Authentication from Weak PRFs with Hidden Auxiliary Input

Daniel Masny *

University of California, Berkeley
daniel.masny@berkeley.edu

Abstract. In this work, we study a class of randomized weak pseudorandom functions, which we call weak PRFs with hidden auxiliary input (HIwPRF). Compared to Learning Parity with Noise (LPN) or Learning with Errors (LWE) based randomized weak PRFs, it provides less algebraic structure such that many known techniques and constructions do not translate to this class.
We investigate the potential of HIwPRFs for secure message and user authentication. We construct a protocol that gives as strong security guarantees when instantiated with a HIwPRF as known from weak PRF, LPN or LWE based protocols.

1 Introduction

A weak pseudorandom function (wPRF) is a weakened notion of a pseudorandom function (PRF). While a PRF has pseudorandom outputs for any set of inputs, wPRF outputs are only indistinguishable from uniform if the inputs are chosen uniformly at random. In particular for adversarially chosen inputs, it might be easy to distinguish wPRF outputs from uniform.

For many assumptions such as DDH or LWR, it is much easier to construct a wPRF, which follows straightforwardly from these assumptions, than a PRF [BPR12,NR04]. In general, the existence of a wPRF also implies the existence of a PRF, since a wPRF implies a pseudorandom generator, which can be combined with a tree structure to construct a PRF using the GGM paradigm [GGM86]. Nevertheless, this transformation requires many wPRF evaluations and large circuit depth, which makes it considerably less efficient.

In fact for many applications not the full strength of a PRF is required and a wPRF is sufficient. One application that has received much attention is authentication, where a prover wants to convince a verifer of its authenticity, given a shared secret key. For security, the verifier needs to reject unauthentic interactions. In this setting, there have been considered roughly three different kinds of adversaries. A passive adversary solely eavesdropping the communication while an active adversary can interact with a prover, but not with a verifier. Passive and active adversaries are successful if they convince in a second phase a verifier of their authenticity. A stronger notion, called man-in-the-middle (MIM) can alter an interaction between a prover and a verifier. Such an adversary is considered already successul when altering at least one message without causing the verifier to reject.

Dodis et. al. [DKPW12] proposed an actively secure 3-round protocol from any wPRF. This was extended by Lyubashevsky and Masny [LM13] to hold against MIM adversaries. Both results follow a line of research to base authentication on the learning parity with noise (LPN) assumption [GKL90,BFKL94] which has started with the HB protocols [HB01] and resulted in 2 and 3-round protocols being secure against various notions of adversaries [JW05,GRS08,KSS10,KPC$^+$11,HKL$^+$12,CKT16].

* Supported by Center for Long-Term Cybersecurity (CLTC, UC Berkeley). Part of the work was done at Ruhr Universit"at Bochum under the support of DFG Research Training Group GRK 1817/1.

The LPN or also the learning with errors (LWE) [Reg05] assumption can be modeled as a distinguishing problem of a special class of wPRFs that are called randomized wPRFs (rwPRF) [ACPS09]. In particular for LPN and LWE, a rwPRF f simply takes a uniformly chosen linear function f' as key and adds a noise term which follows a Bernoulli or discrete Gaussian distribution. For an input a, this rwPRF is defined by $f(a) = f'(a) + e$, where e is a noise term with small norm. This provides a rich algebraic structure that implies key-homomorphism, verifiability [ACPS09] and even public key encryption for sufficiently small noise by rerandomizing input output pairs [Ale03,Reg05]. Here, verifiability means that there is an efficient algorithm that takes two rwPRF evaluations and outputs with overwhelming probability 1 if and only if both outputs are evaluations for the same input, but not necessarily for the same randomness. A wPRF is trivially verifiable and a LPN or LWE based rwPRF as well as long as the noise terms are sufficiently small such that $f'(a) + e_1 - (f'(a) + e_2) = e_1 - e_2$ has still a small norm. This property seems to be crucial for 3-round protocols, since during the proof of security a rewinding argument is used to extract and compare two rwPRF evaluations for the same input, but not necessarily the same randomness.

In this work, we consider a different generalization of wPRFs, which we call weak pseudorandom functions with hidden auxiliary input (HIwPRF). A HIwPRF f takes two inputs, a public label a and a hidden value r. A distinguisher for a HIwPRF is asked to distinguish samples of the form $a,\ f(r, a)$ for uniform a and r from uniform. In case of LPN or LWE, the hidden auxiliary input would be the noise term. Nonetheless, we require for a HIwPRF that the auxiliary input has at most logarithmic size in the security parameter, while the output has at least linear size. This is not the case for LPN and LWE, which require noise with a large support, but it ensures that we exclude trivial constructions like a function that solely outputs the hidden auxiliary input, i.e. $f(r, a) = r$.

Our main contribution is to provide MIM secure message authentication from any HIwPRF, where a MIM adversary has sequential access to a prover and concurrent access to a verifier. This is a strengthening of the security notion compared to previous results [LM13,CKT16] which relied on sequential access. Though it still falls short of full concurrent MIM security. In a different line of research, Damgard and Park [DP14] considered a notion of concurrent, stateful MIM security that led to a secure authentication protocol from PRGs, though it is not secure against adversaries with access to an unbounded amount of provers without using strong underlying primitives. A trivial solution for settings with multiple provers in our as well as in their case could be the use of provers with independent secret keys.

In general, a HIwPRF is non-verifiable, simply since the tuple $a,\ f(r, a),\ f(r', a)$ could be still indistinguishable from uniform. Therefore, by giving a secure authentication protocol, we overcome the paradigm that verifiablity is necessary for cryptographically useful generalizations of wPRFs.

A simple candidate construction of a HIwPRF that is not already a wPRF is obtained by a problem called Random Selection [CKK08,KH12]. As in case of LPN and LWE, it relies on the efficiency and simplicity of linear functions with the difference that instead of adding noise, a random linear function is picked from a small set of linear functions. In more detail, a Random Selection based HIwPRF takes a set of uniform linear functions $S_1, \ldots, S_\ell$ in $\mathbb{Z}_2^{m \times n}$ as key. For any evaluation, it samples a uniform random $1 \leq r \leq \ell$ and outputs $f(r, a) = S_r a \in \mathbb{Z}_2^m$. The best known attack on this assumption is an algebraic attack which has a running time of magnitude $2^{O(\ell \log(n))}$ [KH12]. This HIwPRF seems to be non-verifiable since for keys $S_i,\ S_j,\ f(i, a) - f(j, a) = (S_i - S_j)a$ seems to be still hard to distinguish from uniform unless the trivial case where $i = j$.

2 Preliminaries

We use $x \leftarrow \mathcal{D}$, $x \leftarrow \mathsf{A}$, to sample from a distribution $\mathcal{D}$ or the output distribution of an algorithm A. When $\mathcal{D}$ is a set, it denotes sampling from the uniform distribution over $\mathcal{D}$. We use κ to denote the security parameter and call an algorithm probabilistic polynomial time (*ppt*) if its random tape and running time is polynomial in κ.

Pairwise Independent Functions. We use the following pairwise independent function classes.

Definition 1. [Pairwise Independent Hash Functions (PIH)]. *A family of efficiently computable hash functions* H *from* $\mathbb{D}$ *to* $\mathbb{I}$ *is called pairwise independent if for any* $x \in \mathbb{D}$ *and* $x_2 \in \mathbb{D} \setminus \{x\}$ *and* $d_1, d_2 \in \mathbb{I}$

$$\Pr_{h \leftarrow \mathsf{H}}[h(x) = d_1 \mid h(x_2) = d_2] = \frac{1}{|\mathbb{I}|}.$$

Definition 2. [Pairwise Independent Permutations (PIP)]. *A family of efficiently computable permutations* PIP *from* $\mathbb{D}$ *to* $\mathbb{I} = \mathbb{D}$ *is called pairwise independent if for any* $x \in \mathbb{D}$ *and* $x_2 \in \mathbb{D} \setminus \{x\}$ *and* $d_1, d_2 \neq d_1 \in \mathbb{D}$

$$\Pr_{g \leftarrow \mathsf{PIP}}[g(x) = d_1 \mid g(x_2) = d_2] = \frac{1}{|\mathbb{I}| - 1}.$$

Rewinding. Rewinding resets an algorithm to a previous state in order to receive an additional output for fresh random coins. Therefore, rewinding could be seen as two executions of an algorithm, where both runs use identical random coins til the rewound state and independent coins afterwards. Clearly, the two output distributions are not necessarily independent. We use two simple technical lemmas when using rewinding (for proofs, see Appendix B).

Lemma 1. *Let* A *be an algorithm with random tape* $\mathsf{T} \leftarrow \{0,1\}^\ell$ *and* X_A, X'_A *be random variables over* $\mathsf{T} \leftarrow \{0,1\}^\ell$ *that might depend on* A. *Then, for* X_A *with range* S *and* $y \in \sup(X'_\mathsf{A})$,

$$\Pr[X_\mathsf{A}(\mathsf{T}_0, \mathsf{T}_1) = X_\mathsf{A}(\mathsf{T}_0, \mathsf{T}_2) \mid X'_\mathsf{A}(\mathsf{T}_0, \mathsf{T}_1) = X'_\mathsf{A}(\mathsf{T}_0, \mathsf{T}_2) = y] \geq \frac{1}{|S|},$$

where the probabilities are taken over $(\mathsf{T}_0, \mathsf{T}_1) \leftarrow \{0,1\}^\ell$ *and* T_2, *which has the same distribution as* T_1.

Lemma 2. *Let* E_0, E_1 *be two events over a discrete probability space, then*

1. $\Pr[E_1] \Pr[E_0 \mid E_1]^2 + \Pr[\neg E_1] \Pr[E_0 \mid \neg E_1]^2 \geq \Pr[E_0]^2$,
2. $\Pr[E_1]^2 \Pr[E_0 \mid E_1]^2 + \Pr[\neg E_1]^2 \Pr[E_0 \mid \neg E_1]^2 \geq \frac{1}{2} \Pr[E_0]^2$.

Authentication. We consider 3-round message authentication protocols between a prover P and a verifier V, which share a secret key. The protocol consists of two sets Cmt, Ch, a message space $\mathcal{M}$, and three *ppt* algorithms $(\mathsf{Gen}, \mathsf{Rsp}, \mathsf{Vrfy})$. For $\mathsf{m} \in \mathcal{M}$, $\mathsf{cmt} \in \mathit{Cmt}$, and $\mathsf{ch} \in \mathit{Ch}$,

$\mathsf{Gen}(1^\kappa)$: Generates and outputs a secret key sk.
$\mathsf{Rsp}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch})$: Computes and outputs a response rsp for message m, commitment cmt and challenge ch.

$\mathsf{Vrfy}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp})$: Outputs 1 iff $(\mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp})$ is valid and otherwise 0.

During the protocol, P sends a commitment $\mathsf{cmt} \leftarrow Cmt$ and a message $\mathsf{m} \in \mathcal{M}$ to V, which responds with a challenge $\mathsf{ch} \leftarrow Ch$. P computes and sends $\mathsf{rsp} \leftarrow \mathsf{Rsp}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch})$ and V verifies its authenticity by computing $\mathsf{Vrfy}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp})$. We will call a authentication protocol correct if for any $\mathsf{m} \in \mathcal{M}$,

$$\Pr[1 = \mathsf{Vrfy}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{Rsp}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch}))] \geq 1 - negl(\kappa),$$

where the probability is taken over $\mathsf{sk} \leftarrow \mathsf{Gen}(1^\kappa)$, $\mathsf{cmt} \leftarrow Cmt$, $\mathsf{ch} \leftarrow Ch$, and the random coins of Vrfy and Rsp.

For security, we define two games, $\mathsf{G_{MIM}}$ and $\mathsf{G_U}$, in Figure 1. We call a protocol sequential

$\underline{\mathsf{G_{MIM}}(1^\kappa, \mathsf{A})\text{:}}$
$\mathsf{f}_b := 0;$
$(\mathsf{m_P}, \mathsf{cmt_P}) := (\perp, \perp);$
$\mathsf{L_P} \cup \mathsf{L_V} := \emptyset;$
$\mathsf{sk} \leftarrow \mathsf{Gen}(1^\kappa);$
Invoke $\mathsf{A}^{\mathcal{O}}(1^\kappa);$
Return f_b

$\underline{\mathcal{O}_{\mathsf{P},Cmt}(\mathsf{m})\text{:}}$
$\mathsf{cmt} \leftarrow Cmt;$
$\mathsf{m_P} := \mathsf{m};$
$\mathsf{cmt_P} := \mathsf{cmt};$
Return cmt

$\underline{\mathcal{O}_{\mathsf{P,Rsp}}(\mathsf{ch})\text{:}}$
If$((\mathsf{m_P}, \mathsf{cmt_P}) \in \mathsf{L_P})$ Return $\perp$
$\mathsf{rsp} \leftarrow \mathsf{Rsp}(\mathsf{sk}, \mathsf{m_P}, \mathsf{cmt_P}, \mathsf{ch});$
$\mathsf{L_P} := \mathsf{L_P} \cup \{(\mathsf{m_P}, \mathsf{cmt_P}, \mathsf{ch}, \mathsf{rsp})\};$
Return rsp

$\underline{\mathcal{O}_{\mathsf{V},Ch}(\mathsf{m}, \mathsf{cmt})\text{:}}$
$\mathsf{ch} \leftarrow Ch$
$\mathsf{L_V} := \mathsf{L_V} \cup \{(\mathsf{m}, \mathsf{cmt}, \mathsf{ch})\};$
Return ch

$\underline{\mathcal{O}_{\mathsf{V,Vrfy}}(\mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp})\text{:}}$
If$((\mathsf{m}, \mathsf{cmt}, \mathsf{ch}) \notin \mathsf{L_V})$ Return $\perp$
$b \leftarrow \mathsf{Vrfy}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp});$
$\mathsf{f}_b := \mathsf{f}_b \vee (b \wedge [(\mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp}) \notin \mathsf{L_P}]);$
$\mathsf{L_V} := \mathsf{L_V} \setminus \{(\mathsf{m}, \mathsf{cmt}, \mathsf{ch})\};$
Return b

Fig. 1. The game $\mathsf{G_{MIM}}$, where A has access to the oracles $\mathcal{O}_{\mathsf{P},Cmt}$, $\mathcal{O}_{\mathsf{P,Rsp}}$, $\mathcal{O}_{\mathsf{V},Ch}$, $\mathcal{O}_{\mathsf{V,Vrfy}}$, which we denote with $\mathcal{O}$, and f_b is the output after A terminates. $(\mathsf{m_P}, \mathsf{cmt}_P)$ is the state of the prover P, for which $\mathcal{O}_{\mathsf{P,Rsp}}$ outputs at most one response rsp. $\mathsf{L_V}$ is the state of the verifier V with concurrent access and $\mathsf{L_P}$ stores all valid output tuples generated by P.

prover, concurrent verifier Man-in-the-Middle (MIM) secure if for any *ppt* algorithm A

$$\epsilon_{\mathsf{A}} = \Pr[\mathsf{G_{MIM}}(1^\kappa, \mathsf{A}) = 1] \leq negl(\kappa),$$

where the probability is taken over the random coins of $\mathsf{G_{MIM}}$ and ϵ_{A} is the success probability of A. In Appendix C, we discuss user authentication, which is strictly weaker than message authentication.

3 Weak Pseudorandom Functions with Hidden Auxiliary Input

For a PRF, outputs for any set of inputs are indistinguishable from uniform. A wPRF weakens this notion by demanding the pseudorandomness of outputs only for uniformly chosen inputs, but not necessarily any set of inputs. In this work, we consider a generalization of wPRF's which we call weak pseudorandom functions with hidden auxiliary input (HIwPRF). This class of function takes an additional input from a set R that remains unseen by a distinguisher.

Definition 3. [HIwPRF]. *Let R be a set and F be a family of efficiently computable functions with domain $\mathbb{D} \times R$ and image $\mathbb{I}$, which are functions in κ. F is called a family of HIwPRF's if for $f \leftarrow \mathsf{F}$, oracles $\mathcal{O}_\mathsf{F}$, $\mathcal{O}_\mathsf{U}$,*

$\mathcal{O}_\mathsf{F}$: *Sample $r \leftarrow R$, output $a \leftarrow \mathbb{D}$, $t := f(r, a)$.*
$\mathcal{O}_\mathsf{U}$: *Output $a \leftarrow \mathbb{D}$, $t \leftarrow \mathbb{I}$.*

and any ppt algorithm A

$$\epsilon_\mathsf{A} = |\Pr[\mathsf{A}^{\mathcal{O}_\mathsf{F}} = 1] - \Pr[\mathsf{A}^{\mathcal{O}_\mathsf{U}} = 1]| \leq negl(\kappa),$$

where the probabilities are taken over $f \leftarrow \mathsf{F}$ and the random coins of A, $\mathcal{O}_\mathsf{F}$, $\mathcal{O}_\mathsf{U}$. ϵ_A is the success probability of A.

A HIwPRF is a potentially weaker primitive than a wPRF. If one allows a sufficiently large set R, a HIwPRF f could be constructed unconditionally by defining $f(r, a) := r$. Since such a HIwPRF does not seem to be useful, we require that $|R| = O(\text{poly}(\kappa))$ and $|\mathbb{I}| = \Omega(2^\kappa)$. In this setting, a HIwPRF implies a one-way function and therefore there are generic constructions of a PRF from a HIwPRF, though this would require a high computational depth with many HIwPRF evaluations for a single PRF output.

A natural candidate of a HIwPRF is provided by $\mathsf{Random\ Selection}$. One chooses $|R|$ uniform linear functions $f_1, \ldots, f_{|R|}$ from $\mathbb{D}$ to $\mathbb{I}$ and the HIwPRF candidate f would be $f(r, a) := f_r(a)$. In Appendix A we give a formal definition of $\mathsf{Random\ Selection}$ and state known hardness results.

4 Secure Message Authentication

Let $\mathbb{F}$ be a sufficiently large finite field, e.g. $|\mathbb{F}| = O(2^\kappa)$. For our proposed authentication protocol, we need a PIP family from $\mathbb{F} \to \mathbb{F}$, a PIH family from $Cmt \times \mathcal{M} \to \mathbb{F}$ and a HIwPRF family from $R \times Cmt \to \mathbb{F}$.

We define Gen, Rsp and Vrfy as follows, which is sufficient to fully describe a message authentication protocol.

$\mathsf{Gen}(1^\kappa)$: Sample $f \leftarrow \mathsf{F}$, $h \leftarrow \mathsf{H}$, $g \leftarrow \mathsf{PIP}$ and output $\mathsf{sk} := (f, h, g)$.
$\mathsf{Rsp}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch})$: Sample $r \leftarrow R$ and output

$$\mathsf{rsp} := g(f(r, \mathsf{cmt})) + \mathsf{ch} \cdot h(\mathsf{m}, \mathsf{cmt}) \in \mathbb{F}.$$

$\mathsf{Vrfy}(\mathsf{sk}, \mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp})$: Outputs 1 iff $\exists r \in R$ such that

$$\mathsf{rsp} - \mathsf{ch} \cdot h(\mathsf{m}, \mathsf{cmt}) = g(f(r, \mathsf{cmt}))$$

and otherwise 0.

The scheme is perfectly correct. For a given cmt, ch and m, there are exactly $|R|$ many potential outputs rsp of Rsp and Vrfy will compare rsp to all of them. The running time of Vrfy is polynomial in $|R| = O(\text{poly}(\kappa))$. While this brute forcing approach seems to have a very negative impact on the efficiency, the efficiency heavily depends on the length of the hidden auxiliary input, which is relatively short. Further, this process is highly parallelizeable and in for example the Rfid setting, it is carried out by a reader device for which a low computational complexity is not as critical as it is for an Rfid chip.

5 Security

Theorem 1. *If* F *is* HIwPRF*,* PIP *is* PIP *and* H *is* PIH*, then the proposed protocol is* MIM*-secure. In more detail, for every ppt algorithm* A *there exist a ppt algorithm* D *breaking the* HIwPRF *property of* F *in time* $t_{\mathsf{D}} \leq 2t_{\mathsf{A}}$ *and* $Q \leq 2Q_{Cmt}$ *queries with probability*

$$\epsilon_{\mathsf{D}} \geq \frac{1}{2|R| \cdot Q_{Cmt} \cdot Q_{Ch}^2} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1]^2 - negl(\kappa),$$

where Q_{Ch}*,* Q_{Cmt}*, is an upper bound on the amount of queries that* A *makes to oracle* $\mathcal{O}_{Ch}$*,* $\mathcal{O}_{Cmt}$ *respectively, and* R *is the domain of the random input of* F*.*

Proof. We use three games, $\mathsf{G}_{\mathsf{MIM}}$, G_1 and G_{U}, which are shown in Figure 2. Any efficient adversary with a non negligible success probability in $\mathsf{G}_{\mathsf{MIM}}$ also has a non negligible success probability in G_1 and G_{U}.

An important aspect of G_1 is that the HIwPRF is only evaluated for random inputs, but not on adversarially chosen inputs. Therefore it is easy to replace HIwPRF values in G_1 with uniform values as in G_{U} by using the pseudorandomness assumption of the HIwPRF. In G_{U} crucial parts of the secret key are information theoretically hidden such that no adversary can win the game with a non negligible success probability.

The hardest part of the proof is the transition from $\mathsf{G}_{\mathsf{MIM}}$ to G_1. In $\mathsf{G}_{\mathsf{MIM}}$, the verification oracle requires to evaluate HIwPRF on adversarially chosen inputs to determine the correctness of a response rsp. G_1 guesses the interaction of A with V at which A makes its first successful forgery. This guess denoted with j is correct with good probability. If this is the case, previous responses of A are either trivially correct, i.e. generated by P and hence contained in L_{P}, or incorrect. Still, checking the correctness of A's forgery during interaction j is non-trivial and replaced by three different strategies. An adversary A could simply use a HIwPRF evaluation used already by the prover P. This case is easy to handle. More problematic is when an adversary generates a fresh HIwPRF evaluation, i.e. which was not generated by P, by either choosing a fresh input or choosing solely a fresh hidden auxiliary input. In these two cases, G_1 will rewind A to different states, which we denote with either A_1 or A_2. By a rewinding argument, A and A_1, or A and A_2 will use exactly the same HIwPRF outputs, i.e. with the same input and hidden auxiliary input, with a good probability if A is successful in $\mathsf{G}_{\mathsf{MIM}}$. It turns out that it is sufficient to compare these two evaluations for consistency. Any adversary that does not distinguish HIwPRF outputs from uniform, i.e. behaves like in G_{U}, is not able to answer consistently when being rewinded.

Lemma 3. *For any algorithm* A *making at most* Q_{Ch}*,* Q_{Cmt}*, queries to* $\mathcal{O}_{Ch}$*,* $\mathcal{O}_{Cmt}$ *respectively, within* $\mathsf{G}_{\mathsf{MIM}}$*,*

$$\Pr[\mathsf{G}_1(\mathsf{A}) = 1] \geq \frac{1}{2|R|Q_{Cmt}Q_{Ch}^2} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1]^2,$$

where R *is the domain of the hidden auxiliary input of* F *and the running time and the amount of queries to* $\mathcal{O}_{Cmt}$ *of* A *in* G_1 *is at most two times the running time and the amount of queries* $\mathcal{O}_{Cmt}$ *of* A *in* $\mathsf{G}_{\mathsf{MIM}}$*.*

Proof. In G_1, A, A_1, A_2 are invoked with partially identical random coins. Therefore, we are more explicit about how G_1 uses its random tape T to generate a secret key sk and run A, A_1, A_2 and the oracles $\mathcal{O}$, $\mathcal{O}_1$ and $\mathcal{O}_2$. Let $\mathsf{T} = (\mathsf{T}_0, \mathsf{T}_1, \mathsf{T}_2, \mathsf{T}_5) = (\mathsf{T}_3, \mathsf{T}_4, \mathsf{T}_2, \mathsf{T}_5)$ be the random tape of G_1.

$\underline{\mathsf{G}_{\mathsf{MIM},1,\mathsf{U}}(1^\kappa,\mathsf{A})}$:
$\mathsf{f}_b := 0;$
$(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}) := (\perp, \perp);$
$\mathsf{L}_\mathsf{P} \cup \mathsf{L}_\mathsf{V} := \emptyset;$
$(f, h, g_0) \leftarrow \mathsf{Gen}(1^\kappa);$
$g := g_0; \mathsf{sk} := (f, h, g);$
Return $\mathsf{f}_b \leftarrow \mathsf{A}^{\mathcal{O}}(1^\kappa)$ $\quad\backslash\mathsf{MIM}$
$j \leftarrow [Q_{Ch}]; \ell := 0; \mathsf{rew} := 0;$ $\quad\backslash 1, \mathsf{U}$
$(\mathsf{f}_b, f_j, \mathsf{cmt}_\ell, f_\ell, \mathsf{rew}) \leftarrow \mathsf{A}^{\mathcal{O}}(1^\kappa);$ $\quad\backslash 1, \mathsf{U}$
$\mathsf{If}(\mathsf{rew} = 0)$ Return f_b $\quad\backslash 1, \mathsf{U}$
$\mathsf{If}(\mathsf{rew} = 2)$ $g_2 \leftarrow \mathsf{PIP}; g := g_2;$ $\quad\backslash 1, \mathsf{U}$
$\qquad (\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}) := (\perp, \perp);$ $\quad\backslash 1, \mathsf{U}$
$\qquad \mathsf{L}_\mathsf{P} \cup \mathsf{L}_\mathsf{V} := \emptyset;$ $\quad\backslash 1, \mathsf{U}$
$f'_j \leftarrow \mathsf{A}_{\mathsf{rew}}^{\mathcal{O}_{\mathsf{rew}}}(1^\kappa);$ $\quad\backslash 1, \mathsf{U}$
Return $f'_j = f_j$ $\quad\backslash 1, \mathsf{U}$

$\underline{\mathcal{O}_{\mathsf{P},Cmt}(\mathsf{m})}$:
$\mathsf{cmt} \leftarrow Cmt;$
$r_\mathsf{P} \leftarrow R;$
$u_\mathsf{P} \leftarrow \mathbb{F};$ $\quad\backslash\mathsf{U}$
$\mathsf{m}_\mathsf{P} := \mathsf{m};$
$\mathsf{cmt}_\mathsf{P} := \mathsf{cmt};$
Return cmt

$\underline{\mathcal{O}_{\mathsf{P},Cmt,2}(\mathsf{m})}$:
If $\quad \ell$th query
Then $\mathsf{cmt} := \mathsf{cmt}_\ell;$
Else $\mathsf{cmt} \leftarrow Cmt;$
$r_\mathsf{P} \leftarrow R;$
$u_\mathsf{P} \leftarrow \mathbb{F};$ $\quad\backslash\mathsf{U}$
$\mathsf{m}_\mathsf{P} := \mathsf{m};$
$\mathsf{cmt}_\mathsf{P} := \mathsf{cmt};$
Return cmt

$\underline{\mathcal{O}_{\mathsf{P},\mathsf{Rsp}}(\mathsf{ch})}$:
$\mathsf{If}((\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}) \in \mathsf{L}_\mathsf{P})$ Return $\perp$
$x := g(f(r_\mathsf{P}, \mathsf{cmt}_\mathsf{P}));$ $\quad\backslash\mathsf{MIM}, 1$
$x := g(u_\mathsf{P});$ $\quad\backslash\mathsf{U}$
$\mathsf{rsp} := x + \mathsf{ch} \cdot h(\mathsf{cmt}_\mathsf{P}, \mathsf{m}_\mathsf{P});$
$\mathsf{L}_\mathsf{P} := \mathsf{L}_\mathsf{P} \cup \{(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}, \mathsf{ch}, \mathsf{rsp})\};$
Return rsp

$\underline{\mathcal{O}_{\mathsf{P},\mathsf{Rsp},2}(\mathsf{ch})}$:
$\mathsf{If}((\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}) \in \mathsf{L}_\mathsf{P})$ Return $\perp$
If $\quad \mathsf{cmt}_\mathsf{P} = \mathsf{cmt}_\mathsf{A}$
Then $x := f_\ell;$
Else $\quad x := g(f(r_\mathsf{P}, \mathsf{cmt}_\mathsf{P}));$ $\quad\backslash 1$
$\qquad x := g(u_\mathsf{P});$ $\quad\backslash\mathsf{U}$
$\mathsf{rsp} := x + \mathsf{ch} \cdot h(\mathsf{cmt}_\mathsf{P}, \mathsf{m}_\mathsf{P});$
$\mathsf{L}_\mathsf{P} := \mathsf{L}_\mathsf{P} \cup \{(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}, \mathsf{ch}, \mathsf{rsp})\};$
Return rsp

$\underline{\mathcal{O}_{\mathsf{V},Ch}(\mathsf{m}, \mathsf{cmt})}$:
If $\quad j$th query
Then $\mathsf{m}_j := \mathsf{m}$
$\qquad \mathsf{cmt}_j := \mathsf{cmt}$
$\mathsf{ch} \leftarrow Ch$
$\mathsf{L}_\mathsf{V} := \mathsf{L}_\mathsf{V} \cup \{(\mathsf{m}, \mathsf{cmt}, \mathsf{ch})\};$
Return ch

$\underline{\mathcal{O}_{\mathsf{V},\mathsf{Vrfy}}(\mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp})}$:
$\mathsf{If}((\mathsf{m}, \mathsf{cmt}, \mathsf{ch}) \notin \mathsf{L}_\mathsf{V})$ Return $\perp$
$b := [\exists r \in R : \mathsf{rsp} = g(f(r, \mathsf{cmt})) + \mathsf{ch} \cdot h(\mathsf{cmt}, \mathsf{m})];$ $\quad\backslash\mathsf{MIM}$
$b := [(\mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp}) \in \mathsf{L}_\mathsf{P}];$ $\quad\backslash 1, \mathsf{U}$
If $\quad ((\mathsf{m}_j, \mathsf{cmt}_j) = (\mathsf{m}, \mathsf{cmt})) \wedge (\mathsf{cmt} \in \mathsf{L}_\mathsf{P})$ $\quad\backslash 1, \mathsf{U}$
Then $f_j := g^{-1}(\mathsf{rsp} - h(\mathsf{cmt}, \mathsf{m})\mathsf{ch});$ $\quad\backslash 1, \mathsf{U}$
$\qquad$ Set $\mathsf{m}', \mathsf{ch}', \mathsf{rsp}'$ s.t. $(\mathsf{m}', \mathsf{cmt}, \mathsf{ch}', \mathsf{rsp}') \in \mathsf{L}_\mathsf{P}$ $\quad\backslash 1, \mathsf{U}$
$\qquad f' := g^{-1}(\mathsf{rsp}' - h(\mathsf{cmt}, \mathsf{m}')\mathsf{ch}');$ $\quad\backslash 1, \mathsf{U}$
$\qquad b := [f = f'];$ $\quad\backslash 1, \mathsf{U}$
$\qquad \mathsf{If}$ $\quad b = 0$ $\quad\backslash 1, \mathsf{U}$
$\qquad \mathsf{Then}$ Set ℓ s.t. cmt was ℓth output of $\mathcal{O}_{\mathsf{P},Cmt}$ $\quad\backslash 1, \mathsf{U}$
$\qquad\qquad (\mathsf{cmt}_\ell, f_\ell, \mathsf{rew}) := (\mathsf{cmt}, f', 2);$ $\quad\backslash 1, \mathsf{U}$
If $\quad ((\mathsf{m}_j, \mathsf{cmt}_j) = (\mathsf{m}, \mathsf{cmt})) \wedge (\mathsf{cmt} \notin \mathsf{L}_\mathsf{P})$ $\quad\backslash 1, \mathsf{U}$
Then $f_j := g^{-1}(\mathsf{rsp} - h(\mathsf{cmt}, \mathsf{m})\mathsf{ch});$ $\quad\backslash 1, \mathsf{U}$
$\qquad \mathsf{rew} := 1$ $\quad\backslash 1, \mathsf{U}$
$\mathsf{f}_b := \mathsf{f}_b \vee (b \wedge [(\mathsf{m}, \mathsf{cmt}, \mathsf{ch}, \mathsf{rsp}) \notin \mathsf{L}_\mathsf{P}]);$
$\mathsf{L}_\mathsf{V} := \mathsf{L}_\mathsf{V} \setminus \{(\mathsf{m}, \mathsf{cmt}, \mathsf{ch})\};$
Return b

Fig. 2. The games $\mathsf{G}_{\mathsf{MIM}}$, G_1 and G_U. $(\mathsf{f}_b, f_j, \mathsf{cmt}_\ell, f_\ell, \mathsf{rew}) \leftarrow \mathsf{A}$ denotes the state of $(\mathsf{f}_b, f_j, \mathsf{cmt}_\ell, f_\ell, \mathsf{rew})$ after A's termination. While $\mathcal{O}_1$ is identical to $\mathcal{O}$, $\mathcal{O}_2$ replaces $\mathcal{O}_{\mathsf{P},Cmt}$, $\mathcal{O}_{\mathsf{P},\mathsf{Rsp}}$ with $\mathcal{O}_{\mathsf{P},Cmt,2}$, $\mathcal{O}_{\mathsf{P},\mathsf{Rsp},2}$. A_1 is A rewound to its state after the jth query $\mathsf{m}_j, \mathsf{cmt}_j$ is made to $\mathcal{O}_{\mathsf{V},Ch}$, where the state of P and V, $(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}, \mathsf{L}_\mathsf{P}, \mathsf{L}_\mathsf{V})$, is rewound as well. A_2 is identical to A.

G_1 uses T_0 til A makes the jth query to $\mathcal{O}_{Ch}$. Afterwards it uses T_1 til A's termination. T_2 is used during the execution of A_1. T_3, which is a part of T_0, is used to sample f and h of the secret key sk. For sampling g and during the run of A, T_4 is used, which consists of the remaining parts of $(\mathsf{T}_0, \mathsf{T}_1)$. G_1 uses T_5 to sample g_2 and during the run of A_2.

Now, we can model any interaction between A and the oracles as random variables in T. Of particular interest is A's jth query $(\mathsf{m}_j, \mathsf{cmt}_j)$ to $\mathcal{O}_{\mathsf{V},Ch}$, where random variable j is uniformly distributed over $[Q_{Ch}]$. We use ch_j to denote the output of this query and rsp_j is rsp of A's next query of the form $(\mathsf{m}_j, \mathsf{cmt}_j, \mathsf{ch}_j, \mathsf{rsp})$ to $\mathcal{O}_{\mathsf{V},\mathsf{Vrfy}}$. Here we assume that A does not make trivial queries where $\mathcal{O}_{\mathsf{V},\mathsf{Vrfy}}$ outputs $\perp$. We also consider $\mathcal{O}_{\mathsf{V},\mathsf{Vrfy}}$'s internal variable f_j as a random variable.

Additionally, we define random variable $X_1 \in \{0,1\}$ and $X_2 \in \{0,1\}$,

$$\begin{aligned} X_1 &:= [\exists r : \mathsf{rsp}_j = g(f_j) + \mathsf{ch}_j \cdot h(\mathsf{m}_j, \mathsf{cmt}_j) \wedge (\mathsf{m}_j, \mathsf{cmt}_j, \mathsf{ch}_j, \mathsf{rsp}_j) \notin \mathsf{L}_\mathsf{P}], \\ X_2 &:= [\mathsf{cmt}_j \in \mathsf{L}_\mathsf{P}]. \end{aligned}$$

Note that X_1 is identical to bit f_b of A running in $\mathsf{G}_{\mathsf{MIM}}$. X_2 determines which of the internal "if loops" of $\mathcal{O}_{\mathsf{V},\mathsf{Vrfy}}$ is entered. In particular, $X_2 = 1$ when cmt_j was generated by P and $X_2 = 0$ when cmt_j was generated by A. When $\mathsf{cmt}_j \in \mathsf{L}_\mathsf{P}$, i.e. $X_2 = 1$, then $\mathcal{O}_{\mathsf{P},\mathsf{Rsp}}$ has output a response for cmt_j and $\mathcal{O}_{\mathsf{V},\mathsf{Vrfy}}$ defines random variable f', which is a valid output of function f for input cmt_j.

Since the oracles $\mathcal{O}$, $\mathcal{O}_1$, $\mathcal{O}_2$ have the same output distribution, all the random variables that we have considered so far have the same distribution during the executions of A, A_1, A_2. Therefore, we use the same notation for the random variables, but consider the different underlying random tapes.

From the description of G_1 one can see that $\mathsf{G}_1(\mathsf{T}, \mathsf{A})$ outputs 1, if any of the three following conditions are met.

$$X_1(\mathsf{T}_0, \mathsf{T}_1) \wedge \neg X_2(\mathsf{T}_0) \wedge X_1(\mathsf{T}_0, \mathsf{T}_2) \wedge \big(f_j(\mathsf{T}_0, \mathsf{T}_1) = f_j(\mathsf{T}_0, \mathsf{T}_2)\big), \tag{1}$$

$$\begin{aligned} &X_1(\mathsf{T}_3, \mathsf{T}_4) \wedge X_2(\mathsf{T}_3, \mathsf{T}_4) \wedge X_1(\mathsf{T}_3, \mathsf{T}_5) \wedge X_2(\mathsf{T}_3, \mathsf{T}_5) \\ &\quad \wedge \big(f_j(\mathsf{T}_3, \mathsf{T}_5) = f_j(\mathsf{T}_3, \mathsf{T}_4) \neq f'(\mathsf{T}_3, \mathsf{T}_4)\big), \end{aligned} \tag{2}$$

$$X_1(\mathsf{T}_3, \mathsf{T}_4) \wedge X_2(\mathsf{T}_0) \wedge \big(f_j(\mathsf{T}_3, \mathsf{T}_4) = f'(\mathsf{T}_3, \mathsf{T}_4)\big). \tag{3}$$

In (1), $\mathsf{rew} = 1$ and A, A_1 output the same f_j for the same cmt_j, which is not contained in L_P, i.e. not an output of P. In (2), $\mathsf{rew} = 2$ and A, A_2 output the same f_j for the same cmt_j, which is in L_P, but differs from f' used by P's response for cmt_j. In (3), cmt_j and f_j are both identical to ones used by P, but $(\mathsf{m}_j, \mathsf{cmt}_j, \mathsf{ch}_j, \mathsf{rsp}_j)$ is not in L_P.

Remark that G_1 does not check if $\exists r : \mathsf{rsp}_j = g(f_j) + \mathsf{ch}_j \cdot h(\mathsf{m}_j, \mathsf{cmt}_j)$, therefore it might output 1 even if $X_1 = 0$. This is not a problem, since we lower bound the probability for $\mathsf{G}_1(\mathsf{A}) = 1$ with the probability for $\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1$. In the following, we consider each of the cases (1), (2), and (3) separately and combine the bounds in the end. $\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1$ implies that A makes a query for some m', cmt', ch', rsp' to $\mathcal{O}_{\mathsf{Vrfy}}$ such that there exists a $r' \in R$ such that

$$\mathsf{rsp}' = g(f(r', \mathsf{cmt}')) + \mathsf{ch}' \cdot h(\mathsf{m}', \mathsf{cmt}')),$$

and that will flip f_b from 0 to 1. It is easy to see, that f_b does not change anymore and hence we do not need to take care what happens afterwards.

This query to $\mathcal{O}_{\mathsf{Vrfy}}$ corresponds to a query $(\mathsf{m}', \mathsf{cmt}')$ to $\mathcal{O}_{Ch}$ with output ch'. Let this be the kth query to $\mathcal{O}_{Ch}$. For any A, k is a random variable with range $[Q_{Ch}]$. If G_1 guesses k correctly, i.e. $j = k$, the outcome of $\mathsf{G_{MIM}}$ is identical to X_1 for the same random tape.

We start with the first case, which refers to (1), to lower bound $\mathsf{G_1(A)} = 1$ when $\mathsf{cmt} \notin \mathsf{L_P}$, i.e. $\neg X_2(\mathsf{T}_0)$. By (1),

$$\begin{aligned}
&\Pr_{\mathsf{T}}[\mathsf{G}_1(\mathsf{T},\mathsf{A}) = 1 \mid \neg X_2(\mathsf{T}_0)] \\
&\geq \Pr[X_1(\mathsf{T}_0,\mathsf{T}_1) \wedge X_1(\mathsf{T}_0,\mathsf{T}_2) \wedge (f_j(\mathsf{T}_0,\mathsf{T}_1) = f_j(\mathsf{T}_0,\mathsf{T}_2)) \mid \neg X_2(\mathsf{T}_0)],
\end{aligned}$$

which we split to

$$\Pr[X_1(\mathsf{T}_0,\mathsf{T}_1) \wedge X_1(\mathsf{T}_0,\mathsf{T}_2) \mid \neg X_2(\mathsf{T}_0)] \tag{4}$$

$$\Pr[f_j(\mathsf{T}_0,\mathsf{T}_1) = f_j(\mathsf{T}_0,\mathsf{T}_2) \mid X_1(\mathsf{T}_0,\mathsf{T}_1), X_1(\mathsf{T}_0,\mathsf{T}_2), \neg X_2(\mathsf{T}_0)]. \tag{5}$$

Without loss of generality $0 \in \sup(X_2)$. Then, by Lemma 1

$$\Pr[k(\mathsf{T}_0,\mathsf{T}_1) = k(\mathsf{T}_0,\mathsf{T}_2) \mid \neg X_2(\mathsf{T}_0)] \geq \frac{1}{Q_{Ch}}.$$

Since j is independent and uniform over $[Q_{Ch}]$, this yields

$$\begin{aligned}
&\Pr[X_1(\mathsf{T}_0,\mathsf{T}_1) \wedge X_1(\mathsf{T}_0,\mathsf{T}_2) \mid \neg X_2(\mathsf{T}_0)] \\
&\geq \Pr[X_1(\mathsf{T}_0,\mathsf{T}_1) \wedge X_1(\mathsf{T}_0,\mathsf{T}_2) \wedge j = k(\mathsf{T}_0,\mathsf{T}_1) = k(\mathsf{T}_0,\mathsf{T}_2) \mid \neg X_2(\mathsf{T}_0)] \\
&\geq \frac{1}{Q_{Ch}^2} \Pr[X_1(\mathsf{T}_0,\mathsf{T}_1) \wedge X_1(\mathsf{T}_0,\mathsf{T}_2) \mid j = k(\mathsf{T}_0,\mathsf{T}_1) = k(\mathsf{T}_0,\mathsf{T}_2), \neg X_2(\mathsf{T}_0)].
\end{aligned}$$

$j = k(\mathsf{T}_0,\mathsf{T}_1) = k(\mathsf{T}_0,\mathsf{T}_2)$ means that k is guessed correctly for A and for A_2. Therefore, A, A_2 behave like in $\mathsf{G_{MIM}}$ and

$$\begin{aligned}
&\Pr[X_1(\mathsf{T}_0,\mathsf{T}_1) \wedge X_1(\mathsf{T}_0,\mathsf{T}_2) \mid j = k(\mathsf{T}_0,\mathsf{T}_1) = k(\mathsf{T}_0,\mathsf{T}_2), \neg X_2(\mathsf{T}_0)] \\
&= \Pr[\mathsf{G_{MIM}}(\mathsf{T}_0,\mathsf{T}_1,\mathsf{A}) = 1 \wedge \mathsf{G_{MIM}}(\mathsf{T}_0,\mathsf{T}_2,\mathsf{A}) = 1 \mid \neg X_2(\mathsf{T}_0)]
\end{aligned}$$

and by Jensen's Inequality

$$\begin{aligned}
&\Pr_{\mathsf{T}}[\mathsf{G_{MIM}}(\mathsf{T}_0,\mathsf{T}_1,\mathsf{A}) = 1 \wedge \mathsf{G_{MIM}}(\mathsf{T}_0,\mathsf{T}_2,\mathsf{A}) = 1 \mid \neg X_2(\mathsf{T}_0)] \\
&= \mathbb{E}_{\mathsf{T}_0,\mathsf{T}_1,\mathsf{T}_2}[\mathsf{G_{MIM}}(\mathsf{T}_0,\mathsf{T}_1,\mathsf{A}) = 1 \wedge \mathsf{G_{MIM}}(\mathsf{T}_0,\mathsf{T}_2,\mathsf{A}) = 1 \mid \neg X_2(\mathsf{T}_0)] \\
&= \mathbb{E}_{\mathsf{T}_0}\left(\mathbb{E}_{\mathsf{T}_1}[\mathsf{G_{MIM}}(\mathsf{T}_0,\mathsf{T}_1,\mathsf{A}) = 1 \mid \neg X_2(\mathsf{T}_0)]\right)^2 \\
&\geq \left(\mathbb{E}_{\mathsf{T}_0,\mathsf{T}_1}[\mathsf{G_{MIM}}(\mathsf{T}_0,\mathsf{T}_1,\mathsf{A}) = 1 \mid \neg X_2(\mathsf{T}_0)]\right)^2 \\
&= \Pr_{\mathsf{T}}[\mathsf{G_{MIM}}(\mathsf{A}) = 1 \mid \neg X_2(\mathsf{T}_0)]^2.
\end{aligned}$$

Since (4) is lower bounded by $\Pr[\mathsf{G_{MIM}}(\mathsf{A}) = 1 \mid \neg X_2(\mathsf{T}_0)]^2$, w.l.o.g. $1 \in \sup(X_1(\mathsf{T}_0,\mathsf{T}_1) \wedge X_1(\mathsf{T}_0,\mathsf{T}_2))$.

Notice that for $X_1(\mathsf{T}_0,\mathsf{T}_1) = X_1(\mathsf{T}_0,\mathsf{T}_2) = 1$, $\mathsf{rsp}_j(\mathsf{T}_0,\mathsf{T}_1)$ and $\mathsf{rsp}_j(\mathsf{T}_0,\mathsf{T}_2)$ are valid and hence $f_j(\mathsf{T}_0,\mathsf{T}_1)$ and $f(\mathsf{T}_0,\mathsf{T}_2)$ are valid outputs of function f for input cmt_j. Therefore f_j has a range of size of at most $|R|$, since there are at most $|R|$ many outputs of f for input cmt_j. By Lemma 1

$$\Pr[f_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = f_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_2) \mid X_1(\mathsf{T}_0,\mathsf{T}_1), X_1(\mathsf{T}_0,\mathsf{T}_2), \neg X_2(\mathsf{T}_0)] \geq \frac{1}{|R|}.$$

This yields the bound for case $\neg X_2(\mathsf{T}_0)$, i.e. (1), which is

$$\Pr_{\mathsf{T}}[\mathsf{G}_1(\mathsf{T},\mathsf{A})=1 \mid \neg X_2(\mathsf{T}_0)] \geq \frac{1}{|R|Q_{Ch}^2} \Pr_{\mathsf{T}}[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A})=1 \mid \neg X_2(\mathsf{T}_0)]^2.$$

In the second case, i.e. (2), we include the randomness for sampling cmt_ℓ and f_ℓ in random tape T_3. Then, the oracles $\mathcal{O}$ and $\mathcal{O}_2$ are identical. Notice that $X_2(\mathsf{T}_3,\mathsf{T}_4) \neq X_2(\mathsf{T}_3,\mathsf{T}_5)$ could hold, where by definition $(\mathsf{T}_3,\mathsf{T}_4)=(\mathsf{T}_0,\mathsf{T}_1)$. By Jensen's inequality, the probability that $X_2(\mathsf{T}_3,\mathsf{T}_4) = X_2(\mathsf{T}_3,\mathsf{T}_5)=1$, $f_j(\mathsf{T}_3,\mathsf{T}_4) \neq f_\ell(\mathsf{T}_3)$ and $f_j(\mathsf{T}_3,\mathsf{T}_5) \neq f_\ell(\mathsf{T}_3)$, which we will denote with

$$\begin{aligned} X_6(\mathsf{T}) := [&X_2(\mathsf{T}_3,\mathsf{T}_4) \wedge X_2(\mathsf{T}_3,\mathsf{T}_5) \wedge (f_j(\mathsf{T}_3,\mathsf{T}_4) \neq f_\ell(\mathsf{T}_3)) \\ &\wedge (f_\ell(\mathsf{T}_3,\mathsf{T}_5) \neq f_\ell(\mathsf{T}_3))], \end{aligned}$$

is sufficiently good:

$$\begin{aligned} &\Pr[X_6(\mathsf{T})] \\ &= \mathbb{E}_{\mathsf{T}_3}\left(\mathbb{E}_{\mathsf{T}_4}[X_2(\mathsf{T}_3,\mathsf{T}_4) \wedge (f_j(\mathsf{T}_3,\mathsf{T}_4) \neq f_\ell(\mathsf{T}_3))]\right)^2 \\ &\geq \Pr[X_2(\mathsf{T}_3,\mathsf{T}_4) \wedge (f_j(\mathsf{T}_3,\mathsf{T}_4) \neq f_\ell(\mathsf{T}_3))]^2. \end{aligned}$$

Now, we give a bound for this case, i.e. (2). By (2),

$$\begin{aligned} &\Pr_{\mathsf{T}}[\mathsf{G}_1(\mathsf{T},\mathsf{A})=1 \mid X_6(\mathsf{T})] \\ &\geq \Pr[X_1(\mathsf{T}_3,\mathsf{T}_4) \wedge X_1(\mathsf{T}_3,\mathsf{T}_5) \wedge (f_j(\mathsf{T}_3,\mathsf{T}_4) = f_j(\mathsf{T}_3,\mathsf{T}_5)) \mid X_6(\mathsf{T})], \end{aligned}$$

which splits into

$$\Pr[X_1(\mathsf{T}_3,\mathsf{T}_4) \wedge X_1(\mathsf{T}_3,\mathsf{T}_4) \mid X_6(\mathsf{T})] \tag{6}$$
$$\Pr[f_j(\mathsf{T}_3,\mathsf{T}_4) = f_j(\mathsf{T}_3,\mathsf{T}_5) \mid X_1(\mathsf{T}_3,\mathsf{T}_4), X_1(\mathsf{T}_3,\mathsf{T}_5), X_6(\mathsf{T})]. \tag{7}$$

As previously, we can bound (6) by

$$\begin{aligned} &\Pr[X_1(\mathsf{T}_3,\mathsf{T}_4) \wedge X_1(\mathsf{T}_3,\mathsf{T}_5) \mid X_6(\mathsf{T})] \\ &\geq \frac{1}{Q_{Ch}^2} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A})=1 \mid X_2(\mathsf{T}_3,\mathsf{T}_4), f_j(\mathsf{T}_3,\mathsf{T}_4) \neq f_\ell(\mathsf{T}_3)]^2. \end{aligned}$$

For (7), we consider random variable ℓ.

$$\begin{aligned} &\Pr[f_j(\mathsf{T}_3,\mathsf{T}_4) = f_\ell(\mathsf{T}_3,\mathsf{T}_5) \mid X_1(\mathsf{T}_3,\mathsf{T}_4), X_1(\mathsf{T}_3,\mathsf{T}_5), X_6(\mathsf{T})] \\ &\geq \Pr[(f_j(\mathsf{T}_3,\mathsf{T}_4) = f_j(\mathsf{T}_3,\mathsf{T}_5)) \wedge (\ell(\mathsf{T}_3,\mathsf{T}_4) = \ell(\mathsf{T}_3,\mathsf{T}_5)) \\ &\qquad \mid X_1(\mathsf{T}_3,\mathsf{T}_4), X_1(\mathsf{T}_3,\mathsf{T}_5), X_6(\mathsf{T})]. \end{aligned}$$

W.l.o.g., $1 \in \sup(X_1(\mathsf{T}_3,\mathsf{T}_4) \wedge X_1(\mathsf{T}_3,\mathsf{T}_5) \wedge X_6(\mathsf{T}))$. ℓ has a range of a size of at most Q_{Cmt}. By Lemma 1

$$\Pr_{\mathsf{T}}[\ell(\mathsf{T}_3,\mathsf{T}_4) = \ell(\mathsf{T}_3,\mathsf{T}_5) \mid X_1(\mathsf{T}_3,\mathsf{T}_4) \wedge X_1(\mathsf{T}_3,\mathsf{T}_5) \wedge X_6(\mathsf{T})] \geq \frac{1}{Q_{Cmt}}.$$

Further, under the condition $\ell(\mathsf{T}_3, \mathsf{T}_4) = \ell(\mathsf{T}_3, \mathsf{T}_5)$,

$$\mathsf{cmt}_j(\mathsf{T}_3, \mathsf{T}_4) = \mathsf{cmt}_\ell(\mathsf{T}_3, \mathsf{T}_4) = \mathsf{cmt}_\ell(\mathsf{T}_3, \mathsf{T}_5) = \mathsf{cmt}_j(\mathsf{T}_3, \mathsf{T}_5)$$

holds and therefore for $X_1(\mathsf{T}_3, \mathsf{T}_4) \wedge X_1(\mathsf{T}_3, \mathsf{T}_5)$, $f_j(\mathsf{T}_3, \mathsf{T}_4)$ and $f_j(\mathsf{T}_3, \mathsf{T}_5)$ have the same range of a size of at most $|R|$. W.l.o.g. $1 \in \mathsf{sup}(\ell(\mathsf{T}_3, \mathsf{T}_4) = \ell(\mathsf{T}_3, \mathsf{T}_5))$, which implies by Lemma 1

$$\begin{aligned}&\Pr[f_j(\mathsf{T}_3, \mathsf{T}_4) = f_j(\mathsf{T}_3, \mathsf{T}_5) \mid \ell(\mathsf{T}_3, \mathsf{T}_4) = \ell(\mathsf{T}_3, \mathsf{T}_5), X_1(\mathsf{T}_3, \mathsf{T}_4),\\ &\quad X_1(\mathsf{T}_3, \mathsf{T}_5), X_6(\mathsf{T})] \geq \frac{1}{|R|}.\end{aligned}$$

Hence, we receive a bound for (7)

$$\Pr[f_j(\mathsf{T}_3, \mathsf{T}_4) = f_j(\mathsf{T}_3, \mathsf{T}_5) \mid X_1(\mathsf{T}_3, \mathsf{T}_4), X_1(\mathsf{T}_3, \mathsf{T}_5), X_6(\mathsf{T})] \geq \frac{1}{|R| Q_{Cmt}}$$

and therefore also for (2)

$$\begin{aligned}&\Pr[\mathsf{G}_1(\mathsf{T}, \mathsf{A}) = 1 \mid X_6(\mathsf{T})]\\ &\geq \frac{1}{|R| Q_{Cmt} Q_{Ch}^2} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid X_2(\mathsf{T}_3, \mathsf{T}_4), f_j(\mathsf{T}_3, \mathsf{T}_4) \neq f_\ell(\mathsf{T}_3)]^2.\end{aligned}$$

In the third case, i.e. (3), where $X_2(\mathsf{T}_0) = 1$ and $f_j(\mathsf{T}_0, \mathsf{T}_1) = f_\ell(\mathsf{T}_0, \mathsf{T}_1)$, it is easy to obtain a lower bound, we do not even need to consider rewinding, just j needs to be guessed correctly.

$$\begin{aligned}&\Pr[\mathsf{G}_1(\mathsf{A}) = 1 \mid X_2(\mathsf{T}_0), f_j(\mathsf{T}_0, \mathsf{T}_1) = f_\ell(\mathsf{T}_0, \mathsf{T}_1)]\\ &\geq \frac{1}{Q_{Ch}} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid X_2(\mathsf{T}_0), f_j(\mathsf{T}_0, \mathsf{T}_1) = f_\ell(\mathsf{T}_0, \mathsf{T}_1)].\end{aligned}$$

For the final bounds, we define three probabilities

$$\begin{aligned}\rho_1 &:= \Pr_{\mathsf{T}}[\neg X_2(\mathsf{T}_0)],\\ \rho_2 &:= \Pr_{\mathsf{T}}[X_2(\mathsf{T}_0) \wedge (f_j(\mathsf{T}_0, \mathsf{T}_1) \neq f_\ell(\mathsf{T}_0, \mathsf{T}_1))],\\ \rho_3 &:= \Pr_{\mathsf{T}}[X_2(\mathsf{T}_0) \wedge (f_j(\mathsf{T}_0, \mathsf{T}_1) = f_\ell(\mathsf{T}_0, \mathsf{T}_1))].\end{aligned}$$

and a random variable

$$X_{1\vee 3}(\mathsf{T}_0, \mathsf{T}_1) := [\neg X_2(\mathsf{T}_0) \vee \big(X_2(\mathsf{T}_0) \wedge (f_j(\mathsf{T}_0, \mathsf{T}_1) = f_\ell(\mathsf{T}_0, \mathsf{T}_1))\big)].$$

Notice that $\neg X_{1\vee 3}(\mathsf{T}_0, \mathsf{T}_1) = [X_2(\mathsf{T}_0) \wedge (f_j(\mathsf{T}_0, \mathsf{T}_1) \neq f_\ell(\mathsf{T}_0, \mathsf{T}_1))]$. Using the three bounds, $\sum_{i=1}^3 \rho_i = 1$ and Lemma 2, we obtain

$$\begin{aligned}&(1 - \rho_2) \Pr[\mathsf{G}_1(\mathsf{A}) = 1 \mid X_{1\vee 3}]\\ &= \rho_1 \Pr[\mathsf{G}_1(\mathsf{A}) = 1 \mid \neg X_2] + \rho_3 \Pr[\mathsf{G}_1(\mathsf{A}) = 1 \mid X_2(\mathsf{T}_0) \wedge X_{1\vee 3}]\\ &\geq \frac{\rho_1}{|R| Q_{Ch}^2} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid \neg X_2]^2 + \frac{\rho_3}{Q_{Ch}} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid X_2 \wedge X_{1\vee 3}]\\ &\geq \frac{1}{|R| Q_{Ch}^2} (\rho_1 \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid \neg X_2]^2 + \rho_3 \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid X_2 \wedge X_{1\vee 3}]^2)\\ &\geq \frac{1 - \rho_2}{|R| Q_{Ch}^2} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid X_{1\vee 3}]^2.\end{aligned}$$

We apply again Lemma 2 and conclude with

$$
\begin{aligned}
&\Pr[\mathsf{G}_1(\mathsf{A}) = 1] \\
&= (1-\rho_2)\Pr[\mathsf{G}_1(\mathsf{A}) = 1 \mid X_{1\vee 3}] + \rho_2 \Pr[\mathsf{G}_1(\mathsf{A}) = 1 \mid \neg X_{1\vee 3}] \\
&\geq (1-\rho_2)\Pr[\mathsf{G}_1(\mathsf{A}) = 1 \mid X_{1\vee 3}] + \rho_2^2 \Pr[\mathsf{G}_1(\mathsf{A}) = 1 \mid X_{\mathsf{A},6}] \\
&\geq \frac{(1-\rho_2)^2 \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid X_{1\vee 3}]^2 + \rho_2^2 \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1 \mid \neg X_{1\vee 3}]^2}{|R| Q_{Cmt} Q_{Ch}^2} \\
&\geq \frac{1}{2|R| Q_{Cmt} Q_{Ch}^2} \Pr[\mathsf{G}_{\mathsf{MIM}}(\mathsf{A}) = 1]^2.
\end{aligned}
$$

This concludes the proof of this lemma. □

Now we can replace HIwPRF samples by uniform samples. Based on the assumption that f is a HIwPRF, an adversary cannot distinguish G_1 from G_U. The next lemma proves this observation and is identical to the lemmata of the previous proofs and therefore its proof is omitted.

Lemma 4. *For any algorithm* A*, there is a distinguisher* D *that breaks the* HIwPRF *property of* F *in time* $t_\mathsf{D} \leq 2t_\mathsf{A}$ *for* $Q \leq 2Q_{Cmt}$ *queries with success probability*

$$\epsilon_\mathsf{D} = |\Pr_\mathsf{T}[\mathsf{G}_1(\mathsf{T},\mathsf{A}) = 1] - \Pr_\mathsf{T}[\mathsf{G}_\mathsf{U}(\mathsf{T},\mathsf{A}) = 1]|,$$

where t_A *is the running time of* A*, and* Q_{Cmt} *a bound on its amount of queries to* $\mathcal{O}_{Cmt}$*.*

Proof. Notice that G_1 and G_U are almost identical with the difference that G_1 evaluates f and G_U uses uniform values instead.

We construct a distinguisher D which breaks the HIwPRF hardness assumption if an adversary A has a different success probability in G_1 and G_U. D receives access to an oracle $\mathcal{O}_?$ and needs to distinguish $\mathcal{O}_? = \mathcal{O}_\mathsf{F}$ from $\mathcal{O}_? = \mathcal{O}_\mathsf{U}$. D simply samples h, g_0, g_2 but not f. To evaluate f it uses its access to $\mathcal{O}_?$ and therefore it can answer queries to $\mathcal{O}_{Cmt}$ and $\mathcal{O}_\mathsf{Rsp}$. Due to the rewinding, D needs to query $\mathcal{O}_?$ at most $2Q_{Cmt}$ times. Everything else in G_1 can be done without knowing f.

It is easy to see that if $\mathcal{O}_? = \mathcal{O}_\mathsf{F}$, i.e $\mathcal{O}_?$ outputs samples of the form $\mathsf{cmt}, f(r, \mathsf{cmt})$ for $\mathsf{cmt} \leftarrow Cmt$, $r \leftarrow R$, D simulates G_1 and if $\mathbf{t}$ is uniform, D simulates G_U. Hence, $\Pr[\mathsf{G}_1(\mathsf{A}) = 1]$ is the probability that D outputs 1 when $\mathcal{O}_? = \mathcal{O}_\mathsf{F}$ and $\Pr[\mathsf{G}_\mathsf{U}(\mathsf{A}) = 1]$ is the probability that D outputs 1 when $\mathcal{O}_? = \mathcal{O}_\mathsf{U}$, which concludes the lemma. □

In G_U we exploit that h and g are hidden and an adversary needs to evaluate them in order to forge successfully. From an adversaries view, h and g are not determined and hence to evaluate them is infeasible such that the success probability of any algorithm is negligible in G_U.

Lemma 5. *For any algorithm* A,

$$\Pr_\mathsf{T}[\mathsf{G}_\mathsf{U}(\mathsf{T},\mathsf{A}) = 1] \leq \frac{2}{\mathbb{F}}.$$

Proof. During the execution of A, the oracle $\mathcal{O}_{\mathsf{P},\mathsf{Rsp}}$ has the same output distribution when outputting uniform outputs $u' \leftarrow \mathbb{F}$. The same holds for A_1 with one exception. Let $(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P})$ be the rewound state of P. Let rsp_P be the response for m_P, cmt_P and ch_P during A's execution. Then the response rsp'_P generated for the query m_P to $\mathcal{O}_{\mathsf{P},Cmt}$ with output cmt'_P and for a challenge ch'_P has distribution

$$\mathsf{rsp}'_\mathsf{P} = \mathsf{rsp}_\mathsf{P} + (\mathsf{ch}'_\mathsf{P} - \mathsf{ch}_\mathsf{P}) \cdot h(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}),$$

where rsp_j is uniform in $\mathbb{F}$. As for A_1, there is also a single exception for A_2. Let m_ℓ be the ℓth query to $\mathcal{O}_{\mathsf{P},Cmt}$ with output cmt_ℓ, rsp_ℓ and ch_ℓ the corresponding response and challenge during the execution of A. Let m'_ℓ, rsp'_ℓ, ch'_ℓ the to cmt_ℓ corresponding message, response and challenge during A_2's run. Then rsp'_ℓ has distribution

$$\mathsf{rsp}'_\ell = g_2(f_\ell) + \mathsf{ch}'_\ell \cdot h(\mathsf{m}'_\ell, \mathsf{cmt}_\ell),$$

where $f_\ell = g_0^{-1}(\mathsf{rsp}_\ell - h(\mathsf{m}_\ell, \mathsf{cmt}_\ell)\mathsf{ch}_\ell)$ and rsp_ℓ is uniform in $\mathbb{F}$. Therefore, A, A_1 learns at most output $h(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P})$ of h and A, A_2 learns at most output $g_2(g_0^{-1}(\mathsf{rsp}_\ell - h(\mathsf{m}_\ell, \mathsf{cmt}_\ell)\mathsf{ch}_\ell))$ of g_2.

Let us now consider the winning probability of A. When A_1 is invoked, i.e. $\mathsf{rew} = 1$, G_U outputs 1 if and only if $f_j = f'_j$, which corresponds to

$$\mathsf{rsp}_j - h(\mathsf{m}_j, \mathsf{cmt}_j)\mathsf{ch}_j = \mathsf{rsp}'_j - h(\mathsf{m}_j, \mathsf{cmt}_j)\mathsf{ch}'_j.$$

This determines $h(\mathsf{m}_j, \mathsf{cmt}_j) = (\mathsf{rsp}_j - \mathsf{rsp}'_j)/(\mathsf{ch}_j - \mathsf{ch}'_j)$, if $\mathsf{ch}'_j - \mathsf{ch}_j \neq 0$, which is the case with probability $1/|\mathbb{F}|$, since ch_j, ch'_j are outputs of $\mathcal{O}_{\mathsf{V},Ch}$. Further $\mathsf{cmt}_j \neq \mathsf{cmt}_\mathsf{P}$ and therefore, for any $\mathsf{rsp}_j, \mathsf{rsp}'_j, \mathsf{ch}_j, \mathsf{ch}'_j \neq \mathsf{ch}_j, h(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}) \in \mathbb{F}$, $(\mathsf{m}_j, \mathsf{cmt}_j), (\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P}) \neq (\mathsf{m}_j, \mathsf{cmt}_j) \in \mathbb{D}_\mathsf{H}$,

$$\Pr_{h \leftarrow \mathsf{H}}[h(\mathsf{m}_j, \mathsf{cmt}_j) = (\mathsf{ch}_j - \mathsf{ch}'_j)^{-1}(\mathsf{rsp}_j - \mathsf{rsp}'_j) \mid h(\mathsf{m}_\mathsf{P}, \mathsf{cmt}_\mathsf{P})] \leq \frac{1}{|\mathbb{F}|}$$

and hence, A wins in this case at most with probability $\frac{2}{|\mathbb{F}|}$.

When A_2 is invoked, i.e. $\mathsf{rew} = 2$, A wins if and only if $f'_j = f_j$, or similarly iff

$$g_2(f_j) = \mathsf{rsp}'_j - h(\mathsf{m}'_j, \mathsf{cmt}_\ell)\mathsf{ch}'_j.$$

For this case $f_\ell \neq f_j$ holds, but for any $f_j, f_\ell \neq f_j, \mathsf{rsp}'_j - h(\mathsf{m}'_j, \mathsf{cmt}_\ell)\mathsf{ch}'_j, u := \mathsf{rsp}'_\ell - h(\mathsf{m}'_\ell, \mathsf{cmt}_\ell)\mathsf{ch}'_\ell \in \mathbb{F}$, $(\mathsf{m}_j, \mathsf{cmt}_j) \in \mathbb{D}_\mathsf{H}$,

$$\Pr_{g_2 \leftarrow \mathsf{PIP}}[g_2(f_j) = \mathsf{rsp}'_j - h(\mathsf{m}'_j, \mathsf{cmt}_\ell)\mathsf{ch}'_j \mid g_2(f_\ell) = u] \leq \frac{1}{|\mathbb{F}| - 1},$$

which upper bounds A's winning probability by $\frac{2}{|\mathbb{F}|}$ in this case.

In the last case, where A is not rewound, i.e. $\mathsf{rew} = 0$, there is a tuple $(\mathsf{m}', \mathsf{cmt}_j, \mathsf{ch}', \mathsf{rsp})$ in set L_P. A wins if and only if

$$\mathsf{rsp}_j - h(\mathsf{m}_j, \mathsf{cmt}_j)\mathsf{ch}_j = \mathsf{rsp}' - h(\mathsf{m}', \mathsf{cmt}_j)\mathsf{ch}.$$

and $(\mathsf{m}_j, \mathsf{cmt}_j, \mathsf{ch}_j, \mathsf{rsp}_j) \notin \mathsf{L}_\mathsf{P}$. This implies that $\mathsf{m}_j \neq \mathsf{m}'$ or $\mathsf{ch}_j \neq \mathsf{ch}'$. We first handle the case $\mathsf{ch}_j \neq \mathsf{ch}'$ and $\mathsf{m}_j = \mathsf{m}'$. h information theoretically hidden, therefore A wins for any $\mathsf{rsp}_j, \mathsf{rsp}', \mathsf{ch}_j, \mathsf{ch}' \neq \mathsf{ch}_j \in \mathbb{F}$, $(\mathsf{m}_j, \mathsf{cmt}_j) \in \mathbb{D}_\mathsf{H}$ with probability

$$\Pr_{h \leftarrow \mathsf{H}}[h(\mathsf{m}_j, \mathsf{cmt}_j) = (\mathsf{ch}_j - \mathsf{ch}')^{-1}(\mathsf{rsp}_j - \mathsf{rsp}')] = \frac{1}{|\mathbb{F}|}.$$

Further, ch_j is sampled by $\mathcal{O}_{\mathsf{V},Ch}$ which implies $\mathsf{ch} = 0$ with probability $\frac{1}{|\mathbb{F}|}$. Hence, for any $\mathsf{m}', \mathsf{m}_j \neq \mathsf{m}' \in \mathcal{M}$, $\mathsf{cmt}_j \in Cmt$, $\mathsf{rsp}_j, \mathsf{rsp}', \mathsf{ch}', \mathsf{ch}_j \neq 0, u := h(\mathsf{m}', \mathsf{cmt}_j) \in \mathbb{F}$

$$\Pr_{h \leftarrow \mathsf{H}}[h(\mathsf{m}_j, \mathsf{cmt}_j) = \mathsf{ch}_j^{-1}(\mathsf{rsp}_j - \mathsf{rsp}' + \mathsf{ch}' \cdot u) \mid h(\mathsf{m}', \mathsf{cmt}_j) = u] = \frac{1}{|\mathbb{F}|}.$$

Hence, the success probability of any algorithm A is upper bounded by $\frac{2}{|\mathbb{F}|}$. □

□

6 Acknowledgements

We would like to thank Eike Kiltz for his advice and numerous discussions. We also thank Daniel Wichs for suggesting interactive MACs as a generalization of authentication protocols.

References

ACPS09. Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In Shai Halevi, editor, *CRYPTO 2009*, volume 5677 of *LNCS*, pages 595–618, Santa Barbara, CA, USA, August 16–20, 2009. Springer, Heidelberg, Germany. 2

Ale03. Michael Alekhnovich. More on average case vs approximation complexity. In *44th FOCS*, pages 298–307, Cambridge, MA, USA, October 11–14, 2003. IEEE Computer Society Press. 2

BFKL94. Avrim Blum, Merrick L. Furst, Michael J. Kearns, and Richard J. Lipton. Cryptographic primitives based on hard learning problems. In Douglas R. Stinson, editor, *CRYPTO'93*, volume 773 of *LNCS*, pages 278–291, Santa Barbara, CA, USA, August 22–26, 1994. Springer, Heidelberg, Germany. 1

BPR12. Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 719–737, Cambridge, UK, April 15–19, 2012. Springer, Heidelberg, Germany. 1

CKK08. Jacek Cichon, Marek Klonowski, and Miroslaw Kutylowski. Privacy protection for RFID with hidden subset identifiers. In *Pervasive Computing, 6th International Conference, Pervasive 2008, Sydney, Australia, May 19-22, 2008, Proceedings*, pages 298–314, 2008. 2, 15

CKT16. David Cash, Eike Kiltz, and Stefano Tessaro. Two-round man-in-the-middle security from LPN. In Eyal Kushilevitz and Tal Malkin, editors, *TCC 2016-A, Part I*, volume 9562 of *LNCS*, pages 225–248, Tel Aviv, Israel, January 10–13, 2016. Springer, Heidelberg, Germany. 1, 2

DKPW12. Yevgeniy Dodis, Eike Kiltz, Krzysztof Pietrzak, and Daniel Wichs. Message authentication, revisited. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 355–374, Cambridge, UK, April 15–19, 2012. Springer, Heidelberg, Germany. 1

DP14. Ivan Damgård and Sunoo Park. Towards optimally efficient secret-key authentication from PRG. Cryptology ePrint Archive, Report 2014/426, 2014. `http://eprint.iacr.org/2014/426`. 2

GGM86. Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *Journal of the ACM*, 33(4):792–807, October 1986. 1

GKL90. Oded Goldreich, Hugo Krawczyk, and Michael Luby. On the existence of pseudorandom generators. In Shafi Goldwasser, editor, *CRYPTO'88*, volume 403 of *LNCS*, pages 146–162, Santa Barbara, CA, USA, August 21–25, 1990. Springer, Heidelberg, Germany. 1

GRS08. Henri Gilbert, Matthew J. B. Robshaw, and Yannick Seurin. $\mathrm{HB}^\sharp$: Increasing the security and efficiency of HB^+. In Nigel P. Smart, editor, *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 361–378, Istanbul, Turkey, April 13–17, 2008. Springer, Heidelberg, Germany. 1

HB01. Nicholas J. Hopper and Manuel Blum. Secure human identification protocols. In Colin Boyd, editor, *ASIACRYPT 2001*, volume 2248 of *LNCS*, pages 52–66, Gold Coast, Australia, December 9–13, 2001. Springer, Heidelberg, Germany. 1

HKL+12. Stefan Heyse, Eike Kiltz, Vadim Lyubashevsky, Christof Paar, and Krzysztof Pietrzak. Lapin: An efficient authentication protocol based on ring-LPN. In Anne Canteaut, editor, *FSE 2012*, volume 7549 of *LNCS*, pages 346–365, Washington, DC, USA, March 19–21, 2012. Springer, Heidelberg, Germany. 1

JW05. Ari Juels and Stephen A. Weis. Authenticating pervasive devices with human protocols. In Victor Shoup, editor, *CRYPTO 2005*, volume 3621 of *LNCS*, pages 293–308, Santa Barbara, CA, USA, August 14–18, 2005. Springer, Heidelberg, Germany. 1

KH12. Matthias Krause and Matthias Hamann. The cryptographic power of random selection. In Ali Miri and Serge Vaudenay, editors, *SAC 2011*, volume 7118 of *LNCS*, pages 134–150, Toronto, Ontario, Canada, August 11–12, 2012. Springer, Heidelberg, Germany. 2, 15

KPC+11. Eike Kiltz, Krzysztof Pietrzak, David Cash, Abhishek Jain, and Daniele Venturi. Efficient authentication from hard learning problems. In Kenneth G. Paterson, editor, *EUROCRYPT 2011*, volume 6632 of *LNCS*, pages 7–26, Tallinn, Estonia, May 15–19, 2011. Springer, Heidelberg, Germany. 1

KS09. Matthias Krause and Dirk Stegemann. More on the security of linear RFID authentication protocols. In Michael J. Jacobson Jr., Vincent Rijmen, and Reihaneh Safavi-Naini, editors, *SAC 2009*, volume 5867 of *LNCS*, pages 182–196, Calgary, Alberta, Canada, August 13–14, 2009. Springer, Heidelberg, Germany. 15

KSS10. Jonathan Katz, Ji Sun Shin, and Adam Smith. Parallel and concurrent security of the HB and HB+ protocols. *Journal of Cryptology*, 23(3):402–421, July 2010. 1

LM13. Vadim Lyubashevsky and Daniel Masny. Man-in-the-middle secure authentication schemes from LPN and weak PRFs. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part II*, volume 8043 of *LNCS*, pages 308–325, Santa Barbara, CA, USA, August 18–22, 2013. Springer, Heidelberg, Germany. 1, 2, 16

NR04. Moni Naor and Omer Reingold. Number-theoretic constructions of efficient pseudo-random functions. *Journal of the ACM*, 51(2):231–262, 2004. 1

Reg05. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *37th ACM STOC*, pages 84–93, Baltimore, MA, USA, May 22–24, 2005. ACM Press. 2

A Random Selection

Historically, the CCK protocols [CKK08] were using this principle of randomly selecting a linear function to provide protection against adversaries in the RFID setting but without giving a security reduction. These protocols were broken by Krause and Stegemann [KS09]. They gave an exponential time algorithm to solve the problem of Learning Unions of L Linear Subspaces (LULS). Nevertheless, this algorithm efficiently recovered the secret key of the CCK protocols. Krause and Hamann [KH12] extended this work and introduced the Random Selection problem, which is a special case of the LULS problem. Random Selection is defined as follows.

Definition 4. [Random Selection]. *For parameters $n, m, \ell \in \mathbb{N}$, which are functions in κ, secrets $S_i \leftarrow \mathbb{Z}_2^{m \times n}$ for $i \in [\ell]$ and oracles $\mathcal{O}_{\mathsf{RS}}$, $\mathcal{O}_{\mathsf{U}}$,*

$\mathcal{O}_{\mathsf{RS}}$: *Sample $i \leftarrow [\ell]$, output $a \leftarrow \mathbb{Z}_2^n$, $t := S_i a$.*
$\mathcal{O}_{\mathsf{U}}$: *Output $a \leftarrow \mathbb{Z}_2^n$, $t \leftarrow \mathbb{Z}_2^m$.*

there is a search and a decisional problem:

An algorithm A *solves search* Random Selection *if*

$$\epsilon_{\mathsf{A}} = \Pr[\mathsf{A}^{\mathcal{O}_{\mathsf{RS}}}(1^\kappa) = (S_1, \ldots, S_\ell)] > negl(\kappa).$$

An algorithm D *solves decisional* Random Selection *if*

$$\epsilon_{\mathsf{D}} = |\Pr[\mathsf{D}^{\mathcal{O}_{\mathsf{RS}}}(1^\kappa) = 1] - \Pr[\mathsf{D}^{\mathcal{O}_{\mathsf{U}}}(1^\kappa) = 1]| > negl(\kappa).$$

The algebraic attack by [KH12] on search Random Selection has a running time of magnitude $2^{O(\ell \log n)}$, which is superpolynomial when n is linear and ℓ logarithmic in κ. Yet, it is not clear whether decisional Random Selection is as hard as its search variant. For constructing a HIwPRF, we need to rely on its decisional variant.

B Proof of Lemma 1 and 2

Proof (Proof of Lemma 1). The proof is straightforward using Jensen's Inequality.

$$\begin{aligned}
&\Pr[X_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = X_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_2) \mid X'_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = X'_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_2) = y] \\
&= \mathbb{E}_{\mathsf{T}_0,\mathsf{T}_1,\mathsf{T}_2}[X_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = X_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_2) \mid X'_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = X'_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_2) = y] \\
&= \sum_{s\in S} \mathbb{E}_{\mathsf{T}_0}(\mathbb{E}_{\mathsf{T}_1}[X_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = s \mid X'_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = y])^2 \\
&= |S|(\mathbb{E}_{s\leftarrow S,\mathsf{T}_0}(\mathbb{E}_{\mathsf{T}_1}[X_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = s \mid X'_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = y])^2) \\
&\geq |S|(\mathbb{E}_{s\leftarrow S,\mathsf{T}_0,\mathsf{T}_1}[X_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = s \mid X'_{\mathsf{A}}(\mathsf{T}_0,\mathsf{T}_1) = y])^2 = \frac{1}{|S|}.
\end{aligned}$$

□

Proof (Proof of Lemma 2). From the simple observation

$$\begin{aligned}
&(\Pr[E_0 \mid E_1] - \Pr[E_0 \mid \neg E_1])^2 \geq 0 \\
\Leftrightarrow\ &\Pr[E_0 \mid E_1]^2 + \Pr[E_0 \mid \neg E_1]^2 \geq 2\Pr[E_0 \mid E_1]\Pr[E_0 \mid \neg E_1],
\end{aligned}$$

we obtain for $\rho_1 = \Pr[E_1]$ and $\rho_2 = \Pr[\neg E_1]$

$$\begin{aligned}
&\rho_1 \Pr[E_0 \mid E_1]^2 + \rho_2 \Pr[E_0 \mid \neg E_1]^2 \\
&= \rho_1^2 \Pr[E_0 \mid E_1]^2 + \rho_1\rho_2(\Pr[E_0 \mid E_1]^2 + \Pr[E_0 \mid \neg E_1]^2) + \rho_2^2 \Pr[E_0 \mid \neg E_1]^2 \\
&\geq \rho_1^2 \Pr[E_0 \mid E_1]^2 + 2\rho_1\rho_2 \Pr[E_0 \mid E_1]\Pr[E_0 \mid \neg E_1] + \rho_2^2 \Pr[E_0 \mid \neg E_1]^2 \\
&= (\rho_1 \Pr[E_0 \mid E_1] + \rho_2 \Pr[E_0 \mid \neg E_1])^2 = \Pr[E_0]^2.
\end{aligned}$$

We use a similar approach to show the second inequality.

$$\begin{aligned}
&\rho_1^2 \Pr[E_0 \mid E_1]^2 + \rho_2^2 \Pr[E_0 \mid \neg E_1]^2 \\
&= \rho_1^2 \Pr[E_0 \mid E_1]^2 + \rho_2^2 \Pr[E_0 \mid \neg E_1]^2 + (1-1)\rho_1\rho_2 \Pr[E_0 \mid E_1]\Pr[E_0 \mid \neg E_1] \\
&= \frac{(\rho_1 \Pr[E_0 \mid E_1] + \rho_2 \Pr[E_0 \mid \neg E_1])^2}{2} + \frac{(\rho_1 \Pr[E_0 \mid E_1] - \rho_2 \Pr[E_0 \mid \neg E_1])^2}{2} \\
&\geq \frac{1}{2}\Pr[E_0]^2.
\end{aligned}$$

□

C Secure User Authentication

Once we have a secure message authentication, secure user authentication is trivial. One could simply use the message authentication protocol for a fixed message m. This will result in the protocol in Figure 3. The correctness and security of a message authentication protocol need to hold for any message m and will therefore also hold for any choice of the fixed message in the protocol of Figure 3. Thus, the protocol in Figure 3 is correct and secure based on the correctness and security of the applied message authentication protocol. When choosing the empty message, the protocol is besides the pairwise independent permutation very similar to the Man-in-the-Middle secure user authentication protocols of [LM13] for wPRFs, LPN and LWE.

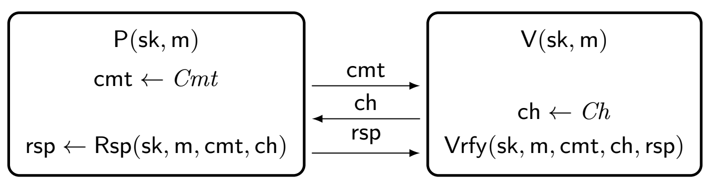

Fig. 3. A prover P and verifier V share a secret key sk and a fixed message m, which might be public. P will convince V that he is authentic by sending a valid response rsp such that Vrfy outputs 1.