HB$^{\#}$: Increasing the Security and Efficiency of HB^{+}

Henri Gilbert, Matthew J.B. Robshaw, and Yannick Seurin

Orange Labs, 38–40 rue General Leclerc, Issy les Moulineaux, France
{henri.gilbert,matt.robshaw,yannick.seurin}@orange-ftgroup.com

This is the full version of the *EUROCRYPT 2008* paper.

Abstract. The innovative HB^{+} protocol of Juels and Weis [10] extends device authentication to low-cost RFID tags. However, despite the very simple on-tag computation there remain some practical problems with HB^{+} and despite an elegant proof of security against some limited active attacks, there is a simple man-in-the-middle attack due to Gilbert *et al.* [8]. In this paper we consider improvements to HB^{+} in terms of both security and practicality. We introduce a new protocol that we denote RANDOM-HB$^{\#}$. This proposal avoids many practical drawbacks of HB^{+}, remains provably resistant to attacks in the model of Juels and Weis, and at the same time is provably resistant to a broader class of active attacks that includes the attack of [8]. We then describe an enhanced variant called HB$^{\#}$ which offers practical advantages over HB^{+}.

Key words: HB^{+}, RFID tags, authentication, LPN, Toeplitz matrix.

1 Introduction

The deployment of low-cost RFID tags is gathering pace. One familiar application is the inventory tracking of consumer items such as clothes, media products, and pharmaceuticals. However since blank tags can be programmed, there are opportunities for an attacker to clone an RFID tag and to introduce counterfeit goods into the supply chain. Thus, in this and other application areas there is much interest in deploying mechanisms for cryptographic tag authentication. However the physical demands for the deployment of cryptography on a cheap tag are substantial. Not only is space limited [10], but the peak and average power consumption often pose a demanding barrier for a tag that derives its power from a reader. Furthermore, since RFID tags pass fleetingly past a reader and are used in multi-tag and multi-reader environments, the communication is limited and its coordination complex.

Juels and Weis introduced HB^{+}, a three-pass symmetric key authentication protocol, at Crypto 2005 [10]. HB^{+} is computationally lightweight—requiring only simple bit-wise operations—and it is supported by a proof of security [10]. There are, however, some practical deficiencies in HB^{+} and the value of the proof of security has been somewhat limited by a simple active attack due to Gilbert *et al.* [8] which we will refer to as the GRS attack. Nevertheless, the simplicity

Tag (secret $\boldsymbol{x}, \boldsymbol{y}$)		Reader (secret $\boldsymbol{x}, \boldsymbol{y}$)
$\nu \in_R \{0, 1 \mid \text{Prob}(\nu = 1) = \eta\}$		
Choose $\boldsymbol{b} \in_R \{0,1\}^k$	$\xrightarrow{\boldsymbol{b}}$	
	$\xleftarrow{\boldsymbol{a}}$	Choose $\boldsymbol{a} \in_R \{0,1\}^k$
Let $z = \boldsymbol{a} \cdot \boldsymbol{x}^t \oplus \boldsymbol{b} \cdot \boldsymbol{y}^t \oplus \nu$	$\xrightarrow{z}$	Check $\boldsymbol{a} \cdot \boldsymbol{x}^t \oplus \boldsymbol{b} \cdot \boldsymbol{y}^t = z$

Fig. 1. One single round of HB^+ [10]. The entire authentication process requires r rounds and, in this basic form, each round consists of the three passes shown. Provided the tag fails less than some threshold t number of rounds, the tag is authenticated.

of both the original proposal and the active attack have led to a number of HB-related publications (see Section 2.2).

In this paper we propose solutions that improve on the practical problems of HB^+ while providing resistance to the GRS attack. The two simple proposals RANDOM-$\text{HB}^\#$ and $\text{HB}^\#$ provide more practical error rates than the original HB^+ and reduce the communication payload by a factor of around 20 (depending on the parameter sets). The protocol RANDOM-$\text{HB}^\#$ is provably secure in the *detection-based* model, the adversarial model used in *all* current proofs of security for HB^+ and its variants. But RANDOM-$\text{HB}^\#$ is also provably secure against the GRS attack and more generally in what we term the GRS-MIM model, an adversarial model that permits an active adversary to manipulate messages from the reader. The related protocol $\text{HB}^\#$ then gives a truly efficient scheme. While the same proofs do not immediately extend in their entirety to $\text{HB}^\#$, we can still say a surprising amount about the scheme in both theory and practice.

Our paper is organised as follows. First we describe HB^+ and some variants. Then, in Section 3, we introduce RANDOM-$\text{HB}^\#$ and provide full security proofs. In Section 4 we describe $\text{HB}^\#$ and its security and practical performance. We then highlight future work and draw our conclusions. Throughout we aim to use established notation. There will be some interplay between vectors $\boldsymbol{x} \in \{0,1\}^k$ (which we always consider to be row vectors) and scalars in GF(2). We use bold type $\boldsymbol{x}$ to indicate a row vector while scalars x are written in normal text. The bitwise addition of two vectors will be denoted $\oplus$ just as for scalars. We denote the *Hamming weight* of $\boldsymbol{x}$ by $\text{Hwt}(\boldsymbol{x})$.

2 HB^+ Variants and Tag Authentication

There are now several protocols based on HB^+ and these offer a variable level of security and practicality. We start by reviewing the original protocol. HB^+ is a three-pass authentication protocol built on the conjectured hardness of the *Learning from Parity with Noise* (LPN) problem [10].

LPN Problem. Let A be a random $(q \times k)$-binary matrix, let $\boldsymbol{x}$ be a random k-bit vector, let $\eta \in]0, \frac{1}{2}[$ be a noise parameter, and let $\boldsymbol{\nu}$ be a random q-bit vector such that $\mathrm{Hwt}(\boldsymbol{\nu}) \leq \eta q$. Given A, η, and $\boldsymbol{z} = A \cdot \boldsymbol{x}^t \oplus \boldsymbol{\nu}^t$, find a k-bit vector $\boldsymbol{y}^t$ such that $\mathrm{Hwt}(A \cdot \boldsymbol{y}^t \oplus \boldsymbol{z}) \leq \eta q$.

The HB^+ protocol is outlined in Figure 1. One doesn't need to look long to see that the goal of low on-tag computation has been achieved. Leaving aside generating $\boldsymbol{b}$ and the bit ν, computation on the tag is reduced to a dot-product (which can be computed bit-wise) and a single bit exclusive-or. Also HB^+ is accompanied by a proof of security. The adversarial model for this proof is referred to as the *detection-based* model [10] and requires that the adversary queries a tag q times and then attempts to pass the HB^+ authentication process by interacting with the reader once. Some commentators are not convinced that this adversarial model is sufficiently strong and an active attack against HB^+ exists when the adversary can interact with both the tag and the reader before attempting to impersonate the tag [8]. That said, the proof of security still has considerable value. The original proof [10] was rather sophisticated and applied to an adversary attempting to fool the reader over a single round of HB^+. This was extended by Katz and Shin [12] who also considered the parallel version of HB^+ with communications batched into one round of a three-pass protocol.

2.1 Some problems with HB^+

While HB^+ is computationally lightweight it still has some practical defects. The possibility of a legitimate tag being rejected has been commented on [12], but other issues such as the complex and extensive tag-reader communication would make HB^+ difficult to use. First, however, we highlight the fact that methods to solve the LPN problem have improved since the original presentation of HB^+.

LPN security and parameter choices. When considering the security and implementation of HB^+ there are four parameters that we need to set:

k :	the length of the secrets,	η :	the noise level,
r :	the number of rounds,	t :	the threshold for tag acceptance.

The first two parameters, k and η, quantify the resistance of the underlying LPN problem to attack. In [11] it is suggested that the parameter sets $k = 224$ and $\eta = 0.25$ provide around 80-bit security. Katz and Shin [12] propose $k \approx 200$ with $\eta = 0.125$, but we note that the reduced level of noise means that the LPN problem instance becomes easier and would necessitate an increase[1] to k.

Since the publication of HB^+ the LPN problem has been studied in more detail and the BKW algorithm cited in [10,12] has been improved. Fossorier *et al.* [6] show that the parameter choices used by [10] offer a level of security no greater than 2^{61} operations rather than the 2^{80} claimed. However, this has been superseded by the work of Levieil and Fouque [16] which suggests that the real security level offered by the parameters in [10] is no more than 2^{52} operations.

[1] However [12] is concerned with security proofs and specific parameter choices are somewhat orthogonal to their work.

Table 1. Error rates and transmission costs for different parameter sets in HB$^+$. The threshold $t = r\eta$ is proposed in [10] so we use $\lceil r\eta \rceil$ in this table. For the other parameters, [10] suggest $k = 224$ and $\eta = 0.25$ (leaving r unspecified) while [12] suggests $k \approx 200$, $\eta = 0.125$, with $40 \leq r \leq 50$. Based on the work of [16], we also consider the data transmission costs when $k = 512$ in the last column.

r	η	k	False reject rate (P_{FR})	False accept rate (P_{FA})	Transmission cost (bits) [k as given]	[$k = 512$]
80	0.25	224	0.44	4×10^{-6}	35,920	82,000
60	0.25	224	0.43	6×10^{-5}	26,984	61,500
40	0.25	224	0.42	1×10^{-3}	17,960	41,000
50	0.125	200	0.44	2×10^{-8}	20,050	51,250
40	0.125	200	0.38	7×10^{-9}	16,040	41,000

Considering [16] we propose alternative parameter values in Section 4.2 that are more consistent with the intended security level. In particular we propose $k = 512$ and $\eta = 0.125$ or, more conservatively, $k = 512$ and $\eta = 0.25$.

Error rates. A false rejection, a legitimate tag being rejected by a legitimate reader, occurs when the number of incorrect authentications exceeds the threshold t. A false acceptance takes place when an illegitimate tag is accepted by a legitimate reader. This occurs when t or fewer verification errors take place and we assume the illegitimate tag is reduced to guessing the reply z at random. The probability of a false rejection, P_{FR}, and a false acceptance, P_{FA}, are given by

$$P_{\mathrm{FR}} = \sum_{i=t+1}^{r} \binom{r}{i} \eta^i (1-\eta)^{r-i} \text{ and } P_{\mathrm{FA}} = \sum_{i=0}^{t} \binom{r}{i} 2^{-r}.$$

Note that both the false rejection and acceptance rate are independent of k, the size of the secrets, while the false acceptance rate is also independent of the noise level η used in HB$^+$. In the original descriptions of HB$^+$ a threshold of $t = r\eta$ is suggested. However (see Table 1) such a choice gives an unacceptably high false rejection rate. It is hard to imagine any practical scenario where a probability higher than 1% of rejecting a legitimate tag could be tolerated.

Transmission costs. HB$^+$ is a three-pass protocol that runs over r rounds. This requires the exchange of $2k + 1$ bits per round and $2rk + r$ bits in total. In the parallel version of the protocol, the data transmission requirements are the same but the data is packed into three passes of rk, rk, and r bits respectively. A three-pass protocol is considerably more practical than a $3r$-pass protocol (this was also mentioned in [12] as a justification for parallel HB$^+$). However the total amount of data transferred in both cases remains unacceptably high. In Table 1 we provide some estimates for the transmission costs in using HB$^+$. In particular we use parameter values that cover those proposed in [10,12]. We

Tag (secret $\boldsymbol{x}, \boldsymbol{y}$)		Reader (secret $\boldsymbol{x}, \boldsymbol{y}$)
$\nu \in \{0, 1 \vert \text{Prob}(\nu = 1) = \eta\}$		
Choose $\boldsymbol{b} \in_R \{0,1\}^k$	$\xrightarrow{\boldsymbol{b}}$	
	$\xleftarrow{\boldsymbol{a'} = \boldsymbol{a} \oplus \boldsymbol{\delta}} \cdots \xleftarrow{\boldsymbol{a}}$	Choose $\boldsymbol{a} \in_R \{0,1\}^k$
Let $z' = \boldsymbol{a'} \cdot \boldsymbol{x}^t \oplus \boldsymbol{b} \cdot \boldsymbol{y}^t \oplus \nu$	$\xrightarrow{z'}$	Check $\boldsymbol{a} \cdot \boldsymbol{x}^t \oplus \boldsymbol{b} \cdot \boldsymbol{y}^t = z'$

Fig. 2. The attack of Gilbert *et al.* [8] on HB^+. The adversary modifies the communications between reader and tag (by adding some perturbation $\boldsymbol{\delta}$) and notes whether authentication is still successful. This reveals one bit of secret information.

also include the transmission costs if we were to use parameter sizes that come closer to providing the intended 80-bit level of security.

An active attack. A simple active attack on HB^+ was provided in [8]. There it is assumed that an adversary can manipulate challenges sent by a legitimate reader to a legitimate tag during the authentication exchange, and can learn whether such manipulation gives an authentication failure. The attack consists of choosing a constant k-bit vector $\boldsymbol{\delta}$ and using it to perturb the challenges sent by a legitimate reader to the tag; $\boldsymbol{\delta}$ is exclusive-or'ed to each authentication challenge for each of the r rounds of authentication. If the authentication process is successful then we must have that $\boldsymbol{\delta} \cdot \boldsymbol{x}^t = 0$ with overwhelming probability. Otherwise $\boldsymbol{\delta} \cdot \boldsymbol{x}^t = 1$ with overwhelming probability and acceptance or rejection by the reader reveals one bit of secret information. The attack is illustrated in Figure 2 for one round of the HB^+ protocol. To retrieve the k-bit secret $\boldsymbol{x}$, one can repeat the attack k times for linearly independent $\boldsymbol{\delta}$'s and solve the resulting system. Conveniently, an adversary can choose $\boldsymbol{\delta}$'s with a single non-zero bit. With $\boldsymbol{x}$ an attacker can impersonate the tag by setting $\boldsymbol{b} = \boldsymbol{0}$. Alternatively, an attacker can emulate a false tag using $\boldsymbol{x}$, send a chosen blinding factor $\boldsymbol{b}$ to a legitimate reader, and return $\boldsymbol{a} \cdot \boldsymbol{x}^t$ to the challenge $\boldsymbol{a}$. If authentication is successful $\boldsymbol{b} \cdot \boldsymbol{y}^t = 0$, otherwise $\boldsymbol{b} \cdot \boldsymbol{y}^t = 1$, with overwhelming probability, and $\boldsymbol{y}$ can be recovered with k linearly independent $\boldsymbol{b}$.

Whether or not the attack is technically easy to mount it is *certificational*. The attack is mathematically simple and fully compromises HB^+. Protocols that resist this attack, while maintaining the computational simplicity of HB^+, would therefore be very attractive.

2.2 Other work on HB^+ and tag authentication

The novelty of the HB^+ protocol has generated considerable interest and much research. We have already mentioned the work of Katz and Shin [12] that closed gaps and extended the original proof of security. Follow-on work by Katz and Smith [13] has further extended these theoretical results to a larger range of noise levels $\frac{1}{4} \leq \eta < \frac{1}{2}$ whereas previous work [12] was only valid for $\eta < \frac{1}{4}$.

Other researchers have considered the active attack of Gilbert *et al.* [8]. Among them Bringer *et al.* [2] have outlined a protocol named HB^{++}. However the resulting protocol has some practical drawbacks. The data transmission costs of HB^{+} remain and the on-tag computation now includes bit-wise rotations and a small-block permutation f. Furthermore, an additional pre-protocol involving a universal hash function h is required to derive new tag/reader secrets at the start of each authentication. All this requires additional hardware and moves away from the essential simplicity of the HB^{+} protocol. Piramuthu [20] proposes a modification to HB^{++} in which the bit-wise rotations are varied for each round of the authentication and the message flow is simplified (saving one bit per authentication round). However the exact security claims are unclear. The variant HB^{*} is proposed by Duc and Kim [4] while another prominent protocol is HB-MP [19]. While both claim to be resistant to the attacks of [8], linear time attacks by the authors [7] show that this is not the case.

Naturally, research into other mechanisms for unilateral and mutual authentication continue in parallel. Schemes based on symmetric cryptography might use a lightweight block cipher [1,21] in a challenge-response protocol while other schemes might use asymmetric techniques such as GPS [9,18]. Other proposals include SQUASH [22] which might be viewed as a dedicated MAC, though the security goals appear to be somewhat reduced when compared to HB^{+} and the proposals RANDOM-$\text{HB}^{\#}$ and $\text{HB}^{\#}$ in this paper.

But this parallel work only serves to emphasize the interest in tag authentication and the importance of understanding the limits of proposals like HB^{+}. Despite the mixed success of current proposals in the literature, HB^{+} still holds much promise. This is due to the exceptionally low on-tag computational requirements and the fact that a proof of security, even if the model is weaker than we might ideally like, is a positive attribute.

3 The Proposal RANDOM-$\text{HB}^{\#}$

We now introduce RANDOM-$\text{HB}^{\#}$ (RANDOM-*HB-sharp*). This goes a long way to fixing many of the practical problems of HB^{+}. Like many other HB^{+}-variants, we prove the security of RANDOM-$\text{HB}^{\#}$ in the *detection-based* model, referred to in what follows as the DET-model. But we go further and prove the security of RANDOM-$\text{HB}^{\#}$ against a class of attacks that includes the GRS attack in what we term the GRS-MIM-model. More details are given in Section 3.1, but this model allows an active attacker to change any message from the reader in any way that they wish and observe the decision of the reader of whether to accept or not.

In RANDOM-$\text{HB}^{\#}$ we generalise HB^{+} and change the form of the secrets $\boldsymbol{x}$ and $\boldsymbol{y}$ from k-bit vectors into $(k_X \times m)$- and $(k_Y \times m)$-binary matrices X and Y. We illustrate RANDOM-$\text{HB}^{\#}$ protocol in Figure 3. One way of looking at RANDOM-$\text{HB}^{\#}$ is to observe that it is equivalent to m iterations of HB^{+}, but each column of X and Y in RANDOM-$\text{HB}^{\#}$ effectively represents a different HB^{+} secret $\boldsymbol{x}$ and $\boldsymbol{y}$. However, while RANDOM-$\text{HB}^{\#}$ carries much of the appearance of the HB^{+} protocol, there are important differences. In particular, the final verification by

Parameters: (k_X, k_Y, m, η, u)		
Tag (secret X, Y)		Reader (secret X, Y)
$\boldsymbol{\nu} \in_R \{\{0,1\}^m \mid$ Prob.$(\nu_i = 1) = \eta$ for $1 \leq i \leq m\}$		
Choose $\boldsymbol{b} \in_R \{0,1\}^{k_Y}$	$\xrightarrow{\boldsymbol{b}}$	
	$\xleftarrow{\boldsymbol{a}}$	Choose $\boldsymbol{a} \in_R \{0,1\}^{k_X}$
Let $\boldsymbol{z} = \boldsymbol{a} \cdot X \oplus \boldsymbol{b} \cdot Y \oplus \boldsymbol{\nu}$	$\xrightarrow{\boldsymbol{z}}$	Check
		$\mathrm{Hwt}(\boldsymbol{a} \cdot X \oplus \boldsymbol{b} \cdot Y \oplus \boldsymbol{z}) \leq um$

Fig. 3. The RANDOM-HB$^{\#}$ authentication protocol where the secrets X and Y are binary random matrices and the protocol has a single round. The verification step requires the comparison of two vectors and yields a PASS/FAIL verdict.

the reader consists of the comparison of two m-bit vectors $\boldsymbol{a} \cdot X \oplus \boldsymbol{b} \cdot Y$ and $\boldsymbol{z}$. For reader-verification we merely count the number of positions e that are in error and if $e \leq t$ for some threshold $t = um$, where $u \in]\eta, \frac{1}{2}[$, then we deduce that the tag is authentic. Thus RANDOM-HB$^{\#}$ and HB$^{\#}$ (see Section 4) consist of a single round.

3.1 Security results for RANDOM-HB$^{\#}$

We now provide security proofs for RANDOM-HB$^{\#}$ in two models. The first is the DET-model used in much of the founding work on HB^{+} [10,12]. Here the adversary is only allowed to query an honest tag without access to the reader. The second permits an active attacker to manipulate messages sent by the reader and will be referred to as the GRS-MIM-model.

Security definitions. In the following, the security parameter will be k, to which the number of rows of the secret matrices X and Y are related by $k_X = \Theta(k)$ and $k_Y = \Theta(k)$. We will say that a function (from positive integers to positive real numbers) is *negligible* if it approaches zero faster than any inverse polynomial, and *noticeable* if it is larger than some inverse polynomial. An algorithm will be *efficient* if it is a *Probabilistic Polynomial-Time* Turing machine. By saying that LPN is a hard problem, we mean that any efficient adversary solves it with only negligible probability.

We will let $\mathcal{T}_{X,Y,\eta}$ denote the algorithm run by an honest tag in the RANDOM-HB$^{\#}$ protocol and $\mathcal{R}_{X,Y,u}$ the algorithm run by the tag reader. We will prove the security of RANDOM-HB$^{\#}$ in two models:

- The DET-model, defined in [10,12], where attacks are carried out in two phases: the adversary first interacts q times with the honest tag. Then the adversary interacts with the reader and tries to impersonate the valid tag.

– The GRS-MIM-model: in a first phase, the adversary can eavesdrop on all communications between an honest tag and an honest reader (including the reader-decision of whether to accept or not) and in addition the attacker can modify any message from the reader to the tag for q executions of the protocol. Then the adversary interacts only with the reader and tries to impersonate the valid tag.

Note that the DET-model is a restriction of the GRS-MIM-model as any attack in the DET-model can easily be converted into an attack in the GRS-MIM-model. By replying at random to a challenge, the probability an adversary impersonating a tag will succeed is the false acceptance rate $P_{\text{FA}} = 2^{-m} \sum_{i=0}^{um} \binom{m}{i}$. This quantity is the best soundness we can achieve for RANDOM-HB$^{\#}$. Note that it is a function of m and u and not of the security parameter k, which will only set how close to P_{FA} the advantage of an adversary is bound to be. Note also that P_{FA} is negligible for any $u \in]\eta, \frac{1}{2}[$ and any $m = \Theta(k)$. We define the advantage of an adversary against the RANDOM-HB$^{\#}$ protocol in the DET and GRS-MIM models as its overhead success probability over P_{FA} in impersonating the tag:

$$\begin{aligned}
&\mathrm{Adv}_{\mathcal{A}}^{\mathrm{DET}}(k_X, k_Y, m, \eta, u, q) \stackrel{\mathrm{def}}{=} \\
&\quad \Pr\left[X \stackrel{\$}{\leftarrow} \mathcal{M}_X, Y \stackrel{\$}{\leftarrow} \mathcal{M}_Y,\ \mathcal{A}^{\mathcal{T}_{X,Y,\eta}}(1^k)\ :\ \langle \mathcal{A}, \mathcal{R}_{X,Y,u} \rangle = \text{ACC}\right] - P_{\text{FA}}; \\
&\mathrm{Adv}_{\mathcal{A}}^{\text{GRS-MIM}}(k_X, k_Y, m, \eta, u, q) \stackrel{\mathrm{def}}{=} \\
&\quad \Pr\left[X \stackrel{\$}{\leftarrow} \mathcal{M}_X, Y \stackrel{\$}{\leftarrow} \mathcal{M}_Y,\ \mathcal{A}^{\mathcal{T}_{X,Y,\eta}, \mathcal{R}_{X,Y,u}}(1^k)\ :\ \langle \mathcal{A}, \mathcal{R}_{X,Y,u} \rangle = \text{ACC}\right] - P_{\text{FA}}.
\end{aligned}$$

where $\mathcal{M}_X$ and $\mathcal{M}_Y$ denote resp. the sets of $(k_X \times m)$- and $(k_Y \times m)$-binary matrices and ACC denotes "accept".

Proof methods. We do not reduce the security of RANDOM-HB$^{\#}$ directly to the LPN problem. A preliminary step of our analysis is to define a natural matrix-based extension of the LPN problem and to prove its hardness. For this we appeal to the theory of "weakly verifiable puzzles". This is a notion introduced by Canetti, Halevi, and Steiner [3] and, informally, refers to a situation where only the entity that generates the puzzle holds secret information enabling the correctness of a candidate solution to be efficiently verified. As noticed by Katz and Shin [12], attacking the one-round HB protocol [10] in the passive model (that is, given q noisy samples $(\boldsymbol{a_i}, \boldsymbol{a_i} \cdot \boldsymbol{x}^t \oplus \nu_i)$, where $\boldsymbol{x}$ is a secret k-bit vector and the $\boldsymbol{a_i}$ are random k-bit vectors, and a random challenge $\boldsymbol{a}$, guess $\boldsymbol{a} \cdot \boldsymbol{x}^t$) may be viewed as a weakly verifiable puzzle. The result by Juels and Weis [10, Lemma 1] asserts, in essence, that this puzzle is $(1 - \frac{1}{2})$-hard if we assume the hardness of the LPN problem, which means that any efficient adversary trying to solve it has a success probability that is negligibly close (in k) to $\frac{1}{2}$. Canetti *et al.* [3] proved that if no efficient algorithm can solve a puzzle with probability more than ϵ, then no efficient algorithm can solve m independent puzzles simultaneously with probability more than ϵ^m. Thus, we define an extension of the HB puzzle that we call the *MHB puzzle*: given q noisy samples $(\boldsymbol{a_i}, \boldsymbol{a_i} \cdot X \oplus \boldsymbol{\nu_i})$, where X is a secret $(k \times m)$-matrix and the $\boldsymbol{a_i}$ are random k-bit vectors, and a random challenge $\boldsymbol{a}$, guess $\boldsymbol{a} \cdot X$. Using Canetti *et al.*'s result, we prove that any efficient adversary trying to solve it has a success probability that is negligibly close (in k) to $\frac{1}{2^m}$ (see Appendices A and B).

The security analysis is carried out in two steps. First we reduce the security of RANDOM-HB$^{\#}$ in the DET-model to the MHB puzzle. Then we reduce the security in the GRS-MIM-model to the security in the DET-model.

Theorem 1 (Security of RANDOM-HB$^{\#}$ in the DET-model). *Let $\mathcal{A}$ be an adversary attacking the* RANDOM-*HB$^{\#}$ protocol with parameters (k_X, k_Y, m, η, u) in the* DET-*model, interacting with the tag in at most q executions of the* RANDOM-*HB$^{\#}$ protocol, running in time T, and achieving advantage greater than δ. Then there is an adversary $\mathcal{A}'$, running in time at most $2mLq(2 + \log_2 q)T$, solving the MHB puzzle with parameters (k_Y, m, η, q'), where $q' = mLq(2 + \log_2 q)$ and $L = \frac{512}{\delta^4(1-2u)^4}(\ln m - \ln \ln 2)$, with success probability $> \left(\frac{1}{2^m} + \frac{\delta}{4}\right)$. Hence, assuming the hardness of the LPN problem, the advantage of any efficient* DET-*adversary against the* RANDOM-*HB$^{\#}$ protocol is negligible in k. As a consequence, for parameters $m = \Theta(k)$, the probability of any efficient* DET-*adversary to impersonate a valid tag is negligible in k.*

Proof. We slightly adapt the proof of Juels and Weis [11, Appendix C]. We denote by $\{(\boldsymbol{b_i}, \boldsymbol{z_i})\}_{1 \le i \le q'}$ the set of samples obtained by $\mathcal{A}'$ from the MHB puzzle generator with secret matrix Y and $\boldsymbol{b}$ the challenge vector for which $\mathcal{A}'$ aims to output $\boldsymbol{z} = \boldsymbol{b} \cdot Y$. $\mathcal{A}'$ uses its samples to simulate a tag algorithm $\mathcal{T}_{X,Y,\eta}$ where X is random with one line equal to $\boldsymbol{z}$. $\mathcal{A}'$ proceeds as follows:

1. Choose a random j, $1 \le j \le k_X$, and construct the $k_X \times m$ matrix X' where all rows are random except the j-th one which is undefined (say, equal to zero). Let $\boldsymbol{x_l}$ denote the l-th row of X'.
2. Divide the $q' = mLq(1+r)$ samples $\{(\boldsymbol{b_i}, \boldsymbol{z_i})\}_{1 \le i \le q'}$ into mL sets of $q(1+r)$ samples. For each bit position $s = 1$ to m, repeat the following L times, considering a fresh set of $q(1+r)$ samples each time:
   (a) For $i = 1$ to q repeat the following: draw a random bit α_i (this is a guess at the j-th bit of the challenge $\boldsymbol{a_i^+}$ which will be sent by the adversary $\mathcal{A}$). If $\alpha_i = 0$, send to $\mathcal{A}$ the blinding vector $\boldsymbol{b_i^+} = \boldsymbol{b_i}$, if $\alpha_i = 1$, send to $\mathcal{A}$ the blinding vector $\boldsymbol{b_i^+} = \boldsymbol{b_i} \oplus \boldsymbol{b}$. $\mathcal{A}$ sends back the challenge $\boldsymbol{a_i^+}$. If the guess was right (i.e. $\alpha_i = \boldsymbol{a_i^+}[j]$), then answer with the vector

   $$\boldsymbol{z_i^+} = \bigoplus_{l \ne j} \left(\boldsymbol{a_i^+}[l] \cdot \boldsymbol{x_l}\right) \oplus \boldsymbol{z_i}.$$

   Otherwise rewind adversary $\mathcal{A}$ to the beginning of its i-th query and try with a new $(\boldsymbol{b_{i'}}, \boldsymbol{z_{i'}})$ chosen among the rq supplementary samples.
   (b) If the rq samples are exhausted before the simulation of the query phase of $\mathcal{A}$ ends, randomly guess $\boldsymbol{z}[s]$.

(c) Otherwise, go to the cloning phase of $\mathcal{A}$: $\mathcal{A}$ sends a blinding vector $\hat{\boldsymbol{b}}$. Choose two random challenge vectors $\hat{\boldsymbol{a}}_{\boldsymbol{1}}$ and $\hat{\boldsymbol{a}}_{\boldsymbol{2}}$ such that they differ in their j-th bit. Transmit $\hat{\boldsymbol{a}}_{\boldsymbol{1}}$ to $\mathcal{A}$, record its response $\hat{\boldsymbol{z}}_{\boldsymbol{1}}$, rewind the adversary, transmit $\hat{\boldsymbol{a}}_{\boldsymbol{2}}$ to $\mathcal{A}$, and record its response $\hat{\boldsymbol{z}}_{\boldsymbol{2}}$ as well.

(d) Compute the guess for $\boldsymbol{z}[s]$ as

$$\hat{\boldsymbol{z}}_{\boldsymbol{1}}[s] \oplus \hat{\boldsymbol{z}}_{\boldsymbol{2}}[s] \oplus \left(\bigoplus_{l \neq j} (\hat{\boldsymbol{a}}_{\boldsymbol{1}}[l] \oplus \hat{\boldsymbol{a}}_{\boldsymbol{2}}[l]) \cdot \boldsymbol{x}_{\boldsymbol{l}}[s] \right).$$

3. Once L guesses have been made for each m bits of $\boldsymbol{z}$, take the majority outcome for each of them and output the answer accordingly.

Let us analyse what $\mathcal{A}'$ achieves. The repeated experiments on $\mathcal{A}$ share some common randomness ω (namely X and Y). Let us denote by ω' the randomness "renewed" at each experiment (that is the randomness used to simulate the tag, the random challenge $\hat{\boldsymbol{a}}$, and $\mathcal{A}$'s internal randomness). By a standard averaging argument, it holds that with probability greater than $P_{\text{FA}} + \frac{\delta}{2}$ over ω, the answer returned by $\mathcal{A}$ is correct in at least $m - t$ positions with probability greater[2] than $\frac{\delta}{2}$ over ω'. Let us assume that this is the case and show that $\mathcal{A}'$ returns a correct answer $\boldsymbol{z}$ with probability greater than $\frac{1}{2}$. The theorem will follow since $P_{\text{FA}} > \frac{2}{2^m}$ as soon as $t > 1$ and the overall probability of success for $\mathcal{A}'$ will be greater than $\frac{P_{\text{FA}}}{2} + \frac{\delta}{4} > \frac{1}{2^m} + \frac{\delta}{4}$.

First we will show that, during phase 2(a), $\mathcal{A}'$ simulates a tag algorithm $\mathcal{T}_{X,Y,\eta}$, where X is the X' matrix with $\boldsymbol{z}$ as j-th row. To see this, observe that when $\alpha_i = \boldsymbol{a}_{\boldsymbol{i}}^{+}[j] = 0$, then

$$\boldsymbol{z}_{\boldsymbol{i}}^{+} = \boldsymbol{a}_{\boldsymbol{i}}^{+} \cdot X \oplus \boldsymbol{b}_{\boldsymbol{i}} \cdot Y \oplus \boldsymbol{\nu}_{\boldsymbol{i}} = \boldsymbol{a}_{\boldsymbol{i}}^{+} \cdot X \oplus \boldsymbol{b}_{\boldsymbol{i}}^{+} \cdot Y \oplus \boldsymbol{\nu}_{\boldsymbol{i}},$$

whereas when $\alpha_i = \boldsymbol{a}_{\boldsymbol{i}}^{+}[j] = 1$, then

$$\boldsymbol{z}_{\boldsymbol{i}}^{+} = \boldsymbol{a}_{\boldsymbol{i}}^{+} \cdot X \oplus \boldsymbol{z} \oplus \boldsymbol{b}_{\boldsymbol{i}} \cdot Y \oplus \boldsymbol{\nu}_{\boldsymbol{i}} = \boldsymbol{a}_{\boldsymbol{i}}^{+} \cdot X \oplus (\boldsymbol{b}_{\boldsymbol{i}} \oplus \boldsymbol{b}) \cdot Y \oplus \boldsymbol{\nu}_{\boldsymbol{i}} = \boldsymbol{a}_{\boldsymbol{i}}^{+} \cdot X \oplus \boldsymbol{b}_{\boldsymbol{i}}^{+} \cdot Y \oplus \boldsymbol{\nu}_{\boldsymbol{i}}.$$

Let us now analyse the advantage $\mathcal{A}'$ enjoys during a single guess for one bit of $\boldsymbol{z}$ during phase 2. First, one can upper bound the probability that $\mathcal{A}'$ enters phase 2(b) by the probability that any one of the q experiments results in the discarding of r pairs of the extra challenge-response pairs, which is $q2^{-r}$. Taking $r = \log_2 q + 1$ yields a probability not greater than $1/2$.

Consider phase 2(d) for a fixed bit position s. The guess of $\mathcal{A}'$ is right when both bits $\hat{\boldsymbol{z}}_{\boldsymbol{1}}[s]$ and $\hat{\boldsymbol{z}}_{\boldsymbol{2}}[s]$ are correct, or when they are both incorrect. Hence we are interested in lower bounding the probability p' of this event. First, we will lower bound the probability p over ω' that the s-th bit of the answer returned by $\mathcal{A}$ is correct. We will assume *w.l.o.g.* that this probability is the same in all positions (otherwise one can "symmetrize" $\mathcal{A}$ by applying a random permutation

[2] Otherwise the probability of success of the adversary would be upper bounded by $(1 - P_{\text{FA}} - \frac{\delta}{2})\frac{\delta}{2} + P_{\text{FA}} + \frac{\delta}{2} < \delta + P_{\text{FA}}$, contradicting the hypothesis on $\mathcal{A}$.

of $\{1, \ldots, m\}$ to the problem). We can lower bound p as follows. Suppose we draw a *random* bit position s. Clearly, this bit is correct with probability p over the choice of s and ω'. At the same time, conditioned on the fact that more than $m-t$ bits are correct, the s-th bit of the answer is correct with probability greater than $1-u$. Consequently, the overall probability for the s-th bit to be correct is greater than $(1-u)\frac{\delta}{2} + \frac{1}{2}(1-\frac{\delta}{2})$, hence $p \geq \frac{1}{2} + \epsilon$ where $\epsilon = \frac{\delta}{2}(\frac{1}{2} - u)$. Juels and Weis proved [10, Lemma 2] that in this case, the probability, *conditioned on the fact that* $\hat{\boldsymbol{a}}_{\mathbf{1}}$ *and* $\hat{\boldsymbol{a}}_{\mathbf{2}}$ *differ in a single bit* j, that both bits $\hat{\boldsymbol{z}}_{\mathbf{1}}[s]$ and $\hat{\boldsymbol{z}}_{\mathbf{2}}[s]$ are correct or incorrect at the same time, is greater than $\frac{1}{2} + \epsilon^3/2 - (\epsilon^3+1)/k_X$. However one can improve on their analysis by using Jensen's inequality[3]. Let γ denote the randomness except for $\hat{\boldsymbol{a}}$ in the experiment ω' we are considering. For a fixed γ, let p_γ denote the probability over $\hat{\boldsymbol{a}}$ that the s-th bit of the answer from $\mathcal{A}$ is correct. We've just proved that $\sum_\gamma p_\gamma \geq \frac{1}{2} + \epsilon$. Let p'_γ denote for a fixed γ, the probability, *conditioned on the fact that* $\hat{\boldsymbol{a}}_{\mathbf{1}}$ *and* $\hat{\boldsymbol{a}}_{\mathbf{2}}$ *differ in a single bit* j, that both bits $\hat{\boldsymbol{z}}_{\mathbf{1}}[s]$ and $\hat{\boldsymbol{z}}_{\mathbf{2}}[s]$ are correct or incorrect at the same time. Following the proof of [10, Lemma 2] we have $p'_\gamma \geq \phi(p_\gamma)$ where

$$\phi(x) = x^2 \left(\frac{k_X + \log_2 x - 1}{k_X}\right) + (1-x)^2 \left(\frac{k_X + \log_2(1-x) - 1}{k_X}\right).$$

As ϕ is convex, one has the following inequalities:

$$p' = \sum_\gamma p'_\gamma \geq \sum_\gamma \phi(p_\gamma) \geq \phi(\sum_\gamma p_\gamma) = \phi(p) \geq \phi(\frac{1}{2} + \epsilon) \geq \frac{1}{2} + 2\epsilon^2 - \frac{1}{k_X}.$$

As $\mathcal{A}'$ enters phase 2(b) with probability less than $1/2$, the probability that $\mathcal{A}'$ guesses bit $\boldsymbol{z}[s]$ correctly is lower-bounded by $\frac{1}{4} + \frac{p'}{2} \geq \frac{1}{2} + \epsilon'$, with $\epsilon' = \epsilon^2 - \frac{1}{2k_X}$.

Using the Chernoff bound, taking the majority outcome of the L experiments allows $\mathcal{A}'$ to guess bit s with probability greater than

$$\pi = \left(1 - e^{\frac{-L\epsilon'^2}{1+2\epsilon'}}\right) \geq \left(1 - e^{\frac{-L\epsilon'^2}{2}}\right).$$

All m bits will be correct with probability greater than $\pi^m \geq \left(1 - e^{\frac{-L\epsilon'^2}{2}}\right)^m$. A probability of success greater than $\frac{1}{2}$ can be attained by taking

$$L = \frac{2}{\epsilon'^2} \ln\left(\frac{1}{1 - e^{-\frac{\ln 2}{m}}}\right) \sim \frac{512}{\delta^4(1-2u)^4}(\ln m - \ln\ln 2).$$

Hence, any efficient DET-adversary achieving a noticeable advantage against the RANDOM-HB$^\#$ protocol can be turned into an efficient solver of the MHB puzzle with a success probability greater than $\frac{1}{2^m} + \delta'$, where δ' is noticeable. This contradicts (Lemma 5 in Appendix B) the assumption that LPN is hard. ☐

[3] Note that this will also improve the security reduction for HB$^+$.

Theorem 2 (Security of RANDOM-**HB**$^{\#}$ **in the** GRS-MIM**-model).** *Let $\mathcal{A}$ be an adversary attacking the* RANDOM-*HB*$^{\#}$ *protocol in the* GRS-MIM-*model, modifying at most q executions of the protocol between an honest tag and an honest reader, running in time T, and achieving advantage greater than δ. Then, under an easily met condition on the parameter set (see the proof and Section 4.2), there is an adversary $\mathcal{A}'$ attacking the* RANDOM-*HB*$^{\#}$ *protocol in the* DET-*model, interacting at most q times with an honest tag, running in time $O(T)$, and impersonating a valid tag with success probability greater than $(P_{FA}+\delta)(1-q\epsilon)$ for some negligible function ϵ. Hence, assuming the hardness of the LPN problem, the advantage of any efficient* GRS-MIM-*adversary against the* RANDOM-*HB*$^{\#}$ *protocol is negligible in k. As a consequence, for parameters $m=\Theta(k)$, the probability of any efficient* GRS-MIM-*adversary to impersonate a valid tag is negligible in k.*

Proof. As $\mathcal{A}'$ has access to an honest tag that it can query freely, there is no difficulty in simulating an honest tag to $\mathcal{A}$. The main challenge comes with the task of simulating the honest reader. Recall that in the GRS-MIM-model, the adversary is only allowed to modify the messages from the reader to the tag. $\mathcal{A}'$ launches the first phase of the adversary $\mathcal{A}$ and simulates the tag and the reader for q times as follows:

1. $\mathcal{A}'$ obtains from the real tag $\mathcal{T}_{X,Y,\eta}$ a blinding vector $\boldsymbol{b_i}$; $\mathcal{A}'$ sends $\boldsymbol{b_i}$ as the blinding vector of the simulated tag to the simulated reader.
2. $\mathcal{A}'$ sends a random vector $\boldsymbol{a_i}$ as the challenge of the simulated reader. $\mathcal{A}$ modifies it into $\boldsymbol{a_i'}=\boldsymbol{a_i}\oplus\boldsymbol{\alpha_i}$. $\mathcal{A}'$ forwards $\boldsymbol{a_i'}$ to the real tag.
3. The real tag returns an answer $\boldsymbol{z_i}=\boldsymbol{a_i'}\cdot X\oplus\boldsymbol{b_i}\cdot Y\oplus\boldsymbol{\nu_i}$ to $\mathcal{A}'$ which uses it as the answer of the simulated tag to the simulated reader.
4. If $\boldsymbol{\alpha_i}$ was the all zero vector, $\mathcal{A}'$ outputs "ACCEPT" as the answer of the simulated reader, otherwise it outputs "REJECT".

After this first phase, $\mathcal{A}'$ launches the cloning phase of $\mathcal{A}$ and replicates its behaviour with the real reader. From the point of view of $\mathcal{A}$, the tag $\mathcal{T}_{X,Y,\eta}$ is perfectly simulated by $\mathcal{A}'$. Let Sim_i denote the event that the reader $\mathcal{R}_{X,Y,u}$ is correctly simulated by $\mathcal{A}$ during the i-th execution of the protocol, and Sim be the event that the reader is correctly simulated for all the q executions of the protocol, $\mathrm{Sim}=\cap_{i=1}^{q}\mathrm{Sim}_i$. Conditioning on this event Sim, the success probability of $\mathcal{A}'$ is the same as the success probability of $\mathcal{A}$, *i.e.* $P_{\mathrm{FA}}+\delta$. Hence, we have to lower bound the probability of Sim.

Consider one execution of the disturbed protocol. When $\boldsymbol{\alpha_i}=\boldsymbol{0}$, $\mathcal{A}'$ clearly fails at simulating the reader with a probability equal to the probability of wrongly rejecting an honest tag, *i.e.* P_{FR}. For the case $\boldsymbol{\alpha_i}\neq\boldsymbol{0}$ we make the following reasoning. Assume that the error vector $\boldsymbol{\alpha_i}\cdot X$ added by $\mathcal{A}$ has a Hamming weight d. This vector is added *before* the Bernoullian noise added by the tag, so that $\boldsymbol{\nu_i}$ is independent of $\boldsymbol{\alpha_i}\cdot X$. Consequently, the resulting error vector $\boldsymbol{\nu_i}\oplus\boldsymbol{\alpha_i}\cdot X$ has a Hamming weight distributed as the sum of d Bernoulli variables taking the value 1 with probability $1-\eta$ and 0 with probability η, and $m-d$ Bernoulli variables taking the value 1 with probability η and 0 with probability $1-\eta$. Hence, the mean value of the Hamming weight of the error vector is $\mu(d) = d(1-\eta) + (m-d)\eta$, and by the Chernoff bound, when $\mu(d) > t$, this weight is less than t with probability less than $e^{-\frac{(\mu-t)^2}{2\mu}}$, which remains true for any $d' \geq d$. Consequently, if the matrix X is such that for any $\boldsymbol{\alpha} \neq \mathbf{0}$, $\mathrm{Hwt}(\boldsymbol{\alpha} \cdot X)$ is high enough, outputting "REJECT" as soon as $\boldsymbol{\alpha_i} \neq \mathbf{0}$ will be a successful strategy. We formalize this as follows.

Let $d_{\min}(X) = \min_{\boldsymbol{\alpha} \neq \mathbf{0}} (\mathrm{Hwt}(\boldsymbol{\alpha} \cdot X))$ denote the minimal distance of the matrix X. We recall the following classical result of coding theory:

Lemma 1. *Let d be an integer in $[1..\lfloor \frac{m}{2} \rfloor]$ and let H be the entropy function $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$. Then*

$$\Pr_X[d_{\min}(X) \leq d] \leq 2^{-\left(1-\frac{k_X}{m}-H(\frac{d}{m})\right)m}.$$

This is a simple consequence of the following upper bound on the number of m-bit vectors of Hamming weight less than d: $\sum_{i=0}^{d} \binom{m}{i} \leq 2^{mH(\frac{d}{m})}$. For any non-zero vector $\boldsymbol{\alpha}$, $\boldsymbol{\alpha} \cdot X$ is uniformly distributed, and hence has Hamming weight less than d with probability less than $2^{m\left(H(\frac{d}{m})-1\right)}$. The lemma follows by a union bound.

Let $\tilde{d}$ be the least integer such that $\mu > t$, *i.e.* $\tilde{d} = 1 + \left\lfloor \frac{t-\eta m}{1-2\eta} \right\rfloor$. Then for any $d \geq \tilde{d}$ when $\boldsymbol{\alpha_i} \neq \mathbf{0}$, one can write

$$\begin{aligned}
\Pr_{X,\boldsymbol{\nu_i}}[\overline{\mathrm{Sim}_i}] &= \Pr_{\boldsymbol{\nu_i}}[\overline{\mathrm{Sim}_i} \mid d_{\min}(X) > d] \cdot \Pr_X[d_{\min}(X) > d] \\
&\qquad + \Pr_{\boldsymbol{\nu_i}}[\overline{\mathrm{Sim}_i} \mid d_{\min}(X) \leq d] \cdot \Pr_X[\min(X) \leq d] \\
&\leq \Pr_{\boldsymbol{\nu_i}}[\overline{\mathrm{Sim}_i} \mid d_{\min}(X) > d] + \Pr_X[d_{\min}(X) \leq d] \\
&\leq e^{-\frac{(\mu-t)^2}{2\mu}} + 2^{-\left(1-\frac{k_X}{m}-H(\frac{d}{m})\right)m}.
\end{aligned}$$

For this upper bound to be useful, the coefficient $\left(1 - \frac{k_X}{m} - H(\frac{d}{m})\right)$ must be positive for some $d \geq \tilde{d}$, in particular for $\tilde{d}$ as it is a decreasing function of d. This is a condition which is easily met for typical values of the parameters (see Section 4.2). Note also that for the asymptotic reduction we have to define $\tilde{d}$ as the least integer such that $\mu(\tilde{d}) > (1+c)t$ for some $c > 0$ in order to ascertain that the first term in the upper bound will be negligible. This way one has, for all $d \geq \tilde{d}$, $e^{-\frac{(\mu-t)^2}{2\mu}} \leq e^{-\frac{uc^2}{2(1+c)}m}$.

Together we have $\Pr[\overline{\mathrm{Sim}_i}] \leq \epsilon$, where ϵ is a negligible function given by

$$\epsilon = \max\left\{P_{\mathrm{FR}}, \min_{d \geq \tilde{d}} \left(e^{-\frac{(\mu-t)^2}{2\mu}} + 2^{-\left(1-\frac{k_X}{m}-H(\frac{d}{m})\right)m}\right)\right\}.$$

Consequently, $\Pr[\mathrm{Sim}] \geq (1 - q\epsilon)$ and $\mathcal{A}'$ has a success probability greater than $(P_{\mathrm{FA}} + \delta)(1 - q\epsilon)$.

If δ is noticeable then $q\epsilon(P_{\mathrm{FA}} + \delta) \leq \delta/2$ for k big enough, and the success probability of $\mathcal{A}'$ is greater than $P_{\mathrm{FA}} + \frac{\delta}{2}$. This contradicts Theorem 1. $\square$

With RANDOM-HB$^{\#}$ we have a surprisingly successful proposal. It is as computationally efficient as HB^{+} since it consists of a series of bitwise dot-product computations. At the same time it is simpler in terms of communication since there is only a single round and the total amount of data transmitted is much less than for HB^{+}. It also possesses a proof of security in the detection-based model, exactly like HB^{+}, but also against man-in-the-middle adversaries of the type used in the GRS attack. However there remains one drawback: storage. We show how to remedy this situation in the next section.

4 The Proposal HB$^{\#}$

In RANDOM-HB$^{\#}$ the tag is required to store two random $(k_X \times m)$- and $(k_Y \times m)$-binary matrices X and Y where k_X, k_Y and m are three-digit figures. The storage costs on the tag would be insurmountable. With this in mind we propose the protocol HB$^{\#}$. This has very modest storage requirements while preserving the computational efficiency of HB^{+}. While there are some subtle technical issues that mean we cannot transfer all the provably security results from RANDOM-HB$^{\#}$ to HB$^{\#}$ we can transfer some. These, together with a plausible conjecture, allow us to claim that HB$^{\#}$ is secure in the GRS-MIM-model. HB$^{\#}$ depends on the notion of a *Toeplitz* matrix. These were used by Krawczyk in message authentication proposals where their good distribution properties and efficient implementation were noted [14,15].

A $(k \times m)$-binary *Toeplitz* matrix M is a matrix for which the entries on every upper-left to lower-right diagonal have the same value. Since the diagonal values of a Toeplitz matrix are fixed, the entire matrix is specified by the top row and the first column. Thus a Toeplitz matrix can be stored in $k + m - 1$ bits rather than the km bits required for a truly random matrix. For any $(k + m - 1)$-bit vector $\boldsymbol{s}$, we denote by $T_{\boldsymbol{s}}$ the Toeplitz matrix whose top row and first column are represented by $\boldsymbol{s}$. HB$^{\#}$ is defined exactly as RANDOM-HB$^{\#}$ except that X and Y are now two random $(k_X \times m)$ and $(k_Y \times m)$-binary Toeplitz matrices.

4.1 Security results for HB$^{\#}$

While there is every indication that HB$^{\#}$ is secure in the DET-model, this remains to be shown. A first obvious step in this direction would be to prove that the Toeplitz variant of the MHB puzzle remains hard. We state the following conjecture to stimulate further research:

Conjecture 1 (Hardness of the Toeplitz-MHB puzzle). Let k be a security parameter, $\eta \in]0, 1/2[$, and m and q be polynomials in k. Let X be a random secret $(k \times m)$-binary *Toeplitz* matrix, and $(\boldsymbol{a_1}, \ldots, \boldsymbol{a_q})$ be q random vectors of length k. Then any efficient algorithm, on input q noisy samples $(\boldsymbol{a_i}, \boldsymbol{a_i} \cdot X \oplus \boldsymbol{\nu_i})$, where each bit of $\boldsymbol{\nu_i}$ is 1 with probability η, and a random vector $\boldsymbol{a}$ of length k, outputs $\boldsymbol{z} = \boldsymbol{a} \cdot X$ with probability negligibly close to $\frac{1}{2^m}$.

Just as for RANDOM-HB$^{\#}$, we can relate the security of the HB$^{\#}$ protocol in the GRS-MIM-model to its security in the DET-model.

Table 2. Practical parameters for HB$^\#$.

HB$^\#$ k_X	k_Y	m	η	t	False reject rate (P_{FR})	False accept rate (P_{FA})	Transmission (bits)	Storage (bits)
80	512	1164	0.25	405	2^{-45}	2^{-83}	1,756	2,918
80	512	441	0.125	113	2^{-45}	2^{-83}	1,033	1,472

Theorem 3 (Security of HB$^\#$ in the GRS-MIM-model). *Let $\mathcal{A}$ be an adversary attacking the HB$^\#$ protocol in the* GRS-MIM*-model, modifying at most q executions of the protocol between an honest tag and an honest reader, running in time T, and achieving advantage greater than δ. Then, under an easily met condition on the parameter set (see proof of Theorem 2 and Section 4.2), there is an adversary $\mathcal{A}'$ attacking the HB$^\#$ protocol in the* DET*-model, interacting at most q times with an honest tag, running in time $O(T)$, and impersonating a valid tag with success probability greater than $(P_{FA}+\delta)(1-q\epsilon)$ for some negligible function ϵ.*

Proof. (Outline) The proof is analogous to that of Theorem 2 and given in Appendix C. It relies on the observation that Lemma 1 remains true when the probability is taken over the set of random $(k_X \times m)$-Toeplitz matrices. □

Hence, the security of HB$^\#$ in the DET-model (which we believe to be a likely conjecture) would directly transfer to the GRS-MIM-model.

4.2 Parameter values for HB$^\#$

When considering the error rates in HB$^\#$, we have considerable flexibility in how we set the acceptance threshold t. Recall that the false rejection rate depends on m, t, and η and the false acceptance rate depends on m and t only. The overall security of the scheme depends on k_X, k_Y and η. However, as already noted by Levieil and Fouque [16] for HB$^+$, and as is clear from the proof of Theorem 1, k_X and k_Y play two different roles: only k_Y is related to the difficulty of the LPN problem, while k_X need only be 80-bit long to achieve 80-bit security.

Some example parameters for different noise levels η are given by Levieil and Fouque [16]. These give very reasonable error rates of $P_{\text{FR}} < 2^{-40}$ and $P_{\text{FA}} < 2^{-80}$. When combined with the larger values of k_Y required for good security with the LPN problem, the HB$^\#$ protocol compares very favourably to HB$^+$. The practical characteristics are summarised in Table 2. The condition necessary for Theorems 2 and 3 to hold is verified for both sets of parameters: for the first one, $\tilde{d} = 229$ and $\left(1 - \frac{k_X}{m} - H(\frac{\tilde{d}}{m})\right) \simeq 0.216$, while for the second one $\tilde{d} = 78$ and $\left(1 - \frac{k_X}{m} - H(\frac{\tilde{d}}{m})\right) \simeq 0.145$. The storage cost of HB$^\#$ is $(k_X + k_Y + 2m - 2)$ bits which is larger than the $2k$ bits required for HB$^+$. However, depending on the choice of m this is not necessarily a substantial increase. The given parameter choices offer 80-bit security (using the latest results on the LPN problem), the false acceptance and rejection rates are less than 2^{-80} and 2^{-40} respectively, and the total communication requirements are around 1,500 bits. This should be compared to error rates of 2^{-1} and 2^{-20} and transmission costs of up to 80,000 bits in the case of HB$^+$ (48,000 bits when $\boldsymbol{x}$ is only 80-bit long) for corresponding parameters. HB$^\#$ requires simple bit operations on-the-tag and thus remains computationally simple.

5 Further work and HB$^\#$ variants

General MIM adversaries. The result of Theorem 3 shows that an adversary successfully mounting an attack on HB$^\#$ must either (i) break HB$^\#$ in the DET-model (which we believe is highly improbable), or (ii) break the LPN problem, or (iii) use an undiscovered active attack involving more than manipulation of the messages from the reader. This raises the question of the security of HB$^\#$ against general man-in-the-middle adversaries allowed to perturb any message of the protocol. Though we do not have a formal proof of such a result, we can make the following heuristic analysis. To provide an appropriate context we might recall earlier work by Krawczyk [14,15]. Let us denote by $\mathcal{H}_\mathcal{T}$, where $\mathcal{T}$ stands for "random Toeplitz" matrix, the (k, m)-family of k-bit to m-bit linear functions $\boldsymbol{a} \mapsto \boldsymbol{a} \cdot T_{\boldsymbol{s}}$ associated with the set of $k \times m$ binary Toeplitz matrices $T_{\boldsymbol{s}}$, each associated with a $(k + m - 1)$-bit vector $\boldsymbol{s}$, and equipped with the uniform probability. The work of Krawczyk [15], which in turn references related work by Mansour *et al.* [17], in effect establishes that $\mathcal{H}_\mathcal{T}$ is $\frac{1}{2^m}$-balanced. In other words, for any non-zero vector $\boldsymbol{a}$, $\boldsymbol{a} \cdot T_{\boldsymbol{s}}$ is uniformly distributed over $\{0, 1\}^m$. This results from the fact that if $\boldsymbol{a}$ is a non-zero vector then $\boldsymbol{a} \cdot T_{\boldsymbol{s}}$ can be re-written as the product of $\boldsymbol{s}$ with a $(k + m - 1) \times m$ matrix derived from $\boldsymbol{a}$ that has rank m.

We can use this property of Toeplitz matrices to argue in favour of the resistance of HB$^\#$ against arbitrary man-in-the-middle adversaries. Consider an attack where the adversary perturbs $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{z}$ by adding respectively three disturbance vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$. The modified error vector is then $\boldsymbol{\nu'} = \boldsymbol{\nu} \oplus \boldsymbol{\alpha} \cdot X \oplus \boldsymbol{\beta} \cdot Y \oplus \boldsymbol{\gamma}$. When $\boldsymbol{\alpha} \neq \mathbf{0}$ or $\boldsymbol{\beta} \neq \mathbf{0}$, then due to the $\frac{1}{2^m}$-balance of $\mathcal{H}_\mathcal{T}$, $\boldsymbol{\nu'}$ is uniformly distributed and the probability that modifications of the communication between tag and reader result in successful authentication is the false acceptance probability P_{FA}. The reader's decision has negligible entropy and hence yields no information on X or Y to the adversary. On the contrary, when $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\mathbf{0}, \mathbf{0})$, the answer $\boldsymbol{z}$ returned by the tag is uniformly random so that $\boldsymbol{\gamma}$ may be considered as independent of X and Y. The reader's decision depends only on $\boldsymbol{\nu} \oplus \boldsymbol{\gamma}$ and again yields no information on X or Y to the adversary. It is helpful to note the essential difference between a man-in-the-middle attack on HB$^\#$ and the same attack on HB$^+$. When attacking HB$^+$, *e.g.* as is done in the GRS attack, the adversary gains 1 bit of information on $\boldsymbol{x}$ at *every* tag and reader HB$^+$ authentication (independently of whether it is successful or not), leading to a linear-time attack. By contrast, in the case of HB$^\#$, whatever the strategy for choosing $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$, the mutual information between the reader's decision and the matrices X and Y is negligible and no efficient adversary can gather noticeable information on X or Y. Though we believe that these observations can be made rigorous, it remains an open problem to extend the technique used in proof of Theorems 2 and 3 to arbitrary man-in-the-middle attacks and to find the right way of simulating the reader when the adversary can also modify $\boldsymbol{b}$ and $\boldsymbol{z}$.

Variants and optimisations. Independently of this theoretical work, there are interesting variants to HB$^{\#}$ that might be of practical value. One interesting option, also mentioned in [12], is for the legitimate tag to test that the noise vector $\boldsymbol{\nu}$ contains no more than t ones before using it. This means the probability of a false rejection would fall to zero. The main advantage of this approach would be to allow the size of m to decrease while maintaining a reasonable false acceptance rate. For instance, with $m = 256$, $\eta = 0.125$, and $t = 48$ we would ordinarily have that $P_{\mathrm{FA}} \approx 2^{-81}$ while $P_{\mathrm{FR}} \approx 2^{-9}$. However, this relatively high false rejection rate can be eliminated by allowing the tag to check $\boldsymbol{\nu}$ before use.

Another possibility to decrease storage and communication costs is to reduce k_Y; for this, it might be interesting to consider the effect of using a larger noise level, *i.e.* to have $\eta > \frac{1}{4}$. In such circumstances k_Y could be reduced—while maintaining the same level of security—thereby leading to storage and communications savings. While it is not immediately clear that this would be a successful approach, when coupled with restrictions to the noise vector $\boldsymbol{\nu}$ this may be worth exploring. Another optimisation could be to use techniques inspired by Krawczyk [14,15] to efficiently re-generate the Toeplitz matrices (*e.g.* by using a *LFSR*). We leave such proposals as topics for future research.

6 Conclusions

In this paper we have presented two new lightweight authentication protocols. While close variants of HB^{+}, these new protocols offer considerable advantages over related work in the literature. RANDOM-HB$^{\#}$ is provably secure in the detection-based model, just like HB^{+}, but it is also provably resistant to a broader class of attacks that includes [8]. The protocol HB$^{\#}$ trades some of the theoretical underpinnings to RANDOM-HB$^{\#}$ and attains a truly practical performance profile. Both RANDOM-HB$^{\#}$ and HB$^{\#}$ offer practical improvements over HB^{+}, and this remains the case even when using the problem sizes required after recent progress on solving the underlying LPN problem.

References

1. A. Bogdanov, L.R. Knudsen, G. Leander, C. Paar, A. Poschmann, M.J.B. Robshaw, Y. Seurin, and C. Vikkelsoe. PRESENT: An Ultra-Lightweight Block Cipher. In *Proceedings of CHES 2007*, LNCS 4727, pp. 450–466, Springer, 2007.
2. J. Bringer, H. Chabanne, and E. Dottax. HB^{++}: A Lightweight Authentication Protocol Secure Against Some Attacks. In *Proceedings of SecPerU 2006*, pp. 28–33, IEEE Computer Society Press, 2006.

3. R. Canetti, S. Halevi and M. Steiner. Hardness Amplification of Weakly Verifiable Puzzles. In *Proceedings of TCC 2005*, LNCS 3378, pp. 17–33, Springer, 2005.
4. D.N. Duc and K. Kim. Securing HB^+ Against GRS Man-in-the-Middle Attack. In *Institute of Electronics, Information and Communication Engineers, Symposium on Cryptography and Information Security*, Jan. 23–26, 2007.
5. M. Feldhofer, S. Dominikus, and J. Wolkerstorfer. Strong Authentication for RFID Systems Using the AES Algorithm. In *Proceedings of CHES 2004*, LNCS 3156, pp. 357–370, Springer, 2004.
6. M.P.C. Fossorier, M.J. Mihaljevic, H. Imai, Y. Cui, and K. Matsuura. A Novel Algorithm for Solving the LPN Problem and its Application to Security Evaluation of the HB Protocol for RFID Authentication. Available from `http://eprint.iacr.org/2006/197.pdf`.
7. H. Gilbert, M.J.B. Robshaw, and Y. Seurin. Good Variants of HB^+ are Hard to Find. In *Proceedings of Financial Crypto 2008*, to appear.
8. H. Gilbert, M.J.B. Robshaw, and H. Sibert. An Active Attack Against HB^+: A Provably Secure Lightweight Authentication Protocol. *IEE Electronics Letters*, volume 41, number 21, pp. 1169–1170, 2005.
9. M. Girault, G. Poupard and J. Stern. On the Fly Authentication and Signature Schemes Based on Groups of Unknown Order. *Journal of Cryptology*, volume 19, number 4, pp. 463–488, 2006.
10. A. Juels and S.A. Weis. Authenticating Pervasive Devices With Human Protocols. In *Proceedings of Crypto 2005*, LNCS 3126, pp. 293–198, Springer, 2005.
11. A. Juels and S.A. Weis. Authenticating Pervasive Devices With Human Protocols. Version of [10] with appendices. Available from `http://saweis.net/pdfs/lpn-paper.pdf`.
12. J. Katz and J. Shin. Parallel and Concurrent Security of the HB and HB^+ Protocols. In *Proceedings of Eurocrypt 2006*, LNCS 4004, pp. 73–87, Springer, 2006.
13. J. Katz and A. Smith. Analysing the HB and HB^+ Protocols in the "Large Error" Case. Available from `http://eprint.iacr.org/2006/326.pdf`.
14. H. Krawczyk. LFSR-based Hashing and Authentication. In *Proceedings of Crypto 1994*, LNCS 839, pp. 129–139, Springer, 1994.
15. H. Krawczyk. New Hash Functions for Message Authentication. In *Proceedings of Eurocrypt 1995*, LNCS 950, pp. 301–310, Springer, 1995.
16. E. Levieil and P.-A. Fouque. An Improved LPN Algorithm. In *Proceedings of SCN 2006*, LNCS 4116, pp. 348–359, Springer, 2006.
17. Y. Mansour, N. Nisan, and P. Tiwari. The Computational Complexity of Universal Hashing. In *Proceedings of STOC '90*, pp. 235–243, 1990.
18. M. McLoone and M.J.B. Robshaw. Public Key Cryptography and RFID. In *Proceedings of CT-RSA 2007*, LNCS 4377, pp. 372–384, Springer, 2007.
19. J. Munilla and A. Peinado. HB-MP: A Further Step in the HB-family of Lightweight Authentication Protocols. *Computer Networks*, volume 51, pp. 2262–2267, 2007.
20. S. Piramuthu. HB and Related Lightweight Authentication Protocols for Secure RFID Tag/Reader Authentication. *CollECTeR Europe Conference*, June 2006.
21. A. Poschmann, G. Leander, K. Schramm, and C. Paar. New Lightweight DES Variants Suited for RFID Applications. In *Proceedings of FSE 2007*, LNCS 4593, pp. 196–210, Springer, 2007.
22. A. Shamir. SQUASH - a New MAC With Provable Security Properties for Highly Constrained Devices Such as RFID Tags. In *Proceedings of FSE 2008*, to appear.

A Weakly Verifiable Puzzles [3]

Here we provide some background on weakly verifiable puzzles; more details can be found in [3]. A (weakly verifiable) puzzle system is a pair of efficient algorithms $\mathcal{Z} = (G, V)$. The puzzle-generator algorithm G, on input the security parameter 1^k, outputs a random puzzle p with some secret check information c, $(p, c) \xleftarrow{\$} G(1^k)$. The puzzle-verifier V is deterministic and on input a puzzle p, check information c, and an answer a, either accepts or rejects. A solver for this puzzle system is an efficient algorithm S that takes a puzzle p as input and outputs an answer a. Its success probability is defined as

$$\mathrm{succ}_{\mathcal{Z}}[S] \stackrel{\mathrm{def}}{=} \Pr_{G,S}\left[(p,c) \xleftarrow{\$} G(1^k), a \xleftarrow{\$} S(p) \; : \; V(p,c,a) = \textsc{acc}\right],$$

where the randomness is taken over G and S. A puzzle system is $(1-\epsilon)$-hard (where ϵ is any function from $\mathbb{N}$ to $]0,1[$) if any efficient solver has success probability upper bounded by $\epsilon + \mathrm{negl}(k)$ where *negl* is a negligible function.

The m-fold repetition of the puzzle system $\mathcal{Z} = (G, V)$ is the puzzle system $\mathcal{Z}^m = (G^m, V^m)$ where G^m, on input the security parameter 1^k, runs G m times and outputs the m puzzles with their check information,

$$((p_1, c_1), \ldots, (p_m, c_m)) \xleftarrow{\$} G^m(1^k).$$

The verifier V^m accepts the answer $(a_1, \ldots, a_m)$ of a solver if, and only if, V accepts for all triplets (p_i, c_i, a_i), $1 \leq i \leq m$.

Attacking the one-round HB protocol can be viewed as a weakly verifiable puzzle as follows.

Definition 1 (HB puzzle). *Let $\eta \in]0, 1/2[$, and q be a polynomial in k. On input the security parameter 1^k, the generator G draws a random secret key $\boldsymbol{x}$ of length k, q random vectors $(\boldsymbol{a_1}, \ldots, \boldsymbol{a_q})$ of length k, computes for $1 \leq i \leq q$ the set of answers $z_i = \boldsymbol{a_i} \cdot \boldsymbol{x}^t \oplus \nu_i$, where $\nu_i = 1$ with probability η, and draws a random vector $\boldsymbol{a}$ of length k constituting the challenge to the adversary. It outputs $\{(\boldsymbol{a_i}, z_i)\}_{1 \leq i \leq q}$ and $\boldsymbol{a}$. The solver returns a single bit z. The secret check information is $\boldsymbol{x}$, and the verifier V accepts if, and only if, $z = \boldsymbol{a} \cdot \boldsymbol{x}^t$.*

The security result by Juels and Weis [10, Lemma 1] giving the security of the one-round HB protocol, can be re-stated in terms of puzzles as:

Lemma 2 ([10], Lemma 1). *Assume the LPN problem is hard. Then the HB puzzle is $(1 - \frac{1}{2})$-hard.*

B Hardness of the MHB Puzzle

Reducing the security of $\textsc{random}$-HB$^{\#}$ directly to the LPN problem (or equivalently, the HB puzzle) is cumbersome. Instead, we will define a natural extension of the HB puzzle which we call MHB (Matrix-based HB) puzzle, and first prove its hardness.

Definition 2 (MHB puzzle). *Let $\eta \in]0, 1/2[$, and m and q be polynomials in k. On input the security parameter 1^k, the generator G draws a random secret $(k \times m)$-binary matrix X, q random vectors $(\boldsymbol{a_1}, \ldots, \boldsymbol{a_q})$ of length k, computes for $1 \le i \le q$ the set of answers $\boldsymbol{z_i} = \boldsymbol{a_i} \cdot X \oplus \boldsymbol{\nu_i}$, where each bit of $\boldsymbol{\nu_i}$ is 1 with probability η, and draws a random vector $\boldsymbol{a}$ of length k constituting the challenge to the adversary. It outputs $\{(\boldsymbol{a_i}, \boldsymbol{z_i})\}_{1 \le i \le q}$ and $\boldsymbol{a}$. The solver returns a vector $\boldsymbol{z}$. The secret check information is X, and the verifier V accepts if, and only if, $\boldsymbol{z} = \boldsymbol{a} \cdot X$.*

We will use the results of Canetti *et al.* on hardness amplification to assert the hardness of the MHB puzzle. However, the MHB puzzle is not the perfect m-fold repetition of the HB puzzle. In the m-fold repetition of the HB puzzle the challenge-sets $\{\boldsymbol{a_i}\}_{1 \le i \le q}$ and $\boldsymbol{a}$ are different for each secret column of X, in the MHB puzzle they are the same. Consequently we will need a slightly modified lemma. We begin by recalling the concrete results of [3], slightly re-formulated by combining their Lemma 1 and Section 3.3.

Lemma 3 ([3], Lemma 1). *Fix efficiently computable functions, $m : \mathbb{N} \to \mathbb{N}$, and ϵ, $\delta : \mathbb{N} \to]0, 1[$. Also fix a puzzle system $\mathcal{Z} = (G, V)$. If there exists a solver $\mathcal{A}$ for $\mathcal{Z}^m$ with success probability $\epsilon^m + \delta$ and running time T, then there exists a solver $\mathcal{A}'$ for $\mathcal{Z}$ with success probability greater than $\epsilon + \frac{\delta}{8m}$ and running time T' polynomial in m, $\frac{1}{\delta}$, $\frac{1}{\epsilon^m + \delta}$ and the running times of $\mathcal{A}$, G, and V.*

We need to adapt this lemma to situations where the repeated puzzles share some common randomness. To do this we generalize the repetition of puzzles in the following way. We suppose that the puzzle generator is composed of two independent algorithms $G = (G_f, G_v)$ (for *fixed* and *variable*). The secret check information is generated by G_v. The success probability $C = \text{succ}_{\mathcal{Z}}[S]$ of a solver is defined as for a simple puzzle:

$$C = \Pr_{(G_f, G_v), S}\left[((p_f, p_v), c) \xleftarrow{\$} G(1^k), a \xleftarrow{\$} S(p_f, p_v) : V((p_f, p_v), c, a) = \text{ACC}\right].$$

We define the m-fold pseudo-repetition $\widetilde{\mathcal{Z}^m}$ of the puzzle $\mathcal{Z} = ((G_f, G_v), V)$ as follows: the generator algorithm $\widetilde{G^m}$, on input the security parameter 1^k, runs G_f a single time: $p_f \xleftarrow{\$} G_f$, runs G_v m times: $p_v^i \xleftarrow{\$} G_v$, $1 \le i \le m$, and outputs the m puzzles (p_f, p_v^i) with their check information

$$(((p_f, p_v^1), c_1), \ldots, ((p_f, p_v^m), c_m)) \xleftarrow{\$} \widetilde{G^m}(1^k).$$

The verifier $\widetilde{V^m}$ accepts if, and only if, V accepts the answer to all the m puzzles. The success probability of a solver for this puzzle $\widetilde{\mathcal{Z}^m}$ is defined naturally. We now re-state Lemma 3 for the pseudo-repetition of puzzles.

Lemma 4. *Fix efficiently computable functions $m : \mathbb{N} \to \mathbb{N}$ and ϵ, $\delta : \mathbb{N} \to]0, 1[$. Also fix a puzzle system $\mathcal{Z} = ((G_f, G_v), V)$.* We make the hypothesis[4] that a

[4] This hypothesis is true for the HB puzzle with $\epsilon = \frac{1}{2}$.

uniformly random answer has a probability ϵ of being accepted by V. *If there exists a solver $\mathcal{A}$ for $\widetilde{\mathcal{Z}^m}$ with success probability $\epsilon^m + \delta'$ and running time T, then there exists a solver $\mathcal{A}'$ for $\mathcal{Z}$ with success probability $\epsilon + \frac{\delta'^2}{32m}$ and running time T' polynomial in m, $\frac{1}{\delta'}$, $\frac{1}{\epsilon^m+\delta'}$ and the running times of $\mathcal{A}$, G, and V.*

Proof. Let (p_f, p_v) be the input puzzle to $\mathcal{A}'$. By a standard averaging argument we know that with probability greater than $\frac{\delta'}{2}$ over p_f, $\mathcal{A}$ solves the m-fold pseudo-repetition of the puzzle with probability greater than $\epsilon^m + \frac{\delta'}{2}$ over the sequence $(p_v^1, \ldots, p_v^m)$ and $\mathcal{A}$'s internal randomness. When this is the case, according to Lemma 3, $\mathcal{A}'$ can use the strategy described by Canetti *et al.* to solve the puzzle with probability greater then $\epsilon + \frac{\delta'}{16m}$. Saying it differently, when p_f is fixed, G_v can be viewed as a classical puzzle generator to which one can apply Lemma 3 for its m-fold (classical) repetition. Otherwise, $\mathcal{A}'$ outputs a random answer. The overall success probability of $\mathcal{A}'$ is then

$$\frac{\delta'}{2}\left(\epsilon + \frac{\delta'}{16m}\right) + (1 - \frac{\delta'}{2})\epsilon = \epsilon + \frac{\delta'^2}{32m}.$$

The running time of $\mathcal{A}'$ is upper bound by a quantity which, as in Lemma 3, is polynomial in m, $\frac{1}{\delta'}$, $\frac{1}{\epsilon^m+\delta'}$ and the running times of $\mathcal{A}$, G, and V. □

We are now ready to state the hardness of the MHB puzzle.

Lemma 5. *Assume the hardness of the LPN problem. Then the MHB puzzle is $(1 - \frac{1}{2^m})$-hard.*

Proof. Assume for a contradiction that there is an efficient adversary $\mathcal{A}$ solving the MHB puzzle with probability greater than $\left(\frac{1}{2^m} + \delta\right)$, where δ is noticeable. According to the definition of the pseudo-repetition of puzzles, the MHB puzzle is the m-fold pseudo-repetition of the HB puzzle where G_f generates the challenges $(\boldsymbol{a_1}, \ldots, \boldsymbol{a_q})$ and $\boldsymbol{a}$, whereas G_v generates the secret $\boldsymbol{x}$ and the noise bits $(\nu_1, \ldots, \nu_q)$. Hence, according to Lemma 4, there is an efficient adversary $\mathcal{A}'$ solving the HB puzzle with success probability greater than $\frac{1}{2} + \delta'$, where $\delta' = \frac{\delta^2}{32m}$ is noticeable. Consequently the HB puzzle cannot be $(1 - \frac{1}{2})$-hard which (Lemma 2) contradicts the hardness of LPN. □

C Proof of Theorem 3

Let us denote by $\mathcal{T}$ the family of k_X-bit to m-bit binary Toeplitz matrices $T_{\boldsymbol{s}}$, each associated with a $(k_X + m - 1)$-bit vector $\boldsymbol{s}$, and equipped with the uniform probability. Mansour *et al.* [17] showed that for any non-zero vector $\boldsymbol{\alpha}$, $\boldsymbol{\alpha} \cdot T_{\boldsymbol{s}}$ is uniformly distributed over $\{0,1\}^m$. This results from the fact that if $\boldsymbol{\alpha}$ is a non-zero vector then $\boldsymbol{\alpha} \cdot T_{\boldsymbol{s}}$ can be re-written as the product of $\boldsymbol{s}$ with a $(k_X + m - 1) \times m$ matrix derived from $\boldsymbol{a}$ that has rank m.

Hence, one can prove exactly as was done for Lemma 1 that

$$\Pr_{X \xleftarrow{\$} \mathcal{T}} [d_{\min}(X) \leq d] \leq 2^{-\left(1 - \frac{k_X}{m} - H(\frac{d}{m})\right)m}.$$

The proof then proceeds exactly as for Theorem 2.