbf{X}(\mathbf{b}+\mathbf{b}')$ twice with different challenges a_0 and a_1 (which happens with probability $1-1/p^k$). In particular, the simulated Reader sends a_0 and the Tag computes u_0 as in Figure 8. Then we rewind $\mathcal{A}$ back to the point it sends $\mathbf{b}+\mathbf{b}'$. This time, a different challenge a_1 is sent, and similarly, a new u_1 is computed.
  The distinguisher now has two responses $u_0 + u_0'$ and $u_1 + u_1'$. Both of them are simultaneously correct with different challenges with probability
  $$\left(\epsilon - 1/p^k\right)^2,$$
  and it follows that
  $$[(u_1 + u_1') - (u_0 + u_0') - h(\mathbf{b}+\mathbf{b}') \cdot (a_1 - a_0)] \cdot s^{-1} \in \mathbb{D}^n,$$
  meaning that in this case the distinguisher always gives the correct response not_random.
- Case 2: Assume that the challenger sends random pairs. We now investigate
  $$\Pr\left[[(u_1 + u_1') - (u_0 + u_0') - h(\mathbf{b}+\mathbf{b}') \cdot (a_1 - a_0)] \cdot s^{-1} \in \mathbb{D}^n\right]. \tag{4}$$
  Let $\tilde{a} = a_1 - a_0$ (and correspondingly for $\tilde{a}', \tilde{u}'$), we can rewrite Equation 4 as
  $$\Pr\left[[\tilde{u}' + h(\mathbf{b}) \cdot (\tilde{a} + \tilde{a}') - h(\mathbf{b}+\mathbf{b}') \cdot \tilde{a}] \cdot s^{-1} \in \mathbb{D}^n\right].$$

As previous unsuccessful queries do not leak information about h and s, $\mathcal{A}$ can be considered choosing $\tilde{u}', \tilde{a}'$ before h and s. For all $\mathbf{t} \in \mathbb{D}^n$,

$$\Pr[h(\mathbf{b}+\mathbf{b}') = h(\mathbf{b}) \cdot (\tilde{a} + \tilde{a}') + \tilde{u}' - \mathbf{t} \cdot s)] = \tfrac{1}{p^n},$$

by definition of pairwise independent function. Therefore,

$$\Pr\left[[\tilde{u}' + h(\mathbf{b}) \cdot (\tilde{a} + \tilde{a}') - h(\mathbf{b}+\mathbf{b}') \cdot \tilde{a}] \cdot s^{-1} \in \mathbb{D}^n\right] = \alpha_{\mathbb{E}}^n.$$

In summary, in this case, one has

$$\left|\Pr\left[\mathbf{X} \leftarrow \mathbb{E}^{k \times n} : D^{\Lambda_k(\mathbf{X})} = 1\right] - \Pr\left[D^{U_{k+1}} = 1\right]\right| = \left(\epsilon - 1/p^k\right)^2 - \alpha_{\mathbb{E}}^n.$$

One can extend the proof to use multiple rewinding as in Theorem 1 to achieve tighter reduction. □

6. Parameters for the MitM design

In contrast to the active-secured design, the RSDP sample is now masked with a secret polynomial s, which makes it challenging to apply any RSDP solvers to retrieve the secret $\mathbf{X}$ (even in the case that $\mathbf{X}$ is a Topeliz matrix). It is reasonable to assume that an Adversary will face a harder problem than just RSDP. Therefore, one can use the same parameters proposed in Table 5 for this design if the cost factor is paramount.

To be more conservative, one can select parameters based on the security reduction. Theorem 2 has tightness $\sqrt{\epsilon}$. In other words, the protocol is only half as secure as the RSDP problem and n has to be big enough so that $\alpha_{\mathbb{E}}^n$ can be deemed negligible. In particular, for d-bit security level, we ask for $\alpha_{\mathbb{E}}^n \leq 2^{-d}$.

TABLE 5: Agressive parameters for the MitM-secured RSDP authentication protocol with p as the field size, and the restricted set is set to be $\mathbb{E} := \{(-2)^i, i = 0, \dots z-1\}$.

Security Level	80	112	128
(p, z)	(127,14)	(127,14)	(127,14)
k	34	54	70
n	26	36	41

TABLE 6: Conservative parameters recommendation for the MitM-secured authentication protocol with p as the field size, and the restricted set is set to be $\mathbb{E} := \{(-2)^i, i = 0, \dots z-1\}$.

Security Level	80	112	128
(p, z)	(127,14)	(127,14)	(127,14)
k	55	79	91
n	96	134	153

There are several ideas when it comes to reducing the cost even more. For instance, selecting the coefficient of s, or secret polynomial h_1, and h_2, to be also in the restricted set. However, more careful work has to be done to understand the security implications of such risky options.

Example 5. Let us look at some efficiency comparisons between LPN and RSDP instantiation for 80-bit security. Since it is unclear how the steepness of reduction translates to security in practice, one can choose the parameters that yield 80-bit security for both the RSDP and LPN problem as in Section 4. In particular, (roughly) $(n, k, \epsilon) = (441, 512, 1/8)$ for LPN and $(n, k, p, z) = (26, 34, 127, 14)$ for RSDP.

- In terms of key storage, assuming we use Toepliz matrices for both cases, LPN uses $k + 4n = 2276$ bits, while RSDP uses $(k+n) \cdot \log(z) + 3n \cdot \log(p) = 770$ bits.
- LPN instantiation uses $3n = 1323$ and RSDP uses $3n \cdot \log(p) = 545$ bits in communication.

Therefore, we can see that RSDP offers significant advantages over traditional LPN-based protocols in terms of both key storage and communication costs.

7. Conclusion

In this paper, we have presented novel authentication protocols based on the Restricted Syndrome Decoding Problem. We show a natural adaptation of the new problem in two constructing directions: HB-family and wPRFs-based authentication protocols. From a theoretical viewpoint, security reductions from the previous works translate to RSDP (with a few additional assumptions, in the case of wPRFs-based). For practical interests, we show that with a well-chosen restricted set $\mathbb{E}$, the proposed protocol yields impressive performance well suited to low-cost cryptographic primitives. More importantly, such a choice of $\mathbb{E}$ does not compromise the security regarding available RSDP solvers.

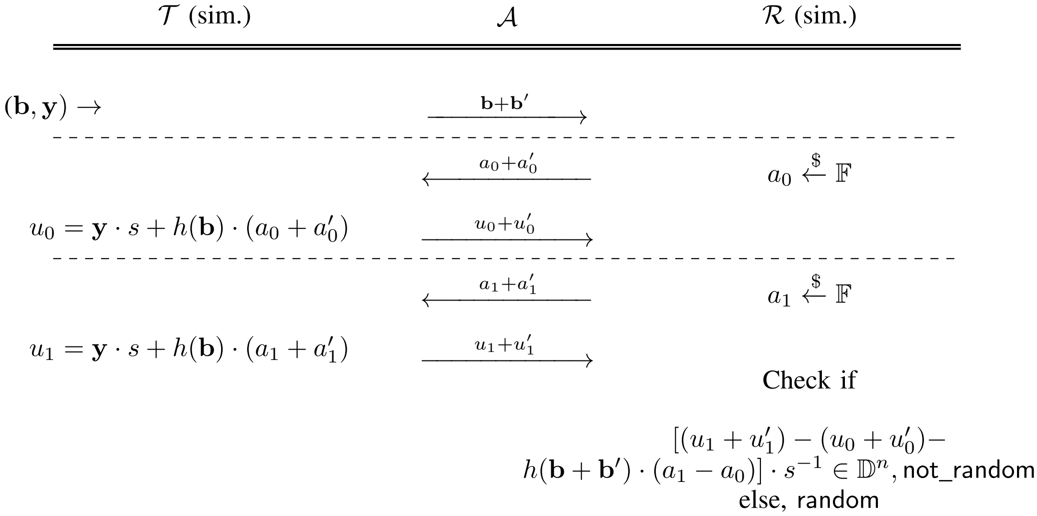

$\mathcal{T}$ (sim.)
$\mathcal{A}$
$\mathcal{R}$ (sim.)
$(\mathbf{b},\mathbf{y}) \rightarrow$
$\xrightarrow{\mathbf{b}+\mathbf{b}'}$
$\xleftarrow{a_0+a_0'}$
$a_0 \xleftarrow{\$} \mathbb{F}$
$u_0 = \mathbf{y}\cdot s + h(\mathbf{b})\cdot(a_0+a_0')$
$\xrightarrow{u_0+u_0'}$
$\xleftarrow{a_1+a_1'}$
$a_1 \xleftarrow{\$} \mathbb{F}$
$u_1 = \mathbf{y}\cdot s + h(\mathbf{b})\cdot(a_1+a_1')$
$\xrightarrow{u_1+u_1'}$
Check if
$[(u_1+u_1')-(u_0+u_0')-$
$h(\mathbf{b}+\mathbf{b}')\cdot(a_1-a_0)]\cdot s^{-1}\in\mathbb{D}^n$, not_random
else, random

Figure 8: Response with the winning query when $\mathbf{b}' \neq \mathbf{0}$.

Acknowledgments

Mustafa Khairallah is funded by the Wallenberg-NTU Presidential Postdoctoral Fellowship.

References

[1] N. J. Hopper and M. Blum, "Secure human identification protocols," in *Advances in Cryptology — ASIACRYPT 2001*, C. Boyd, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 52–66.

[2] A. Juels and S. A. Weis, "Authenticating pervasive devices with human protocols," in *Advances in Cryptology–CRYPTO 2005: 25th Annual International Cryptology Conference, Santa Barbara, California, USA, August 14-18, 2005. Proceedings 25*. Springer, 2005, pp. 293–308.

[3] D. D. Nguyen and K. Kim, "Securing HB+ against GRS man-in-the-middle attack," in *Presented at the Symp. Cryptogr. Inf. Security, Sasebo, Japan.*, 2007.

[4] J. Munilla and A. Peinado, "HB-MP: A further step in the HB-family of lightweight authentication protocols," *Computer Networks*, vol. 51, no. 9, pp. 2262–2267, 2007, (1) Advances in Smart Cards and (2) Topics in Wireless Broadband Systems. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128607000242

[5] J. Bringer and H. Chabanne, "Trusted-hb: a low-cost version of HB+ secure against man-in-the-middle attacks," *CoRR*, vol. abs/0802.0603, 2008. [Online]. Available: http://arxiv.org/abs/0802.0603

[6] H. Gilbert, M. J. B. Robshaw, and Y. Seurin, "HB#: Increasing the security and efficiency of HB+," in *Advances in Cryptology - EUROCRYPT 2008, 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13-17, 2008. Proceedings*, ser. Lecture Notes in Computer Science, vol. 4965. Springer, 2008, pp. 361–378. [Online]. Available: https://iacr.org/archive/eurocrypt2008/49650358/49650358.pdf

[7] H. Gilbert, M. Robshaw, and H. Sibert, "An active attack against HB+ - a provably secure lightweight authentication protocol," 2005, https://eprint.iacr.org/2005/237. [Online]. Available: https://eprint.iacr.org/2005/237

[8] H. Gilbert, M. J. B. Robshaw, and Y. Seurin, "Good variants of HB + are hard to find," in *Financial Cryptography and Data Security*, G. Tsudik, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 156–170.

[9] K. Ouafi, R. Overbeck, and S. Vaudenay, "On the security of HB# against a man-in-the-middle attack," in *Advances in Cryptology - ASIACRYPT 2008, 14th International Conference on the Theory and Application of Cryptology and Information Security, Melbourne, Australia, December 7-11, 2008. Proceedings*, ser. Lecture Notes in Computer Science, vol. 5350. Springer, 2008, pp. 108–124. [Online]. Available: https://www.iacr.org/archive/asiacrypt2008/53500111/53500111.pdf

[10] Z. Li, G. Gong, and Z. Qin, "Secure and efficient lcmq entity authentication protocol," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 4042–4054, 2013.

[11] E. Kiltz, K. Pietrzak, D. Cash, A. Jain, and D. Venturi, "Efficient authentication from hard learning problems," in *Advances in Cryptology – EUROCRYPT 2011*, K. G. Paterson, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 7–26.

[12] V. Lyubashevsky and D. Masny, "Man-in-the-middle secure authentication schemes from lpn and weak prfs," in *Advances in Cryptology – CRYPTO 2013*, R. Canetti and J. A. Garay, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 308–325.

[13] V. Nguyen, T. Johansson, and Q. Guo, "A key-recovery attack on the lcmq authentication protocol," in *2024 IEEE International Symposium on Information Theory (ISIT)*, 2024, pp. 1824–1829.

[14] S. Heyse, E. Kiltz, V. Lyubashevsky, C. Paar, and K. Pietrzak, "Lapin: An efficient authentication protocol based on ring-lpn," in *Fast Software Encryption - 19th International Workshop, FSE 2012, Washington, DC, USA, March 19-21, 2012. Revised Selected Papers*, ser. Lecture Notes in Computer Science, A. Canteaut, Ed., vol. 7549. Springer, 2012, pp. 346–365. [Online]. Available: https://doi.org/10.1007/978-3-642-34047-5_20

[15] H. Gilbert, M. J. B. Robshaw, and Y. Seurin, "Hb$^{\#}$: Increasing the security and efficiency of hb^{+}," in *Advances in Cryptology - EUROCRYPT 2008, 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13-17, 2008. Proceedings*, ser. Lecture Notes in Computer Science, N. P. Smart, Ed., vol. 4965. Springer, 2008, pp. 361–378. [Online]. Available: https://doi.org/10.1007/978-3-540-78967-3_21

[16] M. Baldi, M. Battaglioni, F. Chiaraluce, A. Horlemann-Trautmann, E. Persichetti, P. Santini, and V. Weger, "A new path to code-based signatures via identification schemes with restricted errors," *CoRR*, vol. abs/2008.06403, 2020. [Online]. Available: https://arxiv.org/abs/2008.06403

[17] M. Baldi, S. Bitzer, A. Pavoni, P. Santini, A. Wachter-Zeh, and V. Weger, "Zero knowledge protocols and signatures from the restricted syndrome decoding problem," in *Public-Key Cryptography - PKC 2024 - 27th IACR International Conference on Practice and Theory of Public-Key Cryptography, Sydney, NSW, Australia, April 15-17, 2024, Proceedings, Part II*, ser.

Lecture Notes in Computer Science, Q. Tang and V. Teague, Eds., vol. 14602. Springer, 2024, pp. 243–274. [Online]. Available: https://doi.org/10.1007/978-3-031-57722-2_8

[18] E. R. Berlekamp, R. J. McEliece, and H. C. A. van Tilborg, "On the inherent intractability of certain coding problems (corresp.)," *IEEE Trans. Information Theory*, vol. 24, no. 3, pp. 384–386, 1978. [Online]. Available: https://doi.org/10.1109/TIT.1978.1055873

[19] J. Stern, "A new identification scheme based on syndrome decoding," in *Advances in Cryptology - CRYPTO '93, 13th Annual International Cryptology Conference, Santa Barbara, California, USA, August 22-26, 1993, Proceedings*, ser. Lecture Notes in Computer Science, D. R. Stinson, Ed., vol. 773. Springer, 1993, pp. 13–21. [Online]. Available: https://doi.org/10.1007/3-540-48329-2_2

[20] P. Véron, "Improved identification schemes based on error-correcting codes," *Appl. Algebra Eng. Commun. Comput.*, vol. 8, no. 1, pp. 57–69, 1996. [Online]. Available: https://doi.org/10.1007/s002000050053

[21] N. T. Courtois, M. Finiasz, and N. Sendrier, "How to achieve a mceliece-based digital signature scheme," in *Advances in Cryptology - ASIACRYPT 2001, 7th International Conference on the Theory and Application of Cryptology and Information Security, Gold Coast, Australia, December 9-13, 2001, Proceedings*, ser. Lecture Notes in Computer Science, C. Boyd, Ed., vol. 2248. Springer, 2001, pp. 157–174. [Online]. Available: https://doi.org/10.1007/3-540-45682-1_10

[22] D. Augot, M. Finiasz, and N. Sendrier, "A family of fast syndrome based cryptographic hash functions," in *Progress in Cryptology - Mycrypt 2005, First International Conference on Cryptology in Malaysia, Kuala Lumpur, Malaysia, September 28-30, 2005, Proceedings*, ser. Lecture Notes in Computer Science, E. Dawson and S. Vaudenay, Eds., vol. 3715. Springer, 2005, pp. 64–83. [Online]. Available: https://doi.org/10.1007/11554868_6

[23] D. J. Bernstein, T. Lange, C. Peters, and P. Schwabe, "Really fast syndrome-based hashing," in *Progress in Cryptology - AFRICACRYPT 2011 - 4th International Conference on Cryptology in Africa, Dakar, Senegal, July 5-7, 2011. Proceedings*, ser. Lecture Notes in Computer Science, A. Nitaj and D. Pointcheval, Eds., vol. 6737. Springer, 2011, pp. 134–152. [Online]. Available: https://doi.org/10.1007/978-3-642-21969-6_9

[24] J. Fischer and J. Stern, "An efficient pseudo-random generator provably as secure as syndrome decoding," in *Advances in Cryptology - EUROCRYPT '96, International Conference on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, May 12-16, 1996, Proceeding*, ser. Lecture Notes in Computer Science, U. M. Maurer, Ed., vol. 1070. Springer, 1996, pp. 245–255. [Online]. Available: https://doi.org/10.1007/3-540-68339-9_22

[25] D. J. Bernstein, T. Chou, T. Lange, I. von Maurich, R. Misoczki, R. Niederhagen, E. Persichetti, C. Peters, P. Schwabe, N. Sendrier *et al.*, "Classic mceliece: conservative code-based cryptography," *NIST submissions*, 2017.

[26] N. Aragon, P. Barreto, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, S. Gueron, T. Guneysu, C. A. Melchor, R. Misoczki, E. Persichetti, N. Sendrier, J.-P. Tillich, G. Zemor, V. Vasseur, S. Ghosh, and J. Richter-Brokmann, "BIKE," 2021, available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[27] C. A. Melchor, N. Aragon, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, E. Persichetti, G. Zémor, and I. Bourges, "Hamming quasi-cyclic (hqc)," *NIST PQC Round*, vol. 2, no. 4, p. 13, 2018.

[28] G. A. Melchor, T. Feneuil, N. Gama, S. Gueron, J. Howe, D. Joseph, A. Joux, E. Persichetti, T. Randrianarisoa, M. Rivain, and D. Yue, "Sdith," *Round 1 Additional Signatures to the NIST PostQuantum Cryptography: Digital Signature Schemes Call*, 2023.

[29] G. Banegas, K. Carrier, A. Chailloux, A. Couvreur, T. Debris-Alazard, P. Gaborit, P. Karpman, J. Loyer, R. Niederhagen, N. Sendrier, B. Smith, and J. Tillich, "Wave," *Round 1 Additional Signatures to the NIST PostQuantum Cryptography: Digital Signature Schemes Call*, 2023.

[30] M. Baldi, A. Barenghi, S. Bitzer, P. Karl, F. Manganiello, A. Pavoni, G. Pelosi, P. Santini, J. Schupp, F. Slaughter, A. Wachter-Zeh, and V. Weger, "Cross," *Round 1 Additional Signatures to the NIST PostQuantum Cryptography: Digital Signature Schemes Call*, 2024.

[31] S. Arora and R. Ge, "New algorithms for learning in presence of errors," in *International Colloquium on Automata, Languages, and Programming*. Springer, 2011, pp. 403–415.

[32] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," *J. ACM*, vol. 56, no. 6, sep 2009. [Online]. Available: https://doi.org/10.1145/1568318.1568324

[33] B. Applebaum, D. Cash, C. Peikert, and A. Sahai, "Fast cryptographic primitives and circular-secure encryption based on hard learning problems," in *Advances in Cryptology - CRYPTO 2009*, S. Halevi, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 595–618.

[34] S. Bitzer, A. Pavoni, V. Weger, P. Santini, M. Baldi, and A. Wachter-Zeh, "Generic decoding of restricted errors," in *IEEE International Symposium on Information Theory, ISIT 2023, Taipei, Taiwan, June 25-30, 2023*. IEEE, 2023, pp. 246–251. [Online]. Available: https://doi.org/10.1109/ISIT54713.2023.10206983

[35] W. Beullens, P. Briaud, and M. Øygarden, "A security analysis of restricted syndrome decoding problems," *IACR Cryptol. ePrint Arch.*, p. 611, 2024. [Online]. Available: https://eprint.iacr.org/2024/611

[36] J. Stern, "A method for finding codewords of small weight," in *Coding Theory and Applications, 3rd International Colloquium, Toulon, France, November 2-4, 1988, Proceedings*, 1988, pp. 106–113. [Online]. Available: https://doi.org/10.1007/BFb0019850

[37] A. Becker, A. Joux, A. May, and A. Meurer, "Decoding random binary linear codes in $2n/20$: How $1+1=0$ improves Information Set Decoding," in *Advances in Cryptology – EUROCRYPT 2012*, D. Pointcheval and T. Johansson, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 520–536.

[38] P. Briaud and M. Øygarden, "A new algebraic approach to the regular syndrome decoding problem and implications for PCG constructions," in *Advances in Cryptology - EUROCRYPT 2023 - 42nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Lyon, France, April 23-27, 2023, Proceedings, Part V*, ser. Lecture Notes in Computer Science, C. Hazay and M. Stam, Eds., vol. 14008. Springer, 2023, pp. 391–422. [Online]. Available: https://doi.org/10.1007/978-3-031-30589-4_14

[39] J. Katz and J. S. Shin, "Parallel and concurrent security of the HB and HB+ protocols," in *Advances in Cryptology - EUROCRYPT 2006*, S. Vaudenay, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 73–87.

[40] É. Levieil and P.-A. Fouque, "An improved LPN algorithm," in *Security and Cryptography for Networks: 5th International Conference, SCN 2006, Maiori, Italy, September 6-8, 2006. Proceedings 5*. Springer, 2006, pp. 348–359.