Smooth NIZK Arguments with Applications to Asymmetric UC-PAKE and Threshold-IBE

Charanjit S. Jutla[1] and Arnab Roy[2]

[1] IBM T. J. Watson Research Center, Yorktown Heights, NY, USA
csjutla@us.ibm.com
[2] Fujitsu Laboratories of America, Sunnyvale, CA, USA
aroy@us.fujitsu.com

Abstract. We introduce a novel notion of smooth (-verifier) non- interactive zero-knowledge proofs (NIZK) which parallels the familiar notion of smooth projective hash functions (SPHF). We also show that the recent single group element quasi-adaptive NIZK (QA-NIZK) of Jutla and Roy (CRYPTO 2014) for linear subspaces can be easily extended to be computationally smooth. One important distinction of the new notion from SPHFs is that in a smooth NIZK the public evaluation of the hash on a language member using the projection key does not require the witness of the language member, but instead just requires its NIZK proof. This has the remarkable consequence that in the Gennaro-Lindell paradigm of designing universally-composable password-authenticated key-exchange (UC-PAKE) protocols, if one replaces the traditionally employed SPHFs with the novel smooth QA-NIZK, one gets highly efficient UC-PAKE protocols that are secure even under dynamic corruption. This simpler and modular design methodology allows us to give the first single-round asymmetric UC-PAKE protocol, which is also secure under dynamic corruption in the erasure model.
We also define a related concept of smooth signatures, which we show is black-box equivalent to identity-based encryption (IBE). The novel abstraction allows us to give the first threshold (private-key generation) fully-secure IBE in the standard model.

Keywords: Bilinear pairings, SXDH, MDDH, online attack, server compromise, dual-system, threshold IBE, QA-NIZK, UC-PAKE.

1 Introduction

Ever since the remarkably efficient non-interactive zero knowledge (NIZK) [BFM88] proofs for algebraic statements were developed by Groth and Sahai [GS08], there have been significant efficiency improvements and innovations in the construction of cryptographic protocols. Jutla and Roy [JR13, JR14] and Libert, Peters, Joye and Yung [LPJY14] further improved the efficiency of algebraic NIZK proofs, culminating in *constant* size NIZK proofs for linear subspaces, independent of the number of equations and witnesses. This efficiency improvement came in the weaker Quasi-Adaptive setting [JR13], which nevertheless proved sufficient for many applications.

Quasi-adaptive NIZK (QA-NIZK) proofs were further extended to provide simulation soundness [LPJY14, KW15] and dual-system simulation soundness [JR15], thus lending applicability to many more applications, such as, structure preserving signatures, password authenticated key exchange in the UC model and keyed homomorphic CCA-secure encryption.

In this paper, we further extend QA-NIZK proofs to provide an additional property called *smooth soundness*. The idea is to force the verification step to consist of computing hashes in two different ways and comparing the result. To this end, the verifier is split into three algorithms: a randomized hash key generation algorithm, a public hashing algorithm and a private hashing algorithm. The verification step starts off by generating two hash keys, the private key and the projection key. Next, the setting allows computation of a private hash given the private hash key and the word, and computation of the public hash using the projection key and just a QA-NIZK argument for the word - the witness for the word is *not* required. Completeness states that the private hash is equal to the public hash for a language member and correct QA-NIZK proof. Computational soundness states that it is hard to come up with a proof such that a non-language word passes the same equality check. The new *smoothness* property states that for any non-language word, the private hash algorithm outputs a value (computationally) indistinguishable from uniformly random, even when the projection key is given to the adversary.

Comparison with SPHFs. The new primitive is modeled after smooth projective hash functions (SPHF [CS02]). An SPHF also generates public and private hash keys and defines a private hash and a public hash. Further, similar properties hold where (1) for a member word, private hash equals public hash, (2) for a non-member word, private hash is uniformly random. The crucial difference is that, whereas the SPHF public hash computation requires a witness of the member word, the smooth QA-NIZK public hash requires only a NIZK proof of the word. This allows for hiding of the witness, even when using the public hash key. On the other hand, our constructions only allow computational smooth-soundness, while for SPHFs these properties hold information theoretically. [3]

In this work, we show that the single group element QA-NIZK arguments of [JR14] can be easily extended to be smooth. As a first application, we show that in the Gennaro-Lindell paradigm of designing universally-composable password-authenticated key exchange (UC-PAKE) protocols, if one replaces the traditional SPHFs with the novel smooth QA-NIZK, then one gets highly efficient single-round UC-PAKE protocols which are adaptively secure in the erasure model. At

[3] Trapdoor SPHFs as introduced by [BBC$^+$13] allow a simulation world to have a trapdoor to evaluate a hash over a word without a witness and without having full access to the private hash key. This is different from smooth NIZK proofs which allow public hashing in the real world without a witness, but instead requiring a NIZK proof.

a high level, the UC simulator must emulate each party's outgoing commitment to the password, without knowing the password. This is not difficult, as one can use El-Gamal encryption to achieve a hiding commitment. However, if the party is corrupted after its message has been sent, the simulator is at a loss to produce a witness which each party must retain to eventually compute the SPHF public hash. In our new protocol, the parties need only save the QA-NIZK and not the witness, as that suffices to compute the public hash. The zero-knowledge property of the QA-NIZK assures that the witness is not required at all by the simulator. This nicely captures the main idea behind the recent single-round adaptively secure UC-PAKE of [JR15]. In this work, the novel abstraction further allows us to obtain a single-round adaptively secure *asymmetric* UC-PAKE, which is a much more difficult notion to understand even from a definitional perspective, let alone constructing one. We will introduce in detail the subtleties of asymmetric PAKE in a few paragraphs, but before that we describe our second related contribution.

The QA-NIZK construction of [JR13] also led to standard model signature schemes as well as simpler dual-system [Wat09] identity-based encryption (IBE) schemes. However, although [JR13] gave a black box construction of signatures from any CCA2-encryption scheme and a QA-NIZK for the related ciphertexts, no such black-box construction was forthcoming for the stronger notion of IBE. Recall, Naor gave a simple black-box construction of signatures from IBE (see e.g. [BF01]). The question naturally arises whether smooth QA-NIZKs can enable one to give a similar black-box construction for IBE. However, in the IBE security notion, the adversary only supplies identities as input, and hence the decryption oracle of a CCA2-encryption scheme is useless. Thus, a dual-system argument [Wat09] such as a smooth version of dual-system simulation-sound DSS-QA-NIZK [JR15] seems essential. However, while interesting in its own merit, it is a highly complicated notion to be a high level primitive. Instead, in this work we define a notion of *smooth signatures* which is along the lines of the smooth QA-NIZK defined above. Recall, signatures can be seen as simulation-sound zero-knowledge proofs, but with a designated prover. Following Naor's argument, we show that smooth signatures are black-box *equivalent* to IBE (more precisely, IBE-KEM). We further show that smooth signatures (and hence fully-secure IBE) can be constructed from SPHFs and smooth QA-NIZKs of well-studied languages. With this high-level understanding, we can then give robust threshold smooth signature schemes, which then easily imply threshold full-secure IBE schemes in standard model under simple bilinear group assumptions such as SXDH, and MDDH. The threshold IBE scheme allows the master key to be shared among various parties, and the private keys of identities can then be generated in a robust threshold distributed fashion. The master-key shares can also be generated without a dealer using the standard techniques of [GJKR07]. This is the first fully-secure distributed private-key generation IBE scheme in the standard model. The construction is also highly efficient with ciphertexts consisting of only five group elements in a bilinear group under the SXDH assumption. The previous such schemes were either proven secure only in the random oracle model [BF01], or only provided selective-ID security [BBH06]. We remark here that the original dual-system fully-secure IBE scheme of Waters [Wat09] does not lend itself to be converted to a threshold scheme, and the advances in QA-NIZK are essential for our construction.

(Asymmetric) Password Authenticated Key Exchange. The problem of setting up a secure channel between two parties that only share a human-memorizable password (or a low-entropy secret) was first studied by Bellovin and Merritt [BM92], and Jiang and Gong [JG04]. Since then, this problem has been extensively studied and is called the *password-authenticated key-exchange* (PAKE) problem, while a protocol solving this problem is referred to as a PAKE protocol. Note that neither of the two parties is assumed to have a public key (for instance, if a public key infrastructure is not available or is considered insecure), and one of the main challenges in designing such protocols is the intricacy in the natural security definition which requires that the protocol transcripts cannot be used to launch *offline dictionary attacks*. While an adversary can clearly try to guess the (low-entropy) password and impersonate one of the parties, its advantage from the fact that the password is of low entropy should be limited to such *online* impersonation attacks. An example of an insecure protocol is one where the honest message flow includes a (deterministic) hash of the password, as then an adversary can launch an offline dictionary attack on the hash obtained from a single transcript.

In a subsequent paper, Bellovin and Merritt [BM93] also considered a stronger model of server compromise such that if a server's password file is revealed to the adversary it cannot directly impersonate a client (cf. if the password was stored in the raw at the server). The adversary should be able to impersonate the client only if it succeeds in an offline dictionary attack on the revealed server password file. Clearly, this requires that the server does not store the password as it is (or in some reversibly-encrypted form), and protocols satisfying this stronger security requirement are referred to as *asymmetric PAKE* protocols.

Canetti et al [CHK+05] also considered designing (symmetric) UC-PAKE protocols in the universally-composable (UC) framework [Can01]. One of their main contributions was the definition of a natural UC-PAKE ideal functionality ($\mathcal{F}_{\text{PAKE}}$). Gentry et al [GMR06] extended the functionality of symmetric UC-PAKE [CHK+05] to the asymmetric setting ($\mathcal{F}_{\mathsf{apwKE}}$) and gave a general method of extending any symmetric UC-PAKE protocol to an asymmetric UC-PAKE protocol (from now on referred to as UC-APAKE). Their general method adds an additional round to the UC-PAKE protocol. Moreover, their general two-round method requires that the environment somehow gets to know that in the UC-PAKE protocol both parties remain fresh, and this led them to define the functionality $\mathcal{F}_{\mathsf{apwKE}}$ to have additional TestAbort functions.

Our Contributions. In this paper, we give the first single-round UC-APAKE protocol (realizing $\mathcal{F}_{\mathsf{apwKE}}$). In fact, both parties just send a single message asynchronously. Since this is a single round protocol, we can realize $\mathcal{F}_{\mathsf{apwKE}}$ without the additional (and cumbersome) TestAbort function mentioned above.

The protocol is realized in the (limited programmability [FLR+10]) random-oracle [BR93] hybrid-model under standard static assumptions for bilinear groups, namely SXDH [BBS04] and the general MDDH assumption. Our protocol is also secure against adaptive corruption (in the erasure model) and is very succinct, with each message consisting of only four group elements. Moreover, for each client the server need store only *one* group elements as a "password hash". Many non-UC asymmetric PAKE protocols are at least two rounds [HK98, BPR00, BMP00, Mac01, Boy09]. Benhamouda and Pointcheval [BP13] proposed the first single round asymmetric PAKE protocol, but in a game-based model built on the BPR model [BPR00].

The first single-round UC-secure *symmetric* PAKE protocol was given in [KV11] (using bilinear pairings), which was then further improved (in the number of group elements) in subsequent papers [JR12, BBC+13]. Recently [JR15], a single round UC-PAKE protocol (in the standard model and using bilinear pairings) was also proven secure against adaptive corruption using ideas from the dual-system IBE construction of Waters [Wat09]. However, the [JR15] construction did not employ their Dual-System Simulation Sound QA-NIZK proofs (DSS-QA-NIZK) in a black box manner. Instead, it used ideas from the DSS-QA-NIZK construction and properties as the underlying intuition for the proof.

In this paper, we show that the UC-PAKE of [JR15] can be built in a black box manner using smooth QA-NIZK arguments. The proof only uses the definitional properties of the smooth QA-NIZK, without referring to its specific construction.

Next, we build on the Verifier-based PAKE (VPAKE) construction of [BP13], to construct a new single round UC-APAKE protocol. The intuition behind VPAKEs is as follows. Clearly, the server has to store some form of encryption or (probabilistic or deterministic) hash of the password, so that an adversary on obtaining this server password file has to, at the very least, perform offline tests to recover the password. It is not difficult to see that offline tests suffice as the following argument shows: consider an adversary that has obtained this hash of the password from the server password file. Next, it impersonates the client by guessing a password pw', and impersonates the server using the hash of the password that it has obtained, and checks if both ends compute the same session key to verify if pw' was the correct guess.

Unfortunately, in the UC framework, the simulator has to detect these offline password guesses by an adversary which steals the server password file, and for provable security this seems to inevitably require the random oracle model. Non-UC asymmetric PAKE protocols, do not suffer from the same drawback. In fact, the focus of [BP13] was to propose a security definition and constructions which could be proven secure in the standard model.

In our protocol, each party sends an ElGamal style encryption of the (hash of) the password pw to the other party, along with an SPHF of the underlying language and a projection verification hash key of a smooth QA-NIZK of the underlying language (augmented with the SPHF). If such a message is adversarially inserted, the simulator must have the capability to extract password pw'

from it, so that it can feed the ideal functionality $\mathcal{F}_{\mathsf{apwKE}}$ to test this guess of the password. Thus, the NIZK proof must have simulation-sound extractability. It was shown in [JR15] that dual-system simulation soundness suffices for this purpose (and that makes the protocol very simple). When using smooth QA-NIZK, this dual-system simulation-soundness can be attained by simply sending an SPHF (see above).

Detailed explanations can be found in Section 7 with proof details in Appendix D, where we also explain how the random oracle is used to extract the password efficiently from the exponent. This leads to a security reduction which has an additive computational overhead of $n * m * \mathrm{poly}(q)$, where n is the number of random oracle calls, m is the number of online attacks and q is the security parameter. We remark that the random oracle model uses only limited programmability as defined by Fischlin et al. [FLR$^+$10]. Basically, the output values of the random oracle are all randomly chosen, but different inputs can be assigned dynamically to these outputs.

The rest of the paper is organized as follows. In Section 2, we recall SPHFs. In Section 3, we introduce the new notion of smooth QA-NIZK proofs. In Section 4, we recall the MDDH assumptions and establish a useful technical theorem relating the assumptions. In Section 5, we give the single group element smooth QA-NIZK construction for linear subspaces. In Section 6, we define smooth signatures and give threshold constructions for smooth signatures and the implied threshold-IBE. In Section 7, we describe the ideal functionality $\mathcal{F}_{\mathsf{apwKE}}$ for asymmetric password-authenticated key-exchange and construct the new single-round UC-APAKE protocol. Proofs of many of the theorems are relegated to the Appendix.

2 Preliminaries: Smooth Projective Hash Functions

Since we are interested in distributions of languages, we extend the usual definition of smooth projective hash functions (SPHFs) [CS02] to distribution of languages. So consider a parametrized class of languages $\{L_\rho\}_{\rho \in \mathsf{Lpar}}$ with the parameters coming from an associated parameter language Lpar. An SPHF consists of the following efficient algorithms.

- $\mathsf{hkgen}(\rho)$, which generates two keys, a private key called hk, and a public key called hp.
- $\mathsf{privH}(\mathsf{hk}, x, l)$, computes a hash (in set Π) using the private key on input word x and label l.
- $\mathsf{pubH}(\mathsf{hp}, x, l; w)$ computes a hash (in set Π) using the public key on an input word x with witness w (for language L_ρ) and label l.

The correctness of SPHF family states that for all languages L_ρ in the parametrized class, for all $x \in L_\rho$ (with witness w), and for all labels l,

$$\mathsf{privH}(\mathsf{hk}, x, l) = \mathsf{pubH}(\mathsf{hp}, x, l; w).$$

A projective hash function family is called **smooth** if for all $x \notin L_\rho$, $\mathsf{privH}(\mathsf{hk}, x, l)$ is statistically indistinguishable from a random element in Π, even given hp. It is called **smooth₂** if for all $x \notin L_\rho$, $\mathsf{privH}(\mathsf{hk}, x, l)$ is statistically indistinguishable from a random element in Π, even given hp and one evaluation of $\mathsf{privH}(\mathsf{hk}, x^*, l^*)$ for any $(x^*, l^*) \neq (x, l)$.

3 Smooth Quasi-Adaptive NIZK Proofs

We start by reviewing the definition of Quasi-Adaptive computationally-sound NIZK proofs (QA-NIZK) [JR13]. A witness relation is a binary relation on pairs of inputs, the first called a (potential) language member and the second called a witness. Note that each witness relation R defines a corresponding language L which is the set of all x for which there exists a witness w, such that $R(x, w)$ holds.

We will consider QA-NIZK proofs for a probability distribution $\mathcal{D}$ on a collection of (witness-) relations $\mathcal{R} = \{R_\rho\}$ (with corresponding languages L_ρ). Recall that in a QA-NIZK, the CRS can be set after the language parameter has been chosen according to $\mathcal{D}$. Please refer to [JR13] for detailed definitions.

Definition 1 (QA-NIZK [JR13]). *We call a tuple of efficient algorithms* $(\mathsf{pargen}, \mathsf{crsgen}, \mathsf{prover}, \mathsf{ver})$ *a* quasi-adaptive non-interactive zero-knowledge *(QA-NIZK) proof system for witness-relations* $\mathcal{R}_\lambda = \{R_\rho\}$ *with parameters sampled from a distribution* $\mathcal{D}$ *over associated parameter language* Lpar*, if there exist simulators* crssim *and* sim *such that for all non-uniform PPT adversaries* $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$*, we have (in all of the following probabilistic experiments, the experiment starts by setting* λ *as* $\lambda \leftarrow \mathsf{pargen}(1^m)$*, and choosing* ρ *as* $\rho \leftarrow \mathcal{D}_\lambda$*):*

Quasi-Adaptive Completeness:

$$\Pr\begin{bmatrix} \textsc{crs} \leftarrow \mathsf{crsgen}(\lambda, \rho) \\ (x, w) \leftarrow \mathcal{A}_1(\textsc{crs}, \rho) \\ \pi \leftarrow \mathsf{prover}(\textsc{crs}, x, w) \end{bmatrix} : \begin{matrix} \mathsf{ver}(\textsc{crs}, x, \pi) = 1 \textbf{ if} \\ R_\rho(x, w) \end{matrix} \Bigg] = 1$$

Quasi-Adaptive Soundness:

$$\Pr\begin{bmatrix} \textsc{crs} \leftarrow \mathsf{crsgen}(\lambda, \rho) & : & x \notin L_\rho \textbf{ and} \\ (x, \pi) \leftarrow \mathcal{A}_2(\textsc{crs}, \rho) & : & \mathsf{ver}(\textsc{crs}, x, \pi) = 1] \end{bmatrix} \approx 0$$

Quasi-Adaptive Zero-Knowledge:

$$\Pr\left[\textsc{crs} \leftarrow \mathsf{crsgen}(\lambda, \rho) \ : \ \mathcal{A}_3^{\mathsf{prover}(\textsc{crs}, \cdot, \cdot)}(\textsc{crs}, \rho) = 1\right]$$
$$\approx$$
$$\Pr\left[(\textsc{crs}, \mathsf{trap}) \leftarrow \mathsf{crssim}(\lambda, \rho) \ : \ \mathcal{A}_3^{\mathsf{sim}^*(\textsc{crs}, \mathsf{trap}, \cdot, \cdot)}(\textsc{crs}, \rho) = 1\right],$$

where $\mathsf{sim}^*(\textsc{crs}, \mathsf{trap}, x, w) = \mathsf{sim}(\textsc{crs}, \mathsf{trap}, x)$ *for* $(x, w) \in R_\rho$ *and both oracles (i.e.* prover *and* sim^**) output failure if* $(x, w) \notin R_\rho$*.*

We call a QA-NIZK **smooth (-verifier)** if the verifier ver consists of three efficient algorithms $\mathsf{ver} = (\mathsf{hkgen}, \mathsf{pubH}, \mathsf{privH})$, and it satisfies the following modified completeness and soundness conditions. Here, hkgen is a probabilistic algorithm that takes a CRS as input and outputs two keys, hp, a projection hash key, and hk, a private hash key. The algorithm privH takes as input a word (e.g. a potential language member), and a (private hash) key, and outputs a string. Similarly, the algorithm pubH takes as input a word, a proof (for instance generated by prover), and a (projection hash) key hp, and outputs a string.

The above **completeness** property is now defined as:

$$\Pr\left[\begin{array}{l}\textsc{crs} \leftarrow \mathsf{crsgen}(\lambda, \rho)\\ (x, w) \leftarrow \mathcal{A}_1(\textsc{crs}, \rho)\\ \pi \leftarrow \mathsf{prover}(\textsc{crs}, x, w)\\ (\mathsf{hp}, \mathsf{hk}) \leftarrow \mathsf{hkgen}(\textsc{crs})\end{array} : \begin{array}{c}\mathsf{privH}(\mathsf{hk}, x) = \mathsf{pubH}(\mathsf{hp}, x, \pi)\\ \textbf{if } R_\rho(x, w)\end{array}\right] = 1$$

The QA-NIZK is said to satisfy **smooth-soundness** if for all words $x \notin L_\rho$, $\mathsf{privH}(\mathsf{hk}, x)$ is computationally indistinguishable to the Adversary from uniformly random, even when the Adversary is given hp, and even if it produces x after receiving hp.

More precisely, **Quasi-Adaptive Smooth-Soundness** is the following property (let $\mathcal{U}$ be the uniform distribution on the range of privH, which is assumed to be of cardinality exponential in m): for every two-stage efficient oracle adversary $\mathcal{A}$

$$\Pr\left[\begin{array}{l}\textsc{crs} \leftarrow \mathsf{crsgen}(\lambda, \rho)\\ (\mathsf{hp}, \mathsf{hk}) \leftarrow \mathsf{hkgen}(\textsc{crs})\\ (x^*, \sigma) \leftarrow \mathcal{A}^{\mathcal{O}}(\textsc{crs}, \rho, \mathsf{hp})\\ u \leftarrow \mathcal{U}\end{array} : \mathcal{A}^{\mathcal{O}}(\mathsf{privH}(\mathsf{hk}, x^*), \sigma) = 1 \mid Q\right]$$

$$\approx$$

$$\Pr\left[\begin{array}{l}\textsc{crs} \leftarrow \mathsf{crsgen}(\lambda, \rho)\\ (\mathsf{hp}, \mathsf{hk}) \leftarrow \mathsf{hkgen}(\textsc{crs})\\ (x^*, \sigma) \leftarrow \mathcal{A}^{\mathcal{O}}(\textsc{crs}, \rho, \mathsf{hp})\\ u \leftarrow \mathcal{U}\end{array} : \mathcal{A}^{\mathcal{O}}(u, \sigma) = 1 \mid Q\right]$$

where the oracle $\mathcal{O}$ is instantiated with $\mathsf{privH}(\mathsf{hk}, \cdot)$, and Q is the condition that x^* is not in the language L_ρ **and** all oracle calls by the adversary in both stages are with L_ρ-language members. Here, σ is a local state of $\mathcal{A}$.

Note that as opposed to the information-theoretic smoothness property of projective hash functions, one cannot argue here that $\mathsf{privH}(\mathsf{hk}, x)$ for $x \in L_\rho$ can instead just be computed using hp, as that would also require efficiently computing a witness for x. Hence, the need to provide oracle access to $\mathsf{privH}(\mathsf{hk}, \cdot)$ for language members.

Also, note that smooth-soundness implies the earlier definition of soundness [JR13] if verification of (x, π) is defined as $\mathsf{privH}(\mathsf{hk}, x) = \mathsf{pubH}(\mathsf{hp}, \pi)$.

To differentiate the functionalities of the verifier of a QA-NIZK from similar functionalities of an SPHF, we will prepend the SPHF functionalities with keyword sphf and the QA-NIZK verifier functionalities with the keyword ver.

4 Matrix Decisional Assumptions

We will consider bilinear groups that consist of three cyclic groups of prime order q, $\mathbb{G}_1, \mathbb{G}_2$ and $\mathbb{G}_T$ with an efficient bilinear map $e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_T$. Group elements $\mathbf{g}_1$ and $\mathbf{g}_2$ will typically denote generators of the group $\mathbb{G}_1$ and $\mathbb{G}_2$ respectively. Following [EHK$^+$13], in this section and the next we will use the notations $[a]_1, [a]_2$ and $[a]_T$ to denote $a\mathbf{g}_1, a\mathbf{g}_2$, and $a \cdot e(\mathbf{g}_1, \mathbf{g}_2)$ respectively and use additive notations for group operations. When talking about a general group $\mathbb{G}$ with generator $\mathbf{g}$, we will just use the notation $[a]$ to denote $a\mathbf{g}$. However, in the UC-APAKE constructions, we will switch to multiplicative notation for easy readability.

For two vector or matrices A and B, we will denote the product $A^\top B$ as $A \cdot B$. The pairing product $\mathsf{e}([A]_1, [B]_2)$ evaluates to the matrix product $[AB]_T$ in the target group with pairing as multiplication and target group operation as addition.

We recall the *Matrix Decisional Diffie Hellman* or MDDH assumptions from [EHK$^+$13]. A matrix distribution $\mathcal{D}_{l,k}$, where $l > k$, is defined to be an efficiently samplable distribution on $\mathbb{Z}_q^{l \times k}$ which is full-ranked with overwhelming probability. The $\mathcal{D}_{l,k}$*-MDDH assumption* in group $\mathbb{G}$ states that with samples $\mathbf{A} \leftarrow \mathcal{D}_{l,k}$ and $(\mathbf{s},\ \mathbf{s}') \leftarrow \mathbb{Z}_q^k \times \mathbb{Z}_q^l$, the tuple $([\mathbf{A}], [\mathbf{As}])$ is computationally indistinguishable from $([\mathbf{A}], [\mathbf{s}'])$. A matrix distribution $\mathcal{D}_{k+1,k}$ is simply denoted by $\mathcal{D}_k$.

Intuitively, a $\mathcal{D}_{l,k}$-MDDH assumption allows us to generate l (computationally) independently random group elements from an initial k independently random exponents. A $\mathcal{D}_k$-MDDH assumption allows us to generate one extra random group element. In this section, we will establish that, in fact, a $\mathcal{D}_k$-MDDH assumption can be *boosted* to generate additional (computationally) independently random elements. This will be useful to us in the next section to prove the smoothness property of our construction.

We remark that boosting is different from the random self-reducibility of $\mathcal{D}_{l,k}$-MDDH assumptions, as described by [EHK$^+$13]. While the former aims to generate extra randomness from the same initial sample of vector of random exponents, the latter talks about results from several independent samples of vector of random exponents. Boosting can be seen as an abstraction of the *switching lemma* of [JR14] and follows the same blueprint for the proof.

For an $l \times k$ matrix $\mathbf{A}$, we denote $\bar{\mathbf{A}}$ to be the top $k \times k$ square sub-matrix of $\mathbf{A}$ and $\underline{\mathbf{A}}$ to be the bottom $(l-k) \times k$ sub-matrix of $\mathbf{A}$.

Definition 2. *We say that a matrix distribution $\mathcal{D}_k$ on $\mathbb{Z}_q^{(k+1)\times k}$ is* boostable *to a matrix distribution $\mathcal{D}_{l,k}$ on $\mathbb{Z}_q^{l\times k}$, where $l > k$, if there are efficiently samplable distributions $\mathcal{E}$ on $\mathbb{Z}_q^{(l-k)\times k}$ and $\mathcal{F}$ on $\mathbb{Z}_q^{(l-k)\times(k+1)}$, such that the following hold:*

- *For $\mathbf{A} \leftarrow \mathcal{D}_k, \mathbf{B} \leftarrow \mathcal{D}_{l,k}, \mathbf{E} \leftarrow \mathcal{E}, \mathbf{F} \leftarrow \mathcal{F}$, we have:*

$$\bar{\mathbf{B}} \approx \bar{\mathbf{A}}, \quad \underline{\mathbf{B}} \approx \mathbf{E}\bar{\mathbf{A}} \approx \mathbf{F}\mathbf{A}.$$

- *For $\mathbf{F} \leftarrow \mathcal{F}$, with overwhelming probability, all entries of the rightmost column $\mathbf{F}_r$ of $\mathbf{F}$ are non-zero.*

Theorem 1. *If a matrix distribution $\mathcal{D}_k$ on $\mathbb{Z}_q^{(k+1)\times k}$ is boostable to a matrix distribution $\mathcal{D}_{l,k}$ on $\mathbb{Z}_q^{l\times k}$ then the $\mathcal{D}_k$-MDDH assumption implies the $\mathcal{D}_{l,k}$-MDDH assumption.*

Proof of this theorem can be found in Appendix A.

Corollary 1. *Any $\mathcal{D}_k$ distribution can be boosted to a $\mathcal{D}_{l,k}$ distribution which inherits the distribution of the top $k \times k$ matrix of the samples.*

This can be seen by setting the top $k \times k$ matrix of a $\mathcal{D}_{l,k}$ sample to be the top $k \times k$ matrix of a $\mathcal{D}_k$ sample and setting the bottom $(l-k) \times k$ sub-matrix of the $\mathcal{D}_{l,k}$ sample to be uniformly random in $\mathbb{Z}_q^{(l-k)\times k}$. The required distributions $\mathcal{E}$ and $\mathcal{F}$ are just the uniform distributions on their respective domains.

This corollary allows us to retain the *representation size* of the top square matrix of a $\mathcal{D}_k$ distribution sample, while boosting it to an assumption required for security proofs. In particular, in applications such as this paper, this can lead to shorter public keys.

5 Smooth Quasi-Adaptive NIZK Constructions

In this section we show that the single element QA-NIZK [JR14, KW15] for witness-samplable linear subspaces can easily be extended to be smooth QA-NIZK. Particularly, under SXDH, the public hash key hp generated by $\mathsf{ver.hkgen}$ consists of a single group element. We follow the notations of Kiltz and Wee [KW15] and prove the result under the more general MDDH assumption in bilinear groups.

We follow additive notation for group operations in this section. In later sections we will use product notation.

5.1 Linear Subspace Languages

We consider languages that are linear subspaces of vectors of $\mathbb{G}_1$ elements. In other words, the languages we are interested in can be characterized as languages parametrized by $[\mathbf{M}]_1$ as below:

$$L_{[\mathbf{M}]_1} = \{[\mathbf{M}]_1\mathbf{x} \in \mathbb{G}_1^n \mid \mathbf{x} \in \mathbb{Z}_q^t\}, \text{ where } [\mathbf{M}]_1 \text{ is an } n \times t \text{ matrix of } \mathbb{G}_1 \text{ elements.}$$

Here $[\mathbf{M}]_1$ is an element of the associated *parameter language* Lpar, which is all $n \times t$ matrices of $\mathbb{G}_1$ elements. The parameter language Lpar also has a corresponding witness relation $\mathcal{R}_{\mathrm{par}}$, where the witness is a matrix of $\mathbb{Z}_q$ elements : $\mathcal{R}_{\mathrm{par}}([\mathbf{M}]_1, \mathbf{M}')$ iff $\mathbf{M} = \mathbf{M}'$.

Robust and Efficiently Witness-Samplable Distributions. Let the $t \times n$ dimensional matrix $[\mathbf{M}]_1$ be chosen according to a distribution $\mathcal{D}$ on Lpar. The distribution $\mathcal{D}$ is called *robust* if with probability close to one the left-most t columns of $[\mathbf{M}]_1$ are full-ranked. A distribution $\mathcal{D}$ on Lpar is called *efficiently* *witness-samplable* if there is a probabilistic polynomial time algorithm such that it outputs a pair of matrices $([\mathbf{M}]_1, \mathbf{M}')$ that satisfy the relation $\mathcal{R}_{\text{par}}$ (i.e., $\mathcal{R}_{\text{par}}([\mathbf{M}]_1, \mathbf{M}')$ holds), and further the resulting distribution of the output $[\mathbf{M}]_1$ is same as $\mathcal{D}$. For example, the uniform distribution on Lpar is efficiently witness-samplable, by first picking $\mathbf{M}$ at random, and then computing $[\mathbf{M}]_1$.

Smooth QA-NIZK Construction. We now describe a smooth computationally-sound Quasi-Adaptive NIZK $(\mathsf{K}_0, \mathsf{K}_1, \mathsf{P}, \mathsf{V})$ for linear subspace languages $\{L_{[\mathbf{M}]_1}\}$ with parameters sampled from a robust and efficiently witness-samplable distribution $\mathcal{D}$ over the associated parameter language Lpar and given a $\mathcal{D}_k$-MDDH assumption.

Algorithm K_1: The algorithm K_1 generates the CRS as follows. Let $[\mathbf{M}^{n\times t}]_1$ be the parameter supplied to K_1. It generates an $n \times k$ matrix $\mathbf{K}$ with all elements chosen randomly from $\mathbb{Z}_q$ and a $(k+1)\times k$ matrix $\mathbf{A}$ from the MDDH distribution $\mathcal{D}_k$. Let $\bar{\mathbf{A}}$ be the top $k \times k$ square matrix of $\mathbf{A}$.

The common reference string (CRS) has two parts $\mathbf{CRS}_p$ and $\mathbf{CRS}_v$ which are to be used by the prover and the verifier respectively.

$$\mathbf{CRS}_p^{t\times k} := ([\mathbf{P}]_1 = [\mathbf{M}^\top \mathbf{K}]_1) \qquad\qquad \mathbf{CRS}_v := ([\mathbf{C}]_2^{n\times k} = [\mathbf{K}\bar{\mathbf{A}}]_2, \quad [\bar{\mathbf{A}}]_2^{k\times k})$$

Prover P**:** Given candidate $[\mathbf{y}]_1 = [\mathbf{M}]_1\mathbf{x}$ with witness vector $\mathbf{x}^{t\times 1}$, the prover generates the following proof consisting of k elements in $\mathbb{G}_1$:

$$\pi := \mathbf{x}^\top \mathbf{CRS}_p$$

Verifier V**:** The algorithm hkgen is as follows: Sample $\mathbf{s} \leftarrow \mathbb{Z}_q^k$. Given $\mathbf{CRS}_v$ as above, compute hk and hp as follows:

$$\mathsf{hk} := [\mathbf{C}]_2\ \mathbf{s}, \qquad \mathsf{hp} := [\bar{\mathbf{A}}]_2\ \mathbf{s}$$

The algorithms pubH and privH are as follows: Given candidate $[\mathbf{y}]_1$, and proof π, compute:

$$\mathsf{privH}(\mathsf{hk}, [\mathbf{y}]_1) := e([\mathbf{y}^\top]_1, \mathsf{hk}) \qquad \mathsf{pubH}(\mathsf{hp}, \pi) := e(\pi, \mathsf{hp})$$

Theorem 2. *The above algorithms* $(\mathsf{K}_0, \mathsf{K}_1, \mathsf{P}, \mathsf{V})$ *constitute a smooth computationally -sound Quasi-Adaptive NIZK proof system for linear subspace languages* $\{L_{[\mathbf{M}]_1}\}$ *with parameters* $[\mathbf{M}]_1$ *sampled from a robust and efficiently witness-samplable distribution* $\mathcal{D}$ *over the associated parameter language* Lpar*, given any group generation algorithm for which the* $\mathcal{D}_k$*-MDDH assumption holds for group* $\mathbb{G}_2$*.*

The proofs of completeness, zero knowledge and soundness are same as [KW15]. The proof of smooth soundness is given in Appendix B.

5.2 Smooth Split-CRS QA-NIZK for Tagged Affine Languages.

We now describe a smooth computationally-sound Quasi-Adaptive NIZK $(\mathsf{K}_0, \mathsf{K}_1, \mathsf{P}, \mathsf{V})$ for tagged affine linear subspace languages $\{L_{\mathbf{M}_0,\mathbf{M}_1,\mathbf{M}_2,\mathbf{M}_3,\mathbf{a}}\}$, consisting of words of the form

$$([\mathbf{M}_0\mathbf{x}]_1, [\mathbf{M}_1\mathbf{x} + \mathbf{a}]_1, [(\mathbf{M}_2 + \text{TAG}.\mathbf{M}_3)\mathbf{x}]_1,\ \text{TAG}),$$

with parameters sampled from a robust and efficiently witness-samplable distribution $\mathcal{D}$ over $(\mathbf{M}_0, \mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{a})$ and given a $\mathcal{D}_k$-MDDH assumption. We assume that $\mathbf{M}_0$ is a square matrix and robustness of $\mathcal{D}$ is defined by $\mathbf{M}_0$ being non-singular. The smooth QA-NIZK will be split-CRS [JR13], so that CRS_v is independent of the language parameters.

Algorithm K_1: The algorithm K_1 generates the CRS as follows. Let

$$(\mathbf{M}_0, \mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{a}) \in \mathbb{Z}_q^{t\times t} \times \mathbb{Z}_q^{l\times t} \times \mathbb{Z}_q^{l'\times t} \times \mathbb{Z}_q^{l'\times t} \times \mathbb{Z}_q^{l},$$

be the parameter supplied to K_1. It generates matrices:

$$(\mathbf{K}_1, \mathbf{K}_2, \mathbf{L}_1, \mathbf{L}_2, \mathbf{l}_3) \leftarrow \mathbb{Z}_q^{l\times k} \times \mathbb{Z}_q^{l'\times k} \times \mathbb{Z}_q^{t\times k} \times \mathbb{Z}_q^{t\times k} \times \mathbb{Z}_q^{k}$$

and a $(k+1)\times k$ matrix $\mathbf{A}$ from the MDDH distribution $\mathcal{D}_k$. Let $\bar{\mathbf{A}}$ be the top $k \times k$ square matrix of $\mathbf{A}$.

The common reference string (CRS) has two parts CRS_p and CRS_v which are to be used by the prover and the verifier respectively.

$$\mathsf{CRS}_p := \begin{pmatrix} [\mathbf{P}_1]_1^{t\times k} = [\mathbf{M}_0^\top \mathbf{L}_1 + \mathbf{M}_1^\top \mathbf{K}_1 + \mathbf{M}_2^\top \mathbf{K}_2]_1, \\ [\mathbf{P}_2]_1^{t\times k} = [\mathbf{M}_0^\top \mathbf{L}_2 + \mathbf{M}_3^\top \mathbf{K}_2]_1, \qquad [\mathbf{P}_3]_1^{t\times k} = [\mathbf{l}_3^\top + \mathbf{a}\cdot \mathbf{K}_1]_1 \end{pmatrix}$$

$$\mathsf{CRS}_v := \big([\bar{\mathbf{A}}]_2, \quad [\mathbf{K}_1\bar{\mathbf{A}}]_2, [\mathbf{K}_2\bar{\mathbf{A}}]_2, \quad [\mathbf{L}_1\bar{\mathbf{A}}]_2, [\mathbf{L}_2\bar{\mathbf{A}}]_2, \quad [\mathbf{l}_3\cdot\bar{\mathbf{A}}]_T\big)$$

Prover P**:** Given candidate $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \text{TAG})$, with witness vector $\mathbf{x}^{t\times 1}$, the prover generates the following proof consisting of k elements in $\mathbb{G}_1$:

$$\pi := \mathbf{x}\cdot[\mathbf{P}_1 + \text{TAG}.\mathbf{P}_2]_1 + [\mathbf{P}_3]_1$$

Verifier V**:** The algorithm hkgen additionally takes TAG as input. Sample $\mathbf{s} \leftarrow \mathbb{Z}_q^k$. Given CRS_v as above, compute hk and hp as follows:

$$\text{HK} := \begin{pmatrix} [(\mathbf{L}_1 + \text{TAG}\ \mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{K}_1\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{K}_2\bar{\mathbf{A}}\mathbf{s}]_2, \\ [\mathbf{l}_3\cdot\bar{\mathbf{A}}\mathbf{s}]_T \end{pmatrix}, \quad \text{HP} := [\bar{\mathbf{A}}\mathbf{s}]_2$$

The algorithms pubH and privH are as follows: Given candidate $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \text{TAG})$, and proof π, compute:

$$\mathsf{privH}(\text{HK}, (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \text{TAG})) := \begin{pmatrix} e(\mathbf{y}_1^\top, [(\mathbf{L}_1 + \text{TAG}\ \mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}]_2) + \\ e(\mathbf{y}_2^\top, [\mathbf{K}_1\bar{\mathbf{A}}\mathbf{s}]_2) + e(\mathbf{y}_3^\top, [\mathbf{K}_2\bar{\mathbf{A}}\mathbf{s}]_2) \\ +[\mathbf{l}_3\cdot\bar{\mathbf{A}}\mathbf{s}]_T \end{pmatrix}$$

$$\mathsf{pubH}(\text{HP}, \pi) := e(\pi, \text{HP})$$

Theorem 3. *The above algorithms* $(\mathsf{K}_0, \mathsf{K}_1, \mathsf{P}, \mathsf{V})$ *constitute a smooth computationally-sound Quasi-Adaptive NIZK proof system for linear subspace languages* $\{L_{\mathbf{M}_0,\mathbf{M}_1,\mathbf{M}_2,\mathbf{M}_3,\mathbf{a}}\}$ *with parameters* $(\mathbf{M}_0, \mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{a})$ *sampled from a robust and efficiently witness-samplable distribution* $\mathcal{D}$*, given any group generation algorithm for which the* $\mathcal{D}_k$*-MDDH assumption holds for group* $\mathbb{G}_2$*.*

The proof is in Appendix C. In the next section we will also take advantage of the special structure of the above smooth QA-NIZK. In particular, we can split the private hash key HK as $(\text{HK}_1, \text{HK}_2)$:

$$\text{HK}_1 := \left([(\mathbf{L}_1 + \text{TAG } \mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{K}_1\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{K}_2\bar{\mathbf{A}}\mathbf{s}]_2\right), \quad \text{HK}_2 := [\mathbf{I}_3 \cdot \bar{\mathbf{A}}\mathbf{s}]_T$$

We can also define an additional function privH' as follows: Given candidate $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \text{TAG})$, compute:

$$\mathsf{privH}'(\text{HK}_1, (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \text{TAG})) := \begin{pmatrix} e(\mathbf{y}_1^\top, [(\mathbf{L}_1 + \text{TAG } \mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}]_2) + \\ e(\mathbf{y}_2^\top, [\mathbf{K}_1\bar{\mathbf{A}}\mathbf{s}]_2) + e(\mathbf{y}_3^\top, [\mathbf{K}_2\bar{\mathbf{A}}\mathbf{s}]_2) \end{pmatrix}$$

Finally, we note:

$$\mathsf{privH}(\text{HK}, (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \text{TAG})) := \mathsf{privH}'(\text{HK}_1, (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \text{TAG})) + \text{HK}_2$$

6 Smooth Signatures

Definition 3. *A Smooth Signature scheme is a tuple of efficient algorithms (KeyGen, Sign, HkGen, VerHash), where*

- *KeyGen*(1^m) *generates public verification key* vk*, and a secret signing key* sk*,*
- *Sign*(sk, m)*, generates a signature on message* m*,*
- *HkGen*(vk, m)*, takes the public verification key* vk *and a message* m*, to produce a hash* h *and a projection hash key* hp *(the set of values* h *is exponential in security parameter* m*),*
- *VerHash*(hp, sig)*, takes the projection hash key and a signature to produce a hash.*

The smooth signature scheme is required to have the following two properties ($\mathcal{U}$ *is the uniform distribution on possible hash values* h *output by VkGen): :*

Correctness:

$$\Pr\left[\begin{array}{l}(vk, sk) \leftarrow KeyGen() \\ sig \leftarrow Sign(sk, m) \\ (hp, h) \leftarrow HkGen(vk, m)\end{array} : VerHash(hp, sig) = h\right] = 1$$

Smoothness:

$$\Pr[(vk, sk) \leftarrow KeyGen() : \mathcal{A}^{Sign(.), HkGen(vk,.)}(vk) = 1]$$

$$\approx$$

$$\Pr[(vk, sk) \leftarrow KeyGen() : \mathcal{A}^{Sign(.), HkGen^*(vk,.)}(vk) = 1],$$

$HkGen^*(vk, .)$ calls $HkGen(vk, .)$ to get (hp, h), samples $u \leftarrow \mathcal{U}$ and returns (hp, u). The condition on the oracles is that the HkGen oracle is called only once with a query distinct from all the Sign calls.

It is straightforward to see that a smooth signature scheme is EUF-CMA (existential unforgeability under chosen message attack [GMR88]) secure where the verify algorithm calls $HkGen(vk, m)$ and tests the correctness condition. Essentially, the smooth signature adversary $\mathcal{A}$ would use the signing oracle to respond to the EUF-CMA adversary $\mathcal{B}$. Once $\mathcal{B}$ responds with a signature (m^*, sig^*), $\mathcal{A}$ would query $HkGen(vk, m^*)$ and get (hp, h). If the correctness condition holds then it would output 1 else 0. In the real world, the probability of outputting 1 is the EUF-CMA advantage of $\mathcal{B}$, while in the $HkGen^*$ world, it is negligible as u is chosen uniformly randomly from an exponentially large domain.

We now show that a smooth signature scheme Σ can be used in a black-box way to construct a fully-adaptive CPA-secure IBE scheme (for definitions, see [BF01] (the Naor conversion of IBE to signatures continues to hold from IBE-KEM (key-encapsulation mode [CS98]) to smooth signatures – see e.g. [BF01]):

Setup: Sample $(msk, pk) \leftarrow \Sigma.KeyGen()$ and output that.
Encrypt(pk, i, M)**:** Sample $(hp, h) \leftarrow \Sigma.HkGen(pk, i)$. Output $C := (hp, h + M)$.
KeyGen(msk, i)**:** Sample $K_i \leftarrow \Sigma.Sign(i)$ and output K_i.
Decrypt$(K_i, C = (C_1, C_2))$**:** Output $\Sigma.VerHash(C_1, K_i) - C_2$

Correctness of decryption follows from the correctness of the smooth signature. Adaptive CPA-security follows by constructing a smooth signature attacker $\mathcal{B}$ from a CPA attacker $\mathcal{A}$. The keygen calls of A are simulated by Sign calls by $\mathcal{B}$. The challenge ciphertext is simulated by $\mathcal{B}$ by calling $HkGen$ with the given query and then adding the query message to the second component. Finally, $\mathcal{B}$ outputs whatever $\mathcal{A}$ outputs.

Now note that in the fake $HkGen^*$ world, h is uniformly random, effectively blinding the message M. Also the IBE oracle restrictions respect the smooth signature oracle restrictions, which is not querying for a ciphertext for an id for which a key was obtained. Smoothness of the signature tells us that these two worlds are indistinguishable which translates exactly to CPA security.

The Naor conversion of IBE to signatures [BF01] now continues to hold from IBE-KEM (key-encapsulation mode [CS98]) to smooth signatures. The private key of the signer in the signature scheme is the master secret key of the IBE(-KEM) scheme. The public verification key is the global parameters of the IBE scheme. Signatures on a message m are decryption keys for identity "m". The $HkGen$ function of the smooth signature scheme on input the global parameters and a message m^* produces $(hp, h) := (c^*, k^*)$ by generating a fresh encapsulation key k^* and the corresponding ciphertext $c*$ for identity m^*. Finally, the $VerHash$ function of the smooth signature scheme on input a projection key hp and signature sig^*, outputs the encapsulation key obtained by running IBE-decrypt on decryption key sig^* and ciphertext hp. Smoothness of the signature scheme is assured by noting that fully-adaptive CPA-secure IBE-KEM assures that k^* (i.e. h) is random to an Adversary even when it is given hp for an identity m^* for which decryption keys (or signatures) were not previously obtained.

6.1 Construction

We now construct a smooth signature scheme which has smoothness under $\mathcal{D}_k$-MDDH assumptions in groups $\mathbb{G}_1$ and $\mathbb{G}_2$ of a bilinear group tuple. This construction is similar to dual-system IBE schemes obtained in both [CLL⁺13, JR13], but is nicely abstracted here in terms of SPHF and smooth QA-NIZK.

KeyGen: Sample $\mathbf{B} \leftarrow \mathcal{D}_k$ and $u \leftarrow \mathbb{Z}_q$. Let L be language $\{(R, S) \mid \exists \mathbf{r} : R = [\overline{\mathbf{B}}\mathbf{r}]_1, S = [\underline{\mathbf{B}}\mathbf{r}]_1\}$. Sample $(\mathsf{hp}, \mathsf{hk}) \leftarrow \mathsf{sphf}(L).\mathsf{hkgen}$. Define the language:

$$L^+ = \left\{ (R, S, T) \mid \exists \mathbf{r} : \begin{array}{c} R = [\overline{\mathbf{B}}\mathbf{r}]_1,\ S = [\underline{\mathbf{B}}\mathbf{r} + u]_1, \\ T = \mathsf{sphf}(L).\mathsf{pubH}(\mathsf{hp},\ (R, S - [u]_1),\ m; \mathbf{r}) \end{array} \right\}$$

Let Π be a split-CRS Smooth QA-NIZK, as constructed in Section 5, for L^+. Let $(\mathsf{CRS}_p, \mathsf{CRS}_v) \leftarrow \Pi.\mathsf{crsgen}()$. Set:

$$vk := \mathsf{CRS}_v, \quad sk := ([\mathbf{B}]_1,\ [u]_1,\ \mathsf{hp},\ \mathsf{CRS}_p).$$

Sign(sk, m): Sample $\mathbf{r} \leftarrow \mathbb{Z}_q^k$ and output:

$$sig := \begin{pmatrix} R = [\overline{\mathbf{B}}\mathbf{r}]_1,\ S = [\underline{\mathbf{B}}\mathbf{r} + u]_1, \\ T = \mathsf{sphf}(L).\mathsf{pubH}(\mathsf{hp},\ (R, S - [u]_1),\ m; \mathbf{r}) \\ \pi = \Pi.\mathsf{prover}(\mathsf{CRS}_p,\ (R, S, T),\ \mathbf{r}) \end{pmatrix}$$

HkGen(vk, m): Sample $(\mathrm{HP}, \mathrm{HK} = (\mathrm{HK}_1, \mathrm{HK}_2)) \leftarrow \Pi.HkGen(\mathsf{CRS}_v)$ and output:

$$hp := (\mathrm{HP}, \mathrm{HK}_1),\ h := \mathrm{HK}_2$$

VerHash(hp, sig): Output:

$$\Pi.\mathsf{privH}'(hp_2, (sig_1, sig_2, sig_3)) - \Pi.\mathsf{pubH}(hp_1, sig_4).$$

6.2 Robust Threshold Smooth Signatures

We first define Robust Threshold Smooth Signatures.

Definition 4 (Robust Threshold Smooth Signatures). *A threshold smooth signature scheme is a tuple (Setup, Sign-Share, Sign-Verify, Sign-Combine, HkGen, VerHash) of efficient algorithms described as follows:*

Setup: *Generate vk and shares* $\{(vk_i, sk_i)\}_i$.
Sign-Share: *On input* (vk, j, sk_j, m), *output signature share* sig_j.
Share-Verify: *On input* $(vk, j, vk_j, \tilde{sig}, m)$, *it checks if the signature share* $\tilde{sig}$ *is valid for message* m *and share index* j.

Sign-Combine: *On input vk, the complete vector of vk_i, a message m, and a set of signature shares $\{\tilde{sig_i}\}_{i \in S}$, s.t. $S \subseteq [1..n]$ is of size t, it outputs failure if any of the signature shares fails the Share-Verify test, and otherwise outputs a signature sig for message m.*
HkGen: *On input vk and message m, outputs a pair (hp, h).*
VerHash: *On input hp and signature sig, output a hash vh.*

Definition 5. *A Robust (t, n)-Threshold Smooth Signature scheme has the following four properties:*

Share Correctness: *For all j:*

$$\Pr\left[\begin{array}{l}(vk, \{(vk_i, sk_i)\}_i) \leftarrow KeyGen() \\ sig_j \leftarrow \textit{Sign-Share}(vk, j, sk_j, m)\end{array} : \textit{Share-Verify}(vk, j, vk_j, sig, m) = 1\right] = 1$$

Correctness: *For all m and $S \subseteq [1..n]$, with $|S| = t$:*

$$\Pr\left[\begin{array}{l}(vk, \{(vk_i, sk_i)\}_i) \leftarrow KeyGen() \\ \forall i \in S: \ sig_i \leftarrow \textit{Sign-Share}(vk, i, sk_i, m) \\ sig \leftarrow \textit{Sign-Combine}(vk, \{vk_i\}_i, m, \{sig_i\}_{i \in S}) \\ (hp, h) \leftarrow HkGen(vk, m)\end{array} : VerHash(hp, sig) = h\right] = 1$$

Smoothness: *For all PPT $\mathcal{A}$:*

$$\Pr\left[(vk, \{(vk_i, sk_i)\}_i) \leftarrow KeyGen() : \mathcal{A}^{\textit{Sign-Share}(.,.), HkGen(.), Corrupt(.)}(vk) = 1\right]$$

$$\approx$$

$$\Pr\left[(vk, \{(vk_i, sk_i)\}_i) \leftarrow KeyGen() : \mathcal{A}^{\textit{Sign-Share}(.,.), HkGen^*(.), Corrupt(.)}(vk) = 1\right],$$

$Corrupt(i)$ returns sk_i. $HkGen^*(.)$ calls $HkGen(.)$ to get (hp, h), samples $u \leftarrow \mathcal{U}$ and returns (hp, u). The adversary $\mathcal{A}$ is restricted to calling oracle $HkGen/HkGen^*$ only once and with a query distinct from all the Sign-Share calls; additionally, at most $t - 1$ corruption requests can be made.
Robustness: *For all PPT $\mathcal{A}$:*

$$\Pr\left[\begin{array}{c}(vk, \{(vk_i, sk_i)\}_i) \leftarrow KeyGen() \\ S, m^*, \{sig_i^*\}_{i \in S} \leftarrow \mathcal{A}^{Corrupt(.)}(vk) \\ (hp, h) \leftarrow HkGen(vk, m^*)\end{array} : \begin{array}{c}|S| = t \textbf{ and } sig^* \neq \perp \textbf{ and} \\ \textit{Sign-Combine}\begin{pmatrix}vk, \{vk_i\}_i, \\ m^*, \{sig_i^*\}_{i \in S}\end{pmatrix} \\ = sig^* \\ \textbf{and } VerHash(hp, sig^*) \neq h\end{array}\right] \approx negl.$$

6.3 Robust Threshold Smooth Signature Construction

We now construct a robust (t, n)-threshold smooth signature scheme which has smoothness under given $\mathcal{D}_k$-MDDH assumptions in $\mathbb{G}_1$ and $\mathbb{G}_2$ of a bilinear group tuple.

KeyGen: Sample $\mathbf{B} \leftarrow \mathcal{D}_k$ and $(\mathbf{d},\ \mathbf{e}) \leftarrow \mathbb{Z}_q^k \times \mathbb{Z}_q^k$. Also sample $\mathbf{A} \leftarrow \mathcal{D}_k$ and $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{L}_1, \mathbf{L}_2) \leftarrow \mathbb{Z}_q^k \times \mathbb{Z}_q^k \times \mathbb{Z}_q^{k\times k} \times \mathbb{Z}_q^{k\times k}$. Sample $(p_1, \mathbf{p}_2) \in \mathbb{Z}_q[X] \times \mathbb{Z}_q[X]^k$ as (vector of) degree $(t-1)$ polynomials with coefficients uniformly randomly sampled from $\mathbb{Z}_q$. Define the tagged affine languages L_i for each i as follows:

$$L_i = \left\{ (\mathbf{y}, m) \mid \exists \mathbf{r} \in \mathbb{Z}_q^k : \ \mathbf{y} = \begin{pmatrix} [\overline{\mathbf{B}}\mathbf{r}]_1,\ [\underline{\mathbf{B}}\mathbf{r} + p_1(i)]_1,\ [(\mathbf{d}\cdot\overline{\mathbf{B}} + m\mathbf{e}\cdot\overline{\mathbf{B}})\mathbf{r}]_1, \\ \mathbf{r}\cdot[\overline{\mathbf{B}}^\top(\mathbf{L}_1 + m\mathbf{L}_2) + \underline{\mathbf{B}}^\top\mathbf{k}_1^\top + (\mathbf{d}\cdot\overline{\mathbf{B}} + m\mathbf{e}\cdot\overline{\mathbf{B}})^\top\mathbf{k}_2^\top]_1 \\ +[p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1 \end{pmatrix} \right\}$$

Note that L_i is extension of the language L^+ of Section 6.1, where the fourth component of $\mathbf{y}$ is the smooth QA-NIZK proof of the affine language L^+ with u replaced by $p_1(i)$, and the $[\mathbf{l}_3]_1$ (see Section 5.2) replaced by $[\mathbf{p}_2(i)]_1$. Since QA-NIZK proofs are linear in $\mathbf{r}$, L_i is a tagged affine language, linear in $\mathbf{r}$ with affine components $[p_1(i)]_1$ and $[p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1$ (and tags m).
For each $i \in [1..n]$, let Π_i be a split-CRS QA-NIZK for L_i. Finally, set:

$$vk := \begin{pmatrix} [\mathbf{B}]_1,\ [\mathbf{d}]_1,\ [\mathbf{e}]_1, \\ [\overline{\mathbf{B}}^\top\mathbf{L}_1 + \underline{\mathbf{B}}^\top\mathbf{k}_1^\top + (\mathbf{d}\cdot\overline{\mathbf{B}})^\top\mathbf{k}_2^\top]_1,\ [\overline{\mathbf{B}}^\top\mathbf{L}_2 + (\mathbf{e}\cdot\overline{\mathbf{B}})^\top\mathbf{k}_2^\top]_1, \\ [\overline{\mathbf{A}}]_2,\ [\mathbf{L}_1\overline{\mathbf{A}}]_2,\ [\mathbf{L}_2\overline{\mathbf{A}}]_2,\ [\mathbf{k}_1\cdot\overline{\mathbf{A}}]_2,\ [\mathbf{k}_2\cdot\overline{\mathbf{A}}]_2, [\mathbf{p}_2(0)\cdot\overline{\mathbf{A}}]_T \end{pmatrix}$$

$$\forall i \in [1..n] : \ vk_i := \Pi_i.\mathsf{CRS}_v, \quad sk_i := \left([p_1(i)]_1,\ [p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1,\ \Pi_i.\mathsf{CRS}_p\right).$$

Recall, that the split-CRS QA-NIZK guarantees that CRS_v is independent of the affine components. Moreover, in the CRS_p (see Section 5.2) only $[\mathbf{P}_3]_1 = [p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1$ depends on affine components, and rest of the CRS_p can be made public. Indeed $[\mathbf{P}_1]_1$ and $[\mathbf{P}_2]_1$ are included in vk above. Similarly, when we give a distributed master-key generation protocol in the next sub-section, we will note that $\Pi_i.\mathsf{CRS}_p$ can also be similarly split, and most of it can be made public and only the linear combination of the affine parts kept secret in sk_i.

Sign-Share(sk_i, i, m): Sample $\mathbf{r} \leftarrow \mathbb{Z}_q^k$ and output $sig_i := (sig_{i,0}, \pi)$, where:

$$sig_{i,0} := \begin{pmatrix} [\overline{\mathbf{B}}\mathbf{r}]_1,\ [\underline{\mathbf{B}}\mathbf{r} + p_1(i)]_1,\ [(\mathbf{d}\cdot\overline{\mathbf{B}} + m\mathbf{e}\cdot\overline{\mathbf{B}})\mathbf{r}]_1, \\ \mathbf{r}\cdot[\overline{\mathbf{B}}^\top(\mathbf{L}_1 + m\mathbf{L}_2) + \underline{\mathbf{B}}^\top\mathbf{k}_1^\top + (\mathbf{d}\cdot\overline{\mathbf{B}} + m\mathbf{e}\cdot\overline{\mathbf{B}})^\top\mathbf{k}_2^\top]_1 \\ +[p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1 \end{pmatrix}$$

$$\pi := \Pi_i.\mathsf{prover}(\Pi_i.\mathsf{CRS}_p,\ (sig_{i,0}, m),\ \mathbf{r})$$

Sign-Verify(vk_i, i, sig_i, m): Output the boolean:

$$\Pi_i.\mathsf{ver}(\Pi_i.\mathsf{CRS}_v,\ (sig_{i,0}, m),\ sig_{i,1})$$

Sign-Combine($vk, \{vk_i\}_i, m, \{sig_i\}_{i\in S}$): Check Sign-Verify for each of the $(|S| = t)$ shares and return $\perp$ if any one is false. Otherwise, let ω_i be the Lagrange interpolation co-efficients, which depend only on S, such that for any degree

$(t-1)$ polynomial f, we have $f(0) = \sum_{i \in S} \omega_i f(i)$. Then compute component-wise:

$$sig := \sum_{i \in S} \omega_i sig_{i,0}$$

HkGen(vk, m): Sample $\mathbf{s} \leftarrow \mathbb{Z}_q^k$ and output:

$$hp := \big([(\mathbf{L}_1 + m\mathbf{L}_2)\overline{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_1 \cdot \overline{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_2 \cdot \overline{\mathbf{A}}\mathbf{s}]_2,\ [\overline{\mathbf{A}}\mathbf{s}]_2\big)$$

$$h := [\mathbf{p}_2(0) \cdot \overline{\mathbf{A}}\mathbf{s}]_T$$

VerHash(hp, sig): Output:

$$\sum_{i=1}^{3} \mathsf{e}(sig_i, hp_i) - \mathsf{e}(sig_4, hp_4)$$

Theorem 4. *Under given $\mathcal{D}_k$-MDDH assumptions in groups $\mathbb{G}_1$ and $\mathbb{G}_2$, the above robust (t,n)-threshold smooth signature scheme satisfies share-correctness, correctness, smoothness and the robustness properties.*

While share-correctness and correctness of the threshold smooth signature scheme follows easily from completeness of the underlying SPHF and smooth QA-NIZK schemes, the robustness property follows from the soundness of each of the QA-NIZK Π_i. The proof of the smoothness property is more involved and requires dual-system techniques. Fortunately, the proof of smoothness of the threshold smooth signature scheme is much easier once we understand the proof of smoothness of the basic smooth signature scheme. In turn, the proof of the smooth signature scheme is streamlined because of the underlying smooth split-CRS QA-NIZK for affine languages. The proof of the smooth signature scheme is given in detail in Section 6.7, and the proof of the threshold scheme is deferred to Section 6.8.

6.4 Distributed Master Secret-Key Generation

While the threshold smooth signature scheme described above can easily be implemented using a trusted dealer, we now show that the techniques of [GJKR07] can be generalized and used to generate the master secret key shares in a secure distributed fashion. For discrete log based cryptosystems, [GJKR07] allows for secret sharing of a uniformly random x, and making public the value g^x. The technique can easily be extended to generating various uniformly random and independent values in bilinear groups, say $[x_1]_1$, $[x_2]_1$ ... etc. in group $\mathbb{G}_1$, and $[y_1]_2$, $[y_2]_2$ ... etc. in group $\mathbb{G}_2$, and then to inductively generate random linear combinations of these elements in the two groups using the same random coefficients. At the same time, for some (or all) of these coefficients, the Shamir secret-shares of the coefficients are obtained by the individual parties. The [GJKR07] protocol can also be generalized to generate $g^x h^y$ for given public g, h (or generated randomly and distributedly in a previous round), where the Shamir secret-shares of both x and y are obtained by the individual parties.

Since the $\mathcal{D}_k$-MDDH assumption implies the $\mathcal{U}_k$-MDDH assumption, in the above scheme the matrices $\mathbf{B}$ and $\mathbf{A}$ can be generated uniformly at random. Then, the above generalization of [GJKR07] can be used to generate the public keys vk and $\{vk_i\}_i$ in a distributed fashion, with individual parties retaining the secret keys sk_i. Note, as mentioned in the description of the threshold smooth signature scheme, even $\Pi_i.\mathsf{CRS}_p$ can be split so that only the components dependent on the affine values, i.e., $[p_1(i)]_1,\ [p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1$ need to be part of sk_i and the rest of the components can be public in vk. Thus, vk (excluding $[\mathbf{p}_2(0)\cdot\overline{\mathbf{A}}]_T$) and vk_i (for each $i \in [1..n]$) can be generated distributedly first, and then the shares of p_1 and $\mathbf{p}_2$ can be generated along with public $[\mathbf{p}_2(0)\cdot\overline{\mathbf{A}}]_T$, using the above generalization to $g^x h^y$. Detailed proof will be given in the full version of the paper.

6.5 Threshold IBE

We first define Threshold IBE [BBH06].

Definition 6 ((t,n)-Threshold IBE). *A threshold IBE is a tuple of algorithm (Setup, KeyGen-Share, KeyGen-Verify, KeyGen-Combine, Encrypt, Decrypt):*

Setup: *Generate* pk *and* $\mathbf{MSK}$ *shares (*$\mathbf{MSK}_i$*), as well as master verification keys* $\mathbf{MVK}_i$ *(one for each share* $i \in [1..n]$*).*
KeyGen-Share: *On input* pk, $\mathbf{MSK}_j$*, an identity* id *and* $j \in [1..n]$*, output keyshare* $K_{\mathsf{id},j}$*.*
KeyGen-Verify: *On input* pk*,* i*,* $\mathbf{MVK}_j$*,* $\tilde{K}$*,* id*, it checks if the key-share* $\tilde{K}$ *is valid for identity* id *and share index* i*.*
KeyGen-Combine: *On input* pk*, the complete vector of* $\mathbf{MVK}_i$*, an identity* id*, and a set of keyshares* $\{\tilde{K}_i\}_{i\in S}$*, s.t.* $S \subseteq [1..n]$ *is of size* t*, it outputs failure if any of the key-shares fails the KeyGen-Verify test, and otherwise outputs a key* K *for identity* id*.*
Encrypt: *On input* pk*, and identity* id*, and a message* M*, outputs a ciphertext* C*.*
Decrypt: *On input a ciphertext* C*, and an identity* id*, outputs a message* M*.*

Definition of correctness is standard, i.e., decryption using key for identity id is correct on encryptions made for identity id.

Definition 7. *A* (t,n)*- threshold IBE scheme is (chosen-plaintext-adversary) CPA fully-secure if no PPT adversary has non-negligible advantage in the following game:*

1. *The Challenger* $\mathcal{C}$ *runs Setup to obtain* pk*, vector of master secret key shares* $\mathbf{MSK}_i$ *and verification keys* $\mathbf{MVK}_i$ *(*$i \in [1..n]$*). It gives* pk *and all the verification keys to the Adversary* $\mathcal{A}$ *and retains the master secret key shares.*
2. *The adversary* $\mathcal{A}$ *adaptively makes the following kind of queries:*
   - *Corruption Query.* $\mathcal{A}$ *asks for* $\mathbf{MSK}_j$ *for some* $j \in [1..n]$*.*
   - *Key-Share Request.* $\mathcal{A}$ *asks for share* j *of identity* id*, and challenger returns KeyGen-Share*(id, j)*.*

3. *$\mathcal{A}$ chooses equal length messages M_0, M_1, and an identity id^* different from all previously queried id, and obtains a ciphertext $C^* = Encrypt(\mathsf{pk}, \mathsf{id}^*, M_\gamma))$, for a random bit γ.*
4. *$\mathcal{A}$ makes further queries as in step 2, except key-share requests for identity id^* are not allowed.*
5. *$\mathcal{A}$ outputs a bit γ'.*

Adversary's advantage is the probability of outputting γ' equal to γ (over and above probability $1/2$), and when the total number of corruption queries in step 2 and 4 is at most $t-1$.

Definition 8. *A (t, n)-threshold IBE scheme is said to be consistent if no PPT adversary has non-negligible advantage in the following game: steps 1 and 2 are as in the game above, except there is no limit on how many (share) corruptions it can make. In step 3, $\mathcal{A}$ outputs a ciphertext C meant for an identity id, along with two sets of t key-shares for identity id. $\mathcal{A}$ is considered successful if (1) KeyGen-Verify holds for all the $2 * t$ key-shares, and (2) Decryption of C using KeyGen-Combine using the first set of key shares is different from decryption using KeyGen-Combine using the second set of key shares.*

6.6 The Threshold IBE Scheme

We now give a threshold IBE construction based on a threshold smooth signature scheme Θ.

Setup: Sample $(\mathsf{pk}, \{\mathbf{MVK}_i, \mathbf{MSK}_i\}_i) \leftarrow \Theta.KeyGen()$ and output that.
KeyGen-Share: On input $(\mathsf{pk}, \mathbf{MSK}_j, \mathsf{id}, j)$, sample $K_{\mathsf{id},j} \leftarrow \Theta.Sign{-}Share(\mathsf{pk}, j, \mathbf{MSK}_j, \mathsf{id})$ and output that.
KeyGen-Verify: On input $(\mathsf{pk}, j, \mathbf{MVK}_j, \tilde{K}, \mathsf{id})$, output $Sign{-}Verify(\mathsf{pk}, j, \mathbf{MVK}_j, \tilde{K}, \mathsf{id})$
KeyGen-Combine: On input $(\mathsf{pk}, \{\mathbf{MVK}_i\}_i, \mathsf{id}, \{\tilde{K}\}_{i\in S})$, output $K_{\mathsf{id}} := \Theta.Sign{-}Combine(\mathsf{pk}, \{\mathbf{MVK}_i\}_i, \mathsf{id}, \{\tilde{K}\}_{i\in S})$.
Encrypt: On input $(\mathsf{pk}, \mathsf{id}, M)$, sample $(vk_1, vk_2) \leftarrow \Theta.VkGen(\mathsf{pk}, \mathsf{id})$ and output $C := (vk_1, vk_2 + M)$.
Decrypt: On input $(K_{\mathsf{id}}, C = (C_1, C_2))$, output $M := \Theta.VerHash(C_1, K_{\mathsf{id}}) - C_2$.

Theorem 5. *Under $\mathcal{D}_k$-MDDH assumptions in groups $\mathbb{G}_1$ and $\mathbb{G}_2$, the above threshold IBE scheme is a (t, n)-threshold CPA fully-secure IBE scheme.*

Just as shown in Section 6 it is straightforward to show that (t, n)-threshold smooth signatures imply (t, n)-threshold CPA-fully secure IBE. The IBE scheme so obtained may not satisfy the additional consistency condition. However, the obtained threshold IBE scheme is easily made consistent by attaching a QA-NIZK proof to each ciphertext, proving in zero-knowledge that the ciphertexts are of the correct form. Note that the ciphertexts are nothing but *hp* generated by *HkGen* of the Threshold Smooth Signature Scheme. These *hp* belong to a simple linear in $\mathbf{s}$ tag-based language with parameters $\mathbf{L}_1, \mathbf{L}_2, \mathbf{k}_1, \mathbf{k}_2, \mathbf{A}$. These parameters were generated by KeyGen of the smooth signature scheme (and hence the SetUp of the threshold IBE scheme). Thus, it can also generate the QA-NIZK CRS for this linear tag-based language, and be made part of the the global parameters. The prover CRS would be used in the generation of the QA-NIZK for the ciphertext (which would be just a single group element), and the verifier CRS would be used by the decryption algorithm. The decryption algorithm will try to decrypt only if the QA-NIZK verification succeeds. The soundness of the QA-NIZK assures that the ciphertext is indeed of the correct form. Indeed, if the ciphertext c is of the form

$$c := \big([(\mathbf{L}_1 + m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_1 \cdot \bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_2 \cdot \bar{\mathbf{A}}\mathbf{s}]_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2\big)$$

for some $\mathbf{s}$ and label m, then it is easy to check that VerHash(c, sig) will return $[\mathbf{p}_2(0) \cdot \bar{\mathbf{A}}\mathbf{s}]_T$ regardless of any sig that is returned by Sign-Combine for m.

6.7 Proof of Smoothness of Smooth Signature

Proof. We proceed through a sequence of games where the first game would be in $HkGen$ world, and the last would be in $HkGen^*$ world. The protocol is fully fleshed out in Figure 1 and the essential elements of the game sequence are summarized in Figure 2.

Game $\mathbf{G}_0$: This is the real world.

Game $\mathbf{G}_1$: In this world, hp and h are generated differently when the $HkGen$ oracle is queried with a message m. Sample $(\mathbf{s}, s'_1, s'_2) \leftarrow \mathbb{Z}_q^k \times \mathbb{Z}_q \times \mathbb{Z}_q$, define $\mathbf{t} = (\underline{\mathbf{B}}\bar{\mathbf{B}}^{-1})^\top$, and compute:

$$hp := \left(\begin{bmatrix} (\mathbf{L}_1 + m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s} + \\ \mathbf{t}s'_1 + (\mathbf{d} + m\mathbf{e})s'_2 \end{bmatrix}_2,\ [\mathbf{k}_1 \cdot \bar{\mathbf{A}}\mathbf{s} - s'_1]_2,\ [\mathbf{k}_2 \cdot \bar{\mathbf{A}}\mathbf{s} - s'_2]_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2 \right)$$

$$h := [\mathbf{l}_3 \cdot \bar{\mathbf{A}}\mathbf{s} + us'_1]_T$$

Games $\mathbf{G}_0$ and $\mathbf{G}_1$ are computationally indistinguishable under the $\mathcal{D}_k$-MDDH assumption. This game is same as the Game 2 of the smooth QA-NIZK smoothness proof, where the hkgen algorithm is modified to expose the kernel of the language.

Game $\mathbf{G}_2$: In this world, $\mathbf{d}, \mathbf{e}$ and s'_1 are sampled differently, but preserving their uniform distribution. First sample $(\mathbf{d}_1, d_2, \mathbf{e}_1, e_2) \leftarrow \mathbb{Z}_q^k \times \mathbb{Z}_q \times \mathbb{Z}_q^k \times \mathbb{Z}_q$ and set $\mathbf{d} = \mathbf{d}_1 + d_2\mathbf{t}$ and $\mathbf{e} = \mathbf{e}_1 + e_2\mathbf{t}$. When hkgen is called with message m, then set $s'_1 = -(d_2 + me_2)s'_2$, responding with the following modified hp:

$$hp := \left(\begin{bmatrix} (\mathbf{L}_1 + m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s} \\ +(\mathbf{d}_1 + m\mathbf{e}_1)s'_2 \end{bmatrix}_2,\ [\mathbf{k}_1 \cdot \bar{\mathbf{A}}\mathbf{s} - s'_1]_2,\ [\mathbf{k}_2 \cdot \bar{\mathbf{A}}\mathbf{s} - s'_2]_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2 \right)$$

Note that $\mathbf{t}$ has disappeared from $\mathbb{G}_2$ elements. This step is essentially switching from public SPHF evaluation to private SPHF evaluation with an additional step to absorb the $\mathbf{t}$ vector in group $\mathbb{G}_2$.

KeyGen: Sample $\mathbf{B} \leftarrow \mathcal{D}_k$ and $(u, \mathbf{d}, \mathbf{e}) \leftarrow \mathbb{Z}_q \times \mathbb{Z}_q^k \times \mathbb{Z}_q^k$. Also sample $\mathbf{A} \leftarrow \mathcal{D}_k$ and $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{L}_1, \mathbf{L}_2, \mathbf{l}_3) \leftarrow \mathbb{Z}_q^k \times \mathbb{Z}_q^k \times \mathbb{Z}_q^{k\times k} \times \mathbb{Z}_q^{k\times k} \times \mathbb{Z}_q^k$. Set:

$$vk := ([\bar{\mathbf{A}}]_2,\ [\mathbf{L}_1\bar{\mathbf{A}}]_2,\ [\mathbf{L}_2\bar{\mathbf{A}}]_2,\ [\mathbf{k}_1 \cdot \bar{\mathbf{A}}]_2,\ [\mathbf{k}_2 \cdot \bar{\mathbf{A}}]_2,\ [\mathbf{l}_3 \cdot \bar{\mathbf{A}}]_2)$$

$$sk := (\mathbf{B},\ u,\ \mathbf{d},\ \mathbf{e},\ \mathbf{L}_1,\ \mathbf{L}_2,\ \mathbf{k}_1,\ \mathbf{k}_2,\ \mathbf{l}_3).$$

Sign(m): Sample $\mathbf{r} \leftarrow \mathbb{Z}_q^k$ and output:

$$sig := \begin{pmatrix} [\bar{\mathbf{B}}\mathbf{r}]_1,\ [\underline{\mathbf{B}}\mathbf{r} + u]_1,\ [(\mathbf{d} \cdot \bar{\mathbf{B}} + m\mathbf{e} \cdot \bar{\mathbf{B}})\mathbf{r}]_1, \\ \mathbf{r} \cdot [\bar{\mathbf{B}}^\top(\mathbf{L}_1 + m\mathbf{L}_2) + \underline{\mathbf{B}}^\top \mathbf{k}_1^\top + (\mathbf{d} \cdot \bar{\mathbf{B}} + m\mathbf{e} \cdot \bar{\mathbf{B}})^\top \mathbf{k}_2^\top]_1 + [u\mathbf{k}_1^\top + \mathbf{k}_5^\top]_1 \end{pmatrix}$$

HkGen(m): Sample $\mathbf{s} \leftarrow \mathbb{Z}_q^k$ and output:

$$hp := \big([(\mathbf{L}_1 + m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_1 \cdot \bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_2 \cdot \bar{\mathbf{A}}\mathbf{s}]_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2\big)$$

$$h := [\mathbf{l}_3 \cdot \bar{\mathbf{A}}\mathbf{s}]_T$$

VerHash(hp**, sig):** Output:

$$\sum_{i=1}^{3} \mathsf{e}(sig_i, hp_i) - \mathsf{e}(sig_4, hp_4)$$

Fig. 1. Smooth Signature Scheme in full detail.

Game $\mathbf{G}_3$: In this game, the signature responses are changed from $\mathbf{G}_2$. For every signature query m, freshly sample $(\mathbf{r}, r', r'') \leftarrow \mathbb{Z}_q^k \times \mathbb{Z}_q \times \mathbb{Z}_q$ and compute the signature as:

$$sig := \begin{pmatrix} [\mathbf{r}]_1,\ [r']_1,\ [r'']_1, \\ [\mathbf{r} \cdot (\mathbf{L}_1 + m\mathbf{L}_2) + r'\mathbf{k}_1^\top + r''\mathbf{k}_2^\top + \mathbf{k}_5^\top]_1 \end{pmatrix}$$

Indistinguishability of Games $\mathbf{G}_2$ and $\mathbf{G}_3$ follow by changing the tuples $([\overline{\mathbf{B}}\mathbf{r}]_1, [\underline{\mathbf{B}}\mathbf{r}]_1)$ to $([\mathbf{r}]_1, [r']_1)$ by the $\mathcal{D}_k$-MDDH assumption one by one, while changing the 2-universal SPHF to $[r'']_1$ in each step. The steps are analogous to the Cramer-Shoup CCA encryption scheme [CS02], where we convert the ElGamal encryptions of u to random. The computation of the QA-NIZK component is basically the simulated proof.

Game $\mathbf{G}_4$: In this game, instead of computing h as in Games $\mathbf{G}_1$-$\mathbf{G}_3$, it's just sampled from $\mathbb{Z}_q$ uniformly. This transition is information theoretic as in Game $\mathbf{G}_3$, u is only used to compute h and nothing else that is visible to the adversary.

Games $\mathbf{G}_5$-$\mathbf{G}_7$: We essentially roll back the transitions made in $\mathbf{G}_0$-$\mathbf{G}_3$ in reverse order, while still sampling h randomly as in Game $\mathbf{G}_4$. The final Game $\mathbf{G}_7$ is exactly as the real world, except that *HkGen* has been replaced by *HkGen**. This establishes our claim.

6.8 Proof of Smoothness of Threshold Smooth Signature

Proof. We again proceed through an almost identical sequence of games where the first game would be in *HkGen* world, and the last would be in *HkGen** world. The essential elements of the sequence are summarized in Figure 3.

As the proof follows the same blueprint, we only discuss the distinctive points here. The additional oracle provided is the Corrupt oracle, which is responded to in the same way in each game. In Games $\mathbf{G}_1$-$\mathbf{G}_3$ the *HkGen* response component h is computed as $[\mathbf{p}_2(0)\cdot\bar{\mathbf{A}}\mathbf{s}+p_1(0)s']_2$. In Game $\mathbf{G}_4$, the element $p_1(0)$ is changed to a fresh random element u'. This is possible as in $\mathbf{G}_3$, all the signatures have been randomized, thus removing the polynomial p_1 from the signatures. The only place p_1 appears is in the corruption responses. However, since at most $t-1$ corruption requests are made and $p_1(0)$ is independent of these $t-1$ evaluations of p_1, it can be sampled independently randomly.

Subsequently, in Game $\mathbf{G}_5$, h is just sampled independently randomly, as u' is not used anywhere else in Game $\mathbf{G}_4$.

7 Asymmetric UC-PAKE: UC-APAKE

Based on the UC-PAKE functionality of [CHK$^+$05], Gentry et al [GMR06] gave another UC functionality for asymmetric PAKE (UC-APAKE). A salient feature of the UC-PAKE functionality [CHK$^+$05] is that it models the security requirement that an adversary cannot perform efficient off-line computations on

$$\text{Games 0-1, 6-7:} \quad sig := \begin{pmatrix} [\bar{\mathbf{B}}\mathbf{r}]_1,\ [\underline{\mathbf{B}}\mathbf{r}+u]_1,\ [(\mathbf{d}\cdot\bar{\mathbf{B}}+m\mathbf{e}\cdot\bar{\mathbf{B}})\mathbf{r}]_1, \\ \mathbf{r}\cdot[\bar{\mathbf{B}}^\top(\mathbf{L}_1+m\mathbf{L}_2)+\underline{\mathbf{B}}^\top\mathbf{k}_1^\top+(\mathbf{d}\cdot\bar{\mathbf{B}}+m\mathbf{e}\cdot\bar{\mathbf{B}})^\top\mathbf{k}_2^\top]_1 \\ +[u\mathbf{k}_1^\top+\mathbf{k}_5^\top]_1 \end{pmatrix} \tag{1}$$

$$\text{Games 2, 5:} \quad sig := \begin{pmatrix} [\bar{\mathbf{B}}\mathbf{r}]_1,\ [\underline{\mathbf{B}}\mathbf{r}+u]_1,\ [(\mathbf{d}_1\cdot\bar{\mathbf{B}}+d_2\underline{\mathbf{B}}+m\mathbf{e}_1\cdot\bar{\mathbf{B}}+me_2\underline{\mathbf{B}})\mathbf{r}]_1, \\ \mathbf{r}\cdot\begin{bmatrix} \bar{\mathbf{B}}^\top(\mathbf{L}_1+m\mathbf{L}_2)+\underline{\mathbf{B}}^\top\mathbf{k}_1^\top+ \\ (\mathbf{d}_1\cdot\bar{\mathbf{B}}+d_2\underline{\mathbf{B}}+m\mathbf{e}_1\cdot\bar{\mathbf{B}}+me_2\underline{\mathbf{B}})^\top\mathbf{k}_2^\top \end{bmatrix}_1 \\ +[u\mathbf{k}_1^\top+\mathbf{k}_5^\top]_1 \end{pmatrix} \tag{2}$$

$$\text{Games (2,i,1), (4,i,2):} \quad sig_i := \begin{pmatrix} [\mathbf{r}]_1,\ [r'+u]_1,\ [\mathbf{d}_1\cdot\mathbf{r}+d_2r'+m_i\mathbf{e}_1\cdot\mathbf{r}+m_ie_2r']_1, \\ \begin{bmatrix} \mathbf{r}\cdot(\mathbf{L}_1+m_i\mathbf{L}_2)+r'\mathbf{k}_1^\top+ \\ (\mathbf{r}\cdot\mathbf{d}_1+d_2r'+m_i\mathbf{r}\cdot\mathbf{e}_1+m_ie_2r')\mathbf{k}_2^\top \\ +u\mathbf{k}_1^\top+\mathbf{k}_5^\top \end{bmatrix}_1 \end{pmatrix} \tag{3}$$

$$\text{Games (2,i,2), (4,i,1):} \quad sig_i := \begin{pmatrix} [\mathbf{r}]_1,\ [r']_1,\ [r'']_1, \\ [\mathbf{r}\cdot(\mathbf{L}_1+m_i\mathbf{L}_2)+r'\mathbf{k}_1^\top+r''\mathbf{k}_2^\top+\mathbf{k}_5^\top]_1 \end{pmatrix} \tag{4}$$

$$\text{Games 0, 7:} \quad hp := \left([(\mathbf{L}_1+m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_1\cdot\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_2\cdot\bar{\mathbf{A}}\mathbf{s}]_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2\right) \tag{5}$$

$$\text{Games 1, 6:} \quad hp := \left(\begin{bmatrix} (\mathbf{L}_1+m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}+ \\ \bar{\mathbf{B}}^{-\top}\underline{\mathbf{B}}^\top s_1'+ \\ (\mathbf{d}+m\mathbf{e})s_2' \end{bmatrix}_2,\ [\mathbf{k}_1\cdot\bar{\mathbf{A}}\mathbf{s}-s_1']_2,\ [\mathbf{k}_2\cdot\bar{\mathbf{A}}\mathbf{s}-s_2']_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2 \right) \tag{6}$$

$$\text{Games 2-5:} \quad hp := \left(\begin{bmatrix} (\mathbf{L}_1+m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s} \\ +(\mathbf{d}_1+m\mathbf{e}_1)s_2' \end{bmatrix}_2,\ [\mathbf{k}_1\cdot\bar{\mathbf{A}}\mathbf{s}-s_1']_2,\ [\mathbf{k}_2\cdot\bar{\mathbf{A}}\mathbf{s}-s_2']_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2 \right) \tag{7}$$

$$\text{Game 0:} \quad h := [\mathbf{l}_3\cdot\bar{\mathbf{A}}\mathbf{s}]_T \tag{8}$$

$$\text{Games 1-3:} \quad h := [\mathbf{l}_3\cdot\bar{\mathbf{A}}\mathbf{s}+us_1']_T \tag{9}$$

$$\text{Game 4-7:} \quad h := [s'']_T \tag{10}$$

Fig. 2. Summary of games for the Smooth Signature proof

$$\text{Games 0-1, 7-8:}\quad sig := \begin{pmatrix} [\bar{\mathbf{B}}\mathbf{r}]_1,\ [\underline{\mathbf{B}}\mathbf{r} + p_1(i)]_1,\ [(\mathbf{d}\cdot\bar{\mathbf{B}} + m\mathbf{e}\cdot\bar{\mathbf{B}})\mathbf{r}]_1, \\ \mathbf{r}\cdot[\bar{\mathbf{B}}^\top(\mathbf{L}_1 + m\mathbf{L}_2) + \underline{\mathbf{B}}^\top\mathbf{k}_1^\top + (\mathbf{d}\cdot\bar{\mathbf{B}} + m\mathbf{e}\cdot\bar{\mathbf{B}})^\top\mathbf{k}_2^\top]_1 \\ +[p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1 \end{pmatrix} \tag{11}$$

$$\text{Games 2, 6:}\quad sig := \begin{pmatrix} [\bar{\mathbf{B}}\mathbf{r}]_1,\ [\underline{\mathbf{B}}\mathbf{r} + p_1(i)]_1,\ [(\mathbf{d}_1\cdot\bar{\mathbf{B}} + d_2\underline{\mathbf{B}} + m\mathbf{e}_1\cdot\bar{\mathbf{B}} + me_2\underline{\mathbf{B}})\mathbf{r}]_1, \\ \mathbf{r}\cdot[\bar{\mathbf{B}}^\top(\mathbf{L}_1 + m\mathbf{L}_2) + \underline{\mathbf{B}}^\top\mathbf{k}_1^\top + \\ (\mathbf{d}_1\cdot\bar{\mathbf{B}} + d_2\underline{\mathbf{B}} + m\mathbf{e}_1\cdot\bar{\mathbf{B}} + me_2\underline{\mathbf{B}})^\top\mathbf{k}_2^\top]_1 \\ +[p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1 \end{pmatrix} \tag{12}$$

$$\text{Games (2,j,1), (5,j,2):}\quad sig_j := \begin{pmatrix} [\mathbf{r}]_1,\ [r' + p_1(i)]_1,\ [\mathbf{d}_1\cdot\mathbf{r} + d_2r' + m_j\mathbf{e}_1\cdot\mathbf{r} + m_je_2r']_1, \\ [\mathbf{r}\cdot(\mathbf{L}_1 + m_j\mathbf{L}_2) + r'\mathbf{k}_1^\top + \\ (\mathbf{r}\cdot\mathbf{d}_1 + d_2r' + m_j\mathbf{r}\cdot\mathbf{e}_1 + m_je_2r')\mathbf{k}_2^\top]_1 \\ +[p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)^\top]_1 \end{pmatrix} \tag{13}$$

$$\text{Games (2,j,2), (5,j,1):}\quad sig_j := \begin{pmatrix} [\mathbf{r}]_1,\ [r']_1,\ [r'']_1, \\ [\mathbf{r}\cdot(\mathbf{L}_1 + m_j\mathbf{L}_2) + r'\mathbf{k}_1^\top + r''\mathbf{k}_2^\top]_1 \\ +[\mathbf{p}_2(i)^\top]_1 \end{pmatrix} \tag{14}$$

$$\text{Games 0, 7:}\quad hp := \big([(\mathbf{L}_1 + m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_1\cdot\bar{\mathbf{A}}\mathbf{s}]_2,\ [\mathbf{k}_2\cdot\bar{\mathbf{A}}\mathbf{s}]_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2\big) \tag{15}$$

$$\text{Games 1, 6:}\quad hp := \left(\begin{bmatrix} (\mathbf{L}_1 + m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s} + \\ \bar{\mathbf{B}}^{-\top}\underline{\mathbf{B}}^\top s_1' + \\ (\mathbf{d} + m\mathbf{e})s_2' \end{bmatrix}_2,\ [\mathbf{k}_1\cdot\bar{\mathbf{A}}\mathbf{s} - s_1']_2,\ [\mathbf{k}_2\cdot\bar{\mathbf{A}}\mathbf{s} - s_2']_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2 \right) \tag{16}$$

$$\text{Games 2-5:}\quad hp := \left(\begin{bmatrix} (\mathbf{L}_1 + m\mathbf{L}_2)\bar{\mathbf{A}}\mathbf{s} \\ +(\mathbf{d}_1 + m\mathbf{e}_1)s_2' \end{bmatrix}_2,\ [\mathbf{k}_1\cdot\bar{\mathbf{A}}\mathbf{s} - s_1']_2,\ [\mathbf{k}_2\cdot\bar{\mathbf{A}}\mathbf{s} - s_2']_2,\ [\bar{\mathbf{A}}\mathbf{s}]_2 \right) \tag{17}$$

$$\text{Game 0:}\quad h := [\mathbf{p}_2(0)\cdot\bar{\mathbf{A}}\mathbf{s}]_T \tag{18}$$

$$\text{Games 1-3:}\quad h := [\mathbf{p}_2(0)\cdot\bar{\mathbf{A}}\mathbf{s} + p_1(0)s_1']_T \tag{19}$$

$$\text{Game 4:}\quad h := [\mathbf{p}_2(0)\cdot\bar{\mathbf{A}}\mathbf{s} + u's_1']_T \tag{20}$$

$$\text{Game 5-8:}\quad h := [s'']_T \tag{21}$$

$$\text{Games 0-8:}\quad sk_i := ([p_1(i)]_1,\ [p_1(i)\mathbf{k}_1^\top + \mathbf{p}_2(i)]_1,\ \Pi_i.\mathsf{CRS}_p) \tag{22}$$

Fig. 3. Summary of games for the Threshold Smooth Signature proof

protocol transcripts to verifiably guess the low-entropy password. An adversary can only benefit from the low-entropy of the password by actually conducting an on-line attack (i.e. by impersonating one of the parties with a guessed password). This is modeled in the ideal world with a TestPwd capability available to the ideal world adversary: if TestPwd is called with the correct password, the ideal world adversary is allowed to set the session key. Moreover, in this functionality if any of the parties is corrupted, then the ideal world adversary is given the registered password.

7.1 The UC Ideal Functionality for Asymmetric PAKE

In asymmetric PAKE [GMR06], the ideal functionality also allows an adversary to steal the password file stored at the server (while not necessarily corrupting the server). However, this by itself does not directly provide the actual password to the adversary. However, after this point the adversary is allowed to perform OfflineTestPwd tests to mimic a similar capability in the real world (in fact, the ideal world adversary is even allowed to perform OfflineTestPwd tests before it steals the password file, but it does not get a confirmation of the guess being correct until after it steals the password file).

Moreover, after the "steal password file" event the adversary is also allowed to impersonate the server to a *correctly guessed* client, even without providing the actual password (as it can clearly do so in the real world). However, compromising impersonation of the client still requires providing a correct password. This differentiation in capabilities also becomes important when characterizing the complexity of a simulator in terms of the real world adversary, as we will see later.

The $\mathcal{F}_{\text{PAKE}}$ functionality for UC-PAKE was a single-session functionality. However, asymmetric PAKE requires that a password file be used across multiple sessions, so the $\mathcal{F}_{\mathsf{apwKE}}$ functionality for UC-APAKE is defined as a multiple-session functionality. Note that this cannot be accomplished simply using composition with joint state [CR03] because the functionality itself requires shared state that needs to be maintained between sessions.

The complete UC-APAKE functionality $\mathcal{F}_{\mathsf{apwKE}}$ is described in detail in Fig. 5 in the Appendix.

7.2 UC-APAKE based on VPAKE and Smooth-NIZK

We now design an asymmetric UC-PAKE based on Verifier-based PAKE or VPAKE of Benhamooda and Pointcheval [BP13] and the novel Smooth NIZK proofs. The essential idea of [BP13] is that while the Client holds the actual password, the Server does not hold password in the clear. Instead the Server stores a hard to invert function called PHash (password hash) evaluated over the password and a random "salt" (PSalt) published in the CRS. While executing a session, the client sends encryptions of the password or another function called PPreHash (password pre-hash) evaluated on the password. Correspondingly, the server sends encryptions of the stored PHash.

Generate $\mathbf{g} \leftarrow \mathbb{G}_1$; $a_1, a_2, b_{\mathsf{C}}, b_{\mathsf{S}} \leftarrow \mathbb{Z}_q$ and let $\rho = \{\mathbf{a}_1 = \mathbf{g}^{a_1}, \mathbf{a}_2 = \mathbf{g}^{a_2}, \mathbf{b}_{\mathsf{C}} = \mathbf{g}^{b_{\mathsf{C}}}, \mathbf{b}_{\mathsf{S}} = \mathbf{g}^{b_{\mathsf{S}}}\}$.

Define languages $\begin{bmatrix} L_{\mathsf{C}} = \{(R, S, H) \mid \exists r, p : R = \mathbf{g}^r, S = \mathbf{a}_1^r \mathbf{b}_{\mathsf{C}}^p, H = \mathbf{b}_{\mathsf{S}}^p\} \\ L_{\mathsf{S}} = \{(R, S) \mid \exists r : R = \mathbf{g}^r, S = \mathbf{a}_2^r\} \end{bmatrix}$

Let $(\mathsf{hk}_{\mathsf{C}}, \mathsf{hp}_{\mathsf{C}}) \leftarrow \mathsf{sphf}(L_{\mathsf{C}}).\mathsf{hkgen}$ and $(\mathsf{hp}_{\mathsf{S}}, \mathsf{hk}_{\mathsf{S}}) \leftarrow \mathsf{sphf}(L_{\mathsf{S}}).\mathsf{hkgen}$.

Define languages:

$$\begin{bmatrix} L_{\mathsf{C}}^{+} = \{(R, S, H, T, l) \mid \exists r, p : R = \mathbf{g}^r, S = \mathbf{a}_1^r \mathbf{b}_{\mathsf{C}}^p, H = \mathbf{b}_{\mathsf{S}}^p, T = \mathsf{sphf.pubH}(\mathsf{hp}_{\mathsf{C}}, \langle R, S, H\rangle, l; r, p)\} \\ L_{\mathsf{S}}^{+} = \{(R, S, T, l) \mid \exists r : R = \mathbf{g}^r, S = \mathbf{a}_2^r, T = \mathsf{sphf.pubH}(\mathsf{hp}_{\mathsf{S}}, \langle R, S\rangle, l; r)\} \end{bmatrix}$$

Let $(\mathsf{pargen}_P, \mathsf{crsgen}_P, \mathsf{prover}_P, \mathsf{ver}_P)$ be Smooth QA-NIZKs for languages L_P^{+}, with $P \in \{C, S\}$.
Let $\textsc{crs}_P \leftarrow \mathsf{crsgen}_P(\rho)$ and $\mathcal{H}$ be a collision resistant hash function.
Let $\mathcal{RO}$ be a random oracle and let $\mathrm{phash} = \mathcal{RO}(sid, P_i, P_j, \mathrm{pwd})$.

$$\mathrm{CRS} := (\rho, \mathsf{hp}_{\mathsf{C}}, \mathsf{hp}_{\mathsf{S}}, \textsc{crs}_{\mathsf{C}}, \textsc{crs}_{\mathsf{S}}, \mathcal{H}).$$
$$\text{Server Persistent State} := \mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}.$$

Client P_i	Network
Input $(\mathsf{CltSession}, sid, \mathsf{ssid}, P_i, P_j, \mathrm{pwd})$. Choose $r_1 \leftarrow \mathbb{Z}_q$ and $(\textsc{hk}_1, \textsc{hp}_1) \leftarrow \mathsf{ver}_{\mathsf{S}}.\mathsf{hkgen}(\textsc{crs}_{\mathsf{S}})$. Set $R_1 = \mathbf{g}^{r_1}$, $S_1 = \mathbf{a}_1^{r_1} \mathbf{b}_{\mathsf{C}}^{\mathrm{phash}}$, $T_1 = \mathsf{sphf}_{\mathsf{C}}.\mathsf{pubH}(\mathsf{hp}_{\mathsf{C}}, \langle R_1, S_1, \mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}\rangle, i_1; r_1, \mathrm{phash})$, $W_1 = \mathsf{prover}_{\mathsf{C}}(\textsc{crs}_{\mathsf{C}}, \langle R_1, S_1, \mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}, T_1, i_1\rangle; r_1, \mathrm{phash})$, where $i_1 = \mathcal{H}(sid, \mathsf{ssid}, P_i, P_j, R_1, S_1, \textsc{hp}_1)$. Erase r_1, send $(R_1, S_1, T_1, \textsc{hp}_1)$ and retain $(W_1, \textsc{hk}_1)$.	$\xrightarrow{R_1, S_1, T_1, \textsc{hp}_1}$ P_j
Receive $(R_2', S_2', T_2', \textsc{hp}_2')$. If any of $R_2', S_2', T_2', \textsc{hp}_2'$ is not in their respective group or is 1, set $\mathrm{sk}_1 \xleftarrow{\$} \mathbb{G}_T$, else compute $i_2' = \mathcal{H}(sid, \mathsf{ssid}, P_j, P_i, R_2', S_2', \textsc{hp}_2')$, and $\mathrm{sk}_1 = \mathsf{ver}_{\mathsf{S}}.\mathsf{privH}(\textsc{hk}_1, \langle R_2', S_2'/\mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}, T_2', i_2'\rangle) \cdot \mathsf{ver}_{\mathsf{C}}.\mathsf{pubH}(\textsc{hp}_2', W_1)$. Output $(sid, \mathsf{ssid}, \mathrm{sk}_1)$.	$\xleftarrow{R_2', S_2', T_2', \textsc{hp}_2'}$ P_j

Server P_j	Network
Input $(\mathsf{SvrSession}, sid, \mathsf{ssid}, P_j, P_i, \text{Server Persistent State})$. Choose $r_2 \leftarrow \mathbb{Z}_q$ and $(\textsc{hk}_2, \textsc{hp}_2) \leftarrow \mathsf{ver}_{\mathsf{C}}.\mathsf{hkgen}(\textsc{crs}_{\mathsf{C}})$. Set $R_2 = \mathbf{g}^{r_2}$, $S_2 = \mathbf{a}_2^{r_2} \mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}$, $T_2 = \mathsf{sphf}_{\mathsf{S}}.\mathsf{pubH}(\mathsf{hp}_{\mathsf{S}}, \langle R_2, S_2/\mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}\rangle, i_2; r_2)$, $W_2 = \mathsf{prover}_{\mathsf{S}}(\textsc{crs}_{\mathsf{S}}, \langle R_2, S_2/\mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}, T_2, i_2\rangle; r_2)$, where $i_2 = \mathcal{H}(sid, \mathsf{ssid}, P_j, P_i, R_2, S_2, \textsc{hp}_2)$. Erase r_2, send $(R_2, S_2, T_2, \textsc{hp}_2)$ and retain $(W_2, \textsc{hk}_2)$.	$\xrightarrow{R_2, S_2, T_2, \textsc{hp}_2}$ P_i
Receive $(R_1', S_1', T_1', \textsc{hp}_1')$. If any of $R_1', S_1', T_1', \textsc{hp}_1'$ is not in their respective group or is 1, set $\mathrm{sk}_2 \xleftarrow{\$} \mathbb{G}_T$, else compute $i_1' = \mathcal{H}(sid, \mathsf{ssid}, P_i, P_j, R_1', S_1', \textsc{hp}_1')$, and $\mathrm{sk}_2 = \mathsf{ver}_{\mathsf{C}}.\mathsf{privH}(\textsc{hk}_2, \langle R_1', S_1', \mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}, T_1', i_1'\rangle) \cdot \mathsf{ver}_{\mathsf{S}}.\mathsf{pubH}(\textsc{hp}_1', W_2)$. Output $(sid, \mathsf{ssid}, \mathrm{sk}_2)$.	$\xleftarrow{R_1', S_1', T_1', \textsc{hp}_1'}$ P_i

Fig. 4. Single round RO-hybrid UC-APAKE protocol under SXDH assumption.

Of course, some kind of zero-knowledge proof must accompany these encryptions, and to that end [BP13] can utilize the new smooth projective hash functions (SPHF) for CCA2-encryption [BBC$^+$13] such as Cramer-Shoup encryption [CS02]. In each session, both parties generate fresh SPHF private and projection keys (to be employed on incoming messages). The projection key is sent (piggy-backed) along with the encrypted message. If the encrypted messages use the correct password (meaning both parties have the same password or its PHash), then SPHF computed on the message by the receiving party using the SPHF hash key it generated equals the SPHF computed on the message by the sending party using the SPHF projection key it received. Thus, these SPHF hashes can be used to compute the session key. Smoothness property of the SPHF guarantees security of the VPAKE scheme.

Unfortunately, each party must retain the witness used in the CCA2 encryption, as computing the SPHF projection-hash of its outgoing encrypted message using the received projection key requires this witness. In the strong simulation paradigm of universally composable security, this leads to a problem if an Adversary can corrupt a session dynamically after the outgoing message has been sent and the incoming message has not yet been received. Thus, this SPHF methodology can only handle static corruption. While Jutla and Roy [JR15] have recently given an efficient UC-PAKE protocol which can handle dynamic corruption, the construction uses ideas from dual-system simulation-sound QA-NIZK that they introduce there. These ideas are rather intricate and do not seem to allow a modular or generic design of such UC password-authenticated protocols.

In this paper, we show that the new notion of Smooth QA-NIZK allows easy to understand (and equally efficient) modular or generic design. Just as QA-NIZK proofs can be seen as generalization of projective hash proof systems to public verifiability (and also assuring zero-knowledge), the novel notion of Smooth QA-NIZK naturally generalizes the notion of smooth projective hash functions where instead of the witness, the publicly verifiable proof can be used to evaluate the projection-hash. The zero-knowledge property of this publicly verifiable proof assures that this proof and hence the projection-hash can be generated by a simulator with no access to the witness. In particular, each party in the UC-PAKE protocol can generate an encryption of the password and generate this publicly verifiable QA-NIZK proof, send the encryption to the other party, erase the witness and retain just the proof for later generation of session key.

The natural question that arises is whether one needs a notion of smooth-soundness under simulation. Indeed, one does need some form of unbounded simulation-soundness as the UC simulator generates QA-NIZK proofs on non-language members without access to the password. Unfortunately, the recent efficient unbounded simulation sound QA-NIZK construction of [KW15] does not extend to be smooth under unbounded simulation (or at least current techniques do not seem to allow one to prove so). The dual-system simulation sound QA-NIZK [JR15] does satisfy smoothness property, but it would need introduction of various new intricate definitions and complicated proofs. One may also ask whether CCA2 encryption by itself provides the required simulation soundness, but that is also not the case, as CCA2 encryption by itself does not give a privately-verifiable (say, via its underlying SPHF as in Cramer-Shoup encryption) proof that it is the password that is being encrypted.

In light of this, it turns out that the simplest way to design the UC-APAKE (or UC-PAKE) protocol is to use an El-Gamal encryption of the password (or its $\mathsf{PPreHash}$ or PHash) and augment it with an SPHF proof of its consistency, and finally a Smooth QA-NIZK on this augmented El-Gamal encryption. (If the reader is interested in the simpler UC-PAKE protocol secure under dynamic corruption in the new Smooth QA-NIZK framework, the UC-PAKE definition and protocol are provided in Appendix E).

We will also need the random oracle hybrid model to achieve the goal of a UC-APAKE protocol, as explained next. The focus of [BP13] was to design protocols which can be proven secure in the standard model. They formalized a security notion for APAKEs modifying the game-based BPR model [BPR00]. However, our focus is to construct an APAKE protocol in the UC model. In the UC model of [GMR06], the UC simulator must be able to detect offline password guess attempts of the adversary. This is not possible in the standard model as offline tests can be internally performed by the adversary. In order to intercept offline tests by the adversary, it thus becomes inevitable to use an idealized model, such as the random oracle model.

So in particular, we adapt the random oracle-based password hashing scheme of [BP13]. In the scheme, the public parameters are $param = \mathbf{b}_{\mathsf{C}}, \mathbf{b}_{\mathsf{S}}$ randomly sampled from $\mathbb{G}_1$ and a random oracle $\mathcal{RO}$. Define phash $= \mathcal{RO}(sid$, Client-id, Server-id, pwd), where Client-id, Server-id are the ids of the participating parties, *sid* is the common session-id for all sessions between these parties and pwd is the password of the client. We set:

$$\begin{aligned}&\mathsf{PPreHash}(param, \mathrm{pwd}) = \mathbf{b}_{\mathsf{C}}^{\mathrm{phash}}\\&\mathsf{PSalt}(param) = \mathbf{b}_{\mathsf{S}}\\&\mathsf{PHash}(param, \mathrm{pwd}) = \mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}\end{aligned}$$

Corresponding to the asymmetric storages of the client and the server, we define the following languages, one for each party, which implicitly check the consistency of correct elements being used:

$$\begin{aligned}L_{\mathsf{C}} &= \{(R, S, H) \mid \exists r, p : R = \mathbf{g}^r, S = \mathbf{a}_1^r \mathbf{b}_{\mathsf{C}}^p, H = \mathbf{b}_{\mathsf{S}}^p\}\\L_{\mathsf{S}} &= \{(R, S) \mid \exists r : R = \mathbf{g}^r, S = \mathbf{a}_2^r\}\end{aligned}$$

We now plug these languages into UC-PAKE methodology described above. The client sends ElGamal encryption of $\mathbf{b}_{\mathsf{C}}^p$, as in (R, S) of L_{C}, while the server supplies the last element H for forming a word of L_{C}. The server sends ElGamal encryption of $\mathbf{b}_{\mathsf{S}}^p$, while the client divides out $\mathbf{b}_{\mathsf{S}}^p$ from the second component to form a word of L_{S}.

The CRS provides public smooth$_2$ SPHF keys for the languages L_{C} and L_{S}, which are used by the client and server respectively to compute T_1 and T_2 for their flows.

Lastly, we use Smooth QA-NIZK proofs for generating a public hash key and a private hash key over the above languages augmented with the SPHFs as below:

$$L_{\mathsf{C}}^{+} = \left\{ (R, S, H, T, l) \mid \exists r, p : \begin{array}{c} R = \mathbf{g}^r, S = \mathbf{a}_1^r \mathbf{b}_{\mathsf{C}}^p, H = \mathbf{b}_{\mathsf{S}}^p, \\ T = \mathsf{sphf.pubH}(\mathsf{hp}_{\mathsf{C}}, \langle R, S, H \rangle, l; r, p) \end{array} \right\}$$
$$L_{\mathsf{S}}^{+} = \{ (R, S, T, l) \mid \exists r : R = \mathbf{g}^r, S = \mathbf{a}_2^r, T = \mathsf{sphf.pubH}(\mathsf{hp}_{\mathsf{S}}, \langle R, S \rangle, l; r) \}$$

The client generates a Smooth QA-NIZK verification key pair for the server language L_{S}^{+}, retains the private key HK_1 and sends the public key HP_1 along with the ElGamal encryption and the SPHF. The client computes a QA-NIZK proof W_1 of $(R_1, S_1, \mathbf{b}_{\mathsf{S}}^{\text{phash}}, T_1) \in L_{\mathsf{C}}^{+}$ with label $i_1 = \mathcal{H}(sid, \mathsf{ssid}, P_i, P_j, R_1, S_1, T_1, \text{HP}_1)$ and retains that for later key computation.

Similarly, the server generates a Smooth QA-NIZK verification key pair for the client language L_{C}^{+}, retains the private key HK_2 and sends the public key HP_2 along with the ElGamal encryption and the SPHF. The server computes a QA-NIZK proof W_2 of $(R_2, S_2/\mathbf{b}_{\mathsf{S}}^{\text{phash}}, T_2) \in L_{\mathsf{S}}^{+}$ with label $i_2 = \mathcal{H}(sid, \mathsf{ssid}, P_j, P_i, R_2, S_2, T_2, \text{HP}_2)$ and retains that for later key computation.

In the second part of the protocol, after receiving the peer flow, each party computes the final secret key as the product of the private Smooth QA-NIZK hash of the peer flow with own private Smooth QA-NIZK key and the public Smooth QA-NIZK hash of the (retained) QA-NIZK proof of own flow with the peer public Smooth QA-NIZK hash key. Formally the client computes:

$$\mathsf{ver}_{\mathsf{S}}.\mathsf{privH}(\text{HK}_1, \langle R_2', S_2'/\mathbf{b}_{\mathsf{S}}^{\text{phash}}, T_2', i_2' \rangle) \cdot \mathsf{ver}_{\mathsf{C}}.\mathsf{pubH}(\text{HP}_2', W_1).$$

Similarly, the server computes:

$$\mathsf{ver}_{\mathsf{C}}.\mathsf{privH}(\text{HK}_2, \langle R_1', S_1', \mathbf{b}_{\mathsf{S}}^{\text{phash}}, T_1', i_1' \rangle) \cdot \mathsf{ver}_{\mathsf{S}}.\mathsf{pubH}(\text{HP}_1', W_2).$$

Given the completeness property of the Smooth QA-NIZK, it is not difficult to see that legitimately completed peer sessions end up with equal keys. In the next section, we prove that this protocol securely realizes $\mathcal{F}_{\mathsf{apwKE}}$, as stated in the theorem below.

The complete protocol is described in detail in Figure 4. The SPHF sphf is required to be a smooth_2 projective hash function (see Section 2 for definitions). For simplicity, in this paper we focus on constructions based on $\mathcal{D}_1$-MDDH assumptions, and in particular the SXDH assumption.

Theorem 6. *Under the $\mathcal{D}_1$-MDDH assumption SXDH, the protocol in Fig. 4 securely realizes the $\mathcal{F}_{\mathsf{apwKE}}$ functionality in the $(\mathcal{F}_{\text{CRS}}, \mathcal{F}_{\text{RO}})$-hybrid model, in the presence of adaptive corruption adversaries. The number of unique password arguments passed to* TestPwd *and* OfflineTestPwd *of $\mathcal{F}_{\mathsf{apwKE}}$ combined in the ideal world is at most the number of random oracle calls in the $(\mathcal{F}_{\text{CRS}}, \mathcal{F}_{\text{RO}})$-hybrid world.*

The proof is in Appendix D.

References

[BBC$^+$13] Fabrice Benhamouda, Olivier Blazy, Céline Chevalier, David Pointcheval, and Damien Vergnaud. New techniques for SPHFs and efficient one-round PAKE protocols. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part I*, volume 8042 of *LNCS*, pages 449–475. Springer, Heidelberg, August 2013.

[BBH06] Dan Boneh, Xavier Boyen, and Shai Halevi. Chosen ciphertext secure public key threshold encryption without random oracles. In David Pointcheval, editor, *CT-RSA 2006*, volume 3860 of *LNCS*, pages 226–243. Springer, Heidelberg, February 2006.

[BBS04] Dan Boneh, Xavier Boyen, and Hovav Shacham. Short group signatures. In Matthew Franklin, editor, *CRYPTO 2004*, volume 3152 of *LNCS*, pages 41–55. Springer, Heidelberg, August 2004.

[BF01] Dan Boneh and Matthew K. Franklin. Identity-based encryption from the Weil pairing. In Joe Kilian, editor, *CRYPTO 2001*, volume 2139 of *LNCS*, pages 213–229. Springer, Heidelberg, August 2001.

[BFM88] Manuel Blum, Paul Feldman, and Silvio Micali. Non-interactive zero-knowledge and its applications (extended abstract). In *20th ACM STOC*, pages 103–112. ACM Press, May 1988.

[BM92] Steven M. Bellovin and Michael Merritt. Encrypted key exchange: Password-based protocols secure against dictionary attacks. In *1992 IEEE Symposium on Security and Privacy*, pages 72–84. IEEE Computer Society Press, May 1992.

[BM93] Steven M. Bellovin and Michael Merritt. Augmented encrypted key exchange: A password-based protocol secure against dictionary attacks and password file compromise. In V. Ashby, editor, *ACM CCS 93*, pages 244–250. ACM Press, November 1993.

[BMP00] Victor Boyko, Philip D. MacKenzie, and Sarvar Patel. Provably secure password-authenticated key exchange using Diffie-Hellman. In Bart Preneel, editor, *EUROCRYPT 2000*, volume 1807 of *LNCS*, pages 156–171. Springer, Heidelberg, May 2000.

[Boy09] Xavier Boyen. HPAKE: Password authentication secure against cross-site user impersonation. In Juan A. Garay, Atsuko Miyaji, and Akira Otsuka, editors, *CANS 09*, volume 5888 of *LNCS*, pages 279–298. Springer, Heidelberg, December 2009.

[BP13] Fabrice Benhamouda and David Pointcheval. Verifier-based password-authenticated key exchange: New models and constructions. *IACR Cryptology ePrint Archive*, 2013:833, 2013.

[BPR00] Mihir Bellare, David Pointcheval, and Phillip Rogaway. Authenticated key exchange secure against dictionary attacks. In Bart Preneel, editor, *EUROCRYPT 2000*, volume 1807 of *LNCS*, pages 139–155. Springer, Heidelberg, May 2000.

[BR93] Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In V. Ashby, editor, *ACM CCS 93*, pages 62–73. ACM Press, November 1993.

[Can01] Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *42nd FOCS*, pages 136–145. IEEE Computer Society Press, October 2001.

[CHK+05] Ran Canetti, Shai Halevi, Jonathan Katz, Yehuda Lindell, and Philip D. MacKenzie. Universally composable password-based key exchange. In Ronald Cramer, editor, *EUROCRYPT 2005*, volume 3494 of *LNCS*, pages 404–421. Springer, Heidelberg, May 2005.

[CLL+13] Jie Chen, Hoon Wei Lim, San Ling, Huaxiong Wang, and Hoeteck Wee. Shorter IBE and signatures via asymmetric pairings. In Michel Abdalla and Tanja Lange, editors, *PAIRING 2012*, volume 7708 of *LNCS*, pages 122–140. Springer, Heidelberg, May 2013.

[CR03] Ran Canetti and Tal Rabin. Universal composition with joint state. In Dan Boneh, editor, *CRYPTO 2003*, volume 2729 of *LNCS*, pages 265–281. Springer, Heidelberg, August 2003.

[CS98] Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In Hugo Krawczyk, editor, *CRYPTO'98*, volume 1462 of *LNCS*, pages 13–25. Springer, Heidelberg, August 1998.

[CS02] Ronald Cramer and Victor Shoup. Universal hash proofs and a paradigm for adaptive chosen ciphertext secure public-key encryption. In Lars R. Knudsen, editor, *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 45–64. Springer, Heidelberg, April / May 2002.

[EHK+13] Alex Escala, Gottfried Herold, Eike Kiltz, Carla Ràfols, and Jorge Villar. An algebraic framework for Diffie-Hellman assumptions. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part II*, volume 8043 of *LNCS*, pages 129–147. Springer, Heidelberg, August 2013.

[FLR+10] Marc Fischlin, Anja Lehmann, Thomas Ristenpart, Thomas Shrimpton, Martijn Stam, and Stefano Tessaro. Random oracles with(out) programmability. In Masayuki Abe, editor, *ASIACRYPT 2010*, volume 6477 of *LNCS*, pages 303–320. Springer, Heidelberg, December 2010.

[GJKR07] Rosario Gennaro, Stanislaw Jarecki, Hugo Krawczyk, and Tal Rabin. Secure distributed key generation for discrete-log based cryptosystems. *Journal of Cryptology*, 20(1):51–83, January 2007.

[GMR88] Shafi Goldwasser, Silvio Micali, and Ronald L. Rivest. A digital signature scheme secure against adaptive chosen-message attacks. *SIAM J. Comput.*, 17(2):281–308, 1988.

[GMR06] Craig Gentry, Philip MacKenzie, and Zulfikar Ramzan. A method for making password-based key exchange resilient to server compromise. In Cynthia Dwork, editor, *CRYPTO 2006*, volume 4117 of *LNCS*, pages 142–159. Springer, Heidelberg, August 2006.

[GS08] Jens Groth and Amit Sahai. Efficient non-interactive proof systems for bilinear groups. In Nigel P. Smart, editor, *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 415–432. Springer, Heidelberg, April 2008.

[HK98] Shai Halevi and Hugo Krawczyk. Public-key cryptography and password protocols. In *ACM CCS 98*, pages 122–131. ACM Press, November 1998.

[JG04] Shaoquan Jiang and Guang Gong. Password based key exchange with mutual authentication. In Helena Handschuh and Anwar Hasan, editors, *SAC 2004*, volume 3357 of *LNCS*, pages 267–279. Springer, Heidelberg, August 2004.

[JR12] Charanjit S. Jutla and Arnab Roy. Relatively-sound NIZKs and password-based key-exchange. In Marc Fischlin, Johannes Buchmann, and Mark Manulis, editors, *PKC 2012*, volume 7293 of *LNCS*, pages 485–503. Springer, Heidelberg, May 2012.

[JR13] Charanjit S. Jutla and Arnab Roy. Shorter quasi-adaptive NIZK proofs for linear subspaces. In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT 2013, Part I*, volume 8269 of *LNCS*, pages 1–20. Springer, Heidelberg, December 2013.

[JR14] Charanjit S. Jutla and Arnab Roy. Switching lemma for bilinear tests and constant-size NIZK proofs for linear subspaces. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part II*, volume 8617 of *LNCS*, pages 295–312. Springer, Heidelberg, August 2014.

[JR15] Charanjit S. Jutla and Arnab Roy. Dual-system simulation-soundness with applications to UC-PAKE and more. In Tetsu Iwata and Jung Hee Cheon, editors, *ASIACRYPT 2015, Part I*, volume 9452 of *LNCS*, pages 630–655. Springer, Heidelberg, November / December 2015.

[KV11] Jonathan Katz and Vinod Vaikuntanathan. Round-optimal password-based authenticated key exchange. In Yuval Ishai, editor, *TCC 2011*, volume 6597 of *LNCS*, pages 293–310. Springer, Heidelberg, March 2011.

[KW15] Eike Kiltz and Hoeteck Wee. Quasi-adaptive NIZK for linear subspaces revisited. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part II*, volume 9057 of *LNCS*, pages 101–128. Springer, Heidelberg, April 2015.

[LPJY14] Benoît Libert, Thomas Peters, Marc Joye, and Moti Yung. Non-malleability from malleability: Simulation-sound quasi-adaptive NIZK proofs and CCA2-secure encryption from homomorphic signatures. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 514–532. Springer, Heidelberg, May 2014.

[Mac01] Philip D. MacKenzie. More efficient password-authenticated key exchange. In David Naccache, editor, *CT-RSA 2001*, volume 2020 of *LNCS*, pages 361–377. Springer, Heidelberg, April 2001.

[Wat09] Brent Waters. Dual system encryption: Realizing fully secure IBE and HIBE under simple assumptions. In Shai Halevi, editor, *CRYPTO 2009*, volume 5677 of *LNCS*, pages 619–636. Springer, Heidelberg, August 2009.

A Proof of Boosting Theorem

Theorem 1. (re-stated) If a matrix distribution $\mathcal{D}_k$ on $\mathbb{Z}_q^{(k+1)\times k}$ is boostable to a matrix distribution $\mathcal{D}_{l,k}$ on $\mathbb{Z}_q^{l\times k}$ then the $\mathcal{D}_k$-MDDH assumption implies the $\mathcal{D}_{l,k}$-MDDH assumption.

Proof. We prove this by a sequence of hybrids, where in the i-th hybrid we transform row $k+i$ from that of $[\mathbf{B}\mathbf{s}]$ to uniformly random. We start off with $i=0$, where we have the real output $[\mathbf{B}\mathbf{s}]$ and end with $i=l-k$ where we have the fake output which is uniformly random in $\mathbb{Z}_q^l$.

The i-th hybrid $([\mathbf{B}],[\mathbf{b}])$ is computed as follows. We sample $[\mathbf{A}]$ from $\mathcal{D}_k$ and $\mathbf{s}$ from $\mathbb{Z}_q^k$. We set $[\bar{\mathbf{B}}]$ as $[\bar{\mathbf{A}}]$ and, if $i \neq 0$, the row i of $[\underline{\mathbf{B}}]$ as the row i of $\mathbf{F}[\mathbf{A}]$. All other rows $j \neq i$ of $[\underline{\mathbf{B}}]$ are set to the j-th row of $\mathbf{E}[\bar{\mathbf{A}}]$. We set the top k elements of $[\mathbf{b}]$ to be $[\bar{\mathbf{A}}\mathbf{s}]$ and choose all the $(k+j)$-th elements, where $j<i$, of $[\mathbf{b}]$ uniformly at random from $\mathbb{Z}_q$. If $i \neq 0$, we set the $(k+i)$-th element of $[\mathbf{b}]$ to be the i-th element of $\mathbf{F}[\mathbf{A}\mathbf{s}]$. For all $j>i$, we set the $(k+j)$-th element of $[\mathbf{b}]$ to be the j-th element of $\mathbf{E}[\bar{\mathbf{A}}\mathbf{s}]$. To summarize, $[\mathbf{b}]$ is computed as:

$$\begin{bmatrix} [\bar{\mathbf{A}}\mathbf{s}] \\ \$ \\ \vdots \\ \$ \\ (\mathbf{F}[\mathbf{A}\mathbf{s}])_i \\ (\mathbf{E}[\bar{\mathbf{A}}\mathbf{s}])_{j=(i+1) \text{ to } (l-k)} \end{bmatrix}$$

We observe that the 0-th hybrid has the distribution of $([\mathbf{B}],[\mathbf{B}\mathbf{s}])$ and the $(l-k)$-th hybrid has the distribution of $([\mathbf{B}],[\mathbf{s}'])$, with $\mathbf{s}'$ uniform in $\mathbb{Z}_q^l$.

Now, $(\mathbf{F}[\mathbf{A}\mathbf{s}])_i = (\mathbf{F}_l)_i[\bar{\mathbf{A}}\mathbf{s}] + (\mathbf{F}_r)_i[\underline{\mathbf{A}}\mathbf{s}]$, where $\mathbf{F}_l$ is the first k-column submatrix of $\mathbf{F}$ and $\mathbf{F}_r$ is the last column of $\mathbf{F}$. Suppose we are given a $\mathcal{D}_k$-MDDH challenge $([\mathbf{A}], \boldsymbol{\chi} = [\mathbf{A}\mathbf{s}] \textbf{ or } [\mathbf{s}'])$. If $\boldsymbol{\chi} = [\mathbf{A}\mathbf{s}]$, then $(\mathbf{F}\boldsymbol{\chi})_i$ is distributed as $(\mathbf{F}[\mathbf{A}\mathbf{s}])_i$. Else, if $\boldsymbol{\chi} = [\mathbf{s}']$, then $(\mathbf{F}\boldsymbol{\chi})_i$ is distributed uniformly randomly in $\mathbb{Z}_q$, since $(\mathbf{F}_r)_i$ is overwhelmingly non-zero by design. Next we transition to an intermediate hybrid i' where $[\mathbf{b}]$ is computed as:

$$\begin{bmatrix} [\bar{\mathbf{A}}\mathbf{s}] \\ \$ \\ \vdots \\ \$ \\ \$ \\ (\mathbf{E}[\bar{\mathbf{A}}\mathbf{s}])_{j=(i+1) \text{ to } (l-k)} \end{bmatrix}$$

As shown above, the hybrid i' is indistinguishable from hybrid i by the $\mathcal{D}_k$-MDDH assumption. Next we transition to the hybrid $i+1$ where $[\mathbf{b}]$ is computed as:

$$\begin{bmatrix} [\bar{\mathbf{A}}\mathbf{s}] \\ \$ \\ \vdots \\ \$ \\ \$ \\ (\mathbf{F}[\mathbf{A}\mathbf{s}])_{(i+1)} \\ (\mathbf{E}[\bar{\mathbf{A}}\mathbf{s}])_{j=(i+2) \text{ to } (l-k)} \end{bmatrix}$$

The hybrid $i+1$ is indistinguishable from hybrid i', as $\mathbf{E}\bar{\mathbf{A}}$ is identically distributed as $\mathbf{FA}$. The theorem is thus established by chaining all the hybrids.

B Proof of Smooth Soundness of QA-NIZK for Linear Subspaces

Theorem 2 The algorithms $(\mathsf{K}_0, \mathsf{K}_1, \mathsf{P}, \mathsf{V})$ constitute a smooth computationally - sound Quasi-Adaptive NIZK proof system for linear subspace languages $\{L_{[\mathbf{M}]_1}\}$ with parameters $[\mathbf{M}]_1$ sampled from a robust and efficiently witness-samplable distribution $\mathcal{D}$ over the associated parameter language Lpar, given any group generation algorithm for which the $\mathcal{D}_k$-MDDH assumption holds for group $\mathbb{G}_2$.

The proofs of completeness, zero knowledge and soundness are same as [KW15]. The proof of smooth soundness follows.

Smooth Soundness: First, note that the range of privH is exponential in the security parameter, for otherwise an adversarial circuit can compute discrete logarithms with non-negligible probability. We prove smoothness by transforming the system over a sequence of games. Game $\mathbf{G}_0$ just replicates the construction, but samples $\mathbf{A}$ from a distribution $\mathcal{D}_{k+n-t,k}$ obtained by boosting the given distribution $\mathcal{D}_k$ by Corollary 1. The construction only uses the top $k \times k$ sub-matrix $\bar{\mathbf{A}}$ of the sample which is distributed identically for both $\mathcal{D}_k$ and $\mathcal{D}_{k+n-t,k}$. Let $\underline{\mathbf{A}}$ be the bottom $(n-t) \times k$ sub-matrix of $\mathbf{A}$.

In Game $\mathbf{G}_1$, the challenger efficiently samples $[\mathbf{M}]_1$ according to distribution $\mathcal{D}$, along with witness $\mathbf{M}$ (since $\mathcal{D}$ is an efficiently witness samplable distribution). Since $\mathbf{M}$ is an $n \times t$ dimensional rank t matrix, there is a rank $n-t$ matrix $\mathbf{M}^\perp$ of dimension $n \times (n-t)$ whose columns form a complete basis for the kernel of $\mathbf{M}^\top$, which means $\mathbf{M}^\top \mathbf{M}^\perp = 0^{t \times (n-t)}$. In this game, the NIZK CRS is computed as follows: Generate matrix $\mathbf{K'}^{\,n \times k}$ and compute the matrix $\mathbf{T}^{(n-t)\times k}$, such that $\mathbf{T}\bar{\mathbf{A}} = \underline{\mathbf{A}}$. Implicitly set: $\mathbf{K} = \mathbf{K}' + \mathbf{M}^\perp \mathbf{T}$. Therefore we have,

$$\mathbf{CRS}_p^{t\times k} = [\mathbf{M}^\top \mathbf{K}]_1 = [\mathbf{M}^\top(\mathbf{K}' + \mathbf{M}^\perp \mathbf{T})]_1 = [\mathbf{M}^\top \mathbf{K}']_1$$

$$[\mathbf{C}]_2^{n\times k} = [(\mathbf{K}' + \mathbf{M}^\perp \mathbf{T})\bar{\mathbf{A}}]_2 = \mathbf{K}'[\bar{\mathbf{A}}]_2 + \mathbf{M}^\perp[\underline{\mathbf{A}}]_2,$$

$$\mathsf{hk} = [\mathbf{C}]_2\ \mathrm{s}, \quad \mathsf{hp} = [\bar{\mathbf{A}}]_2\ \mathrm{s}$$

In Game $\mathbf{G}_2$, we sample fresh random vectors $\mathbf{s}'$ in $\mathbb{Z}_q^k$ and $\mathbf{s}''$ in $\mathbb{Z}_q^{n-t}$ and modify the simulated computations as follows:

$$\mathsf{CRS}_p^{t\times k} = [\mathbf{M}^\top \mathbf{K}']_1, \qquad [\mathbf{C}]_2^{n\times k} = \mathbf{K}'[\bar{\mathbf{A}}]_2 + \mathbf{M}^\perp[\underline{\mathbf{A}}]_2,$$

$$\mathsf{hk} = \mathbf{K}'[\mathbf{s}']_2 + \mathbf{M}^\perp[\mathbf{s}'']_2, \quad \mathsf{hp} = [\mathbf{s}']_2$$

Given a $\mathcal{D}_{k+n-t,k}$ challenge which is either "real": $([\mathbf{A}]_2, [\bar{\mathbf{A}}\mathbf{s}]_2, [\underline{\mathbf{A}}\mathbf{s}]_2)$ or "fake": $([\mathbf{A}]_2, [\mathbf{s}']_2, [\mathbf{s}'']_2)$, we observe that the real tuple can be used to simulate Game $\mathbf{G}_1$, while the fake tuple can be used to simulate Game $\mathbf{G}_2$. Thus the games $\mathbf{G}_1$ and $\mathbf{G}_2$ are indistinguishable by the $\mathcal{D}_{k+n-t,k}$-MDDH assumption, which in turn is implied by the $\mathcal{D}_k$-MDDH assumption by Theorem 1.

Now in Game $\mathbf{G}_2$ we have,

$$\mathsf{privH}(\mathsf{hk}, [\mathbf{y}^*]_1) = e\left([\mathbf{y}^{*\top}]_1, \mathbf{K}'[\mathbf{s}']_2 + \mathbf{M}^\perp[\mathbf{s}'']_2\right)$$

For the oracle queries where $[\mathbf{y}^*]_1 \in L_{[\mathbf{M}]_1}$, we have $\mathbf{y}^{*\top}\mathbf{M}^\perp = 0^{1\times(n-t)}$. Hence the simulator responds with $e\left([\mathbf{y}^*]_1^\top, \mathbf{K}'[\mathbf{s}']_2\right)$. Note that $\mathbf{s}''$ does not appear in this response.

For the adversary supplied $[\mathbf{y}^*]_1 \notin L_{[\mathbf{M}]_1}$, we have $\mathbf{y}^{*\top}\mathbf{M}^\perp \neq 0^{1\times(n-t)}$. Therefore $\mathsf{privH}(\mathsf{hk}, \mathbf{y}^*)$ is uniformly random, as $\mathbf{s}''$ is independently random of everything else given to the adversary. □

C Tagged Affine Smooth QA-NIZK Proof

We now give a proof for smooth soundness. The proofs of completeness, zero knowledge and soundness are similar to Theorem 2.

Smooth Soundness: We prove smooth soundness by transforming the system over a sequence of games. Game $\mathbf{G}_0$ just replicates the construction, but samples $\mathbf{A}$ from a distribution $\mathcal{D}_{k+l+l',k}$ obtained by boosting the given distribution $\mathcal{D}_k$ by Corollary 1. The construction only uses the top $k\times k$ sub-matrix $\bar{\mathbf{A}}$ of the sample which is distributed identically for both $\mathcal{D}_k$ and $\mathcal{D}_{k+l+l',k}$. Let $\underline{\mathbf{A}}_1$ and $\underline{\mathbf{A}}_2$ be the next $l\times k$ and bottom $l'\times k$ sub-matrices of $\mathbf{A}$, respectively.

In Game $\mathbf{G}_1$, the challenger efficiently samples $(\mathbf{M}_0, \mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{a})$ according to distribution $\mathcal{D}$. In this game, the NIZK CRS is computed as follows: Generate matrix $\mathbf{K}'^{\,n\times k}$ and compute the matrices $\mathbf{T}_1$ and $\mathbf{T}_2$, such that $\mathbf{T}_1\bar{\mathbf{A}} = \underline{\mathbf{A}}_1$ and $\mathbf{T}_2\bar{\mathbf{A}} = \underline{\mathbf{A}}_2$. Implicitly set:

$$\begin{gathered}\mathbf{K}_1 = \mathbf{K}'_1 + \mathbf{T}_1, \quad \mathbf{K}_2 = \mathbf{K}'_2 + \mathbf{T}_2,\\ \mathbf{L}_1 = \mathbf{L}'_1 - \mathbf{M}_0^{-\top}\mathbf{M}_1^\top\mathbf{T}_1 - \mathbf{M}_0^{-\top}\mathbf{M}_2^\top\mathbf{T}_2, \quad \mathbf{L}_2 = \mathbf{L}'_2 - \mathbf{M}_0^{-\top}\mathbf{M}_3^\top\mathbf{T}_2,\\ \mathbf{l}_3 = \mathbf{l}'_3 - \mathbf{a}\cdot\mathbf{T}_1\end{gathered}$$

Therefore we have,

$$\mathsf{CRS}_p := \begin{pmatrix} [\mathbf{P}_1]_1^{t\times k} = [\mathbf{M}_0^\top\mathbf{L}'_1 + \mathbf{M}_1^\top\mathbf{K}'_1 + \mathbf{M}_2^\top\mathbf{K}'_2]_1, \\ [\mathbf{P}_2]_1^{t\times k} = [\mathbf{M}_0^\top\mathbf{L}'_2 + \mathbf{M}_3^\top\mathbf{K}'_2]_1, \qquad [\mathbf{P}_3]_1^{t\times k} = [\mathbf{l}_3'^\top + \mathbf{a}\cdot\mathbf{K}'_1]_1 \end{pmatrix}$$

$$\mathsf{CRS}_v := \begin{pmatrix} [\bar{\mathbf{A}}]_2, \quad [\mathbf{K}_1'\bar{\mathbf{A}} + \underline{\mathbf{A}}_1]_2, [\mathbf{K}_2'\bar{\mathbf{A}} + \underline{\mathbf{A}}_2]_2, \\ [\mathbf{L}_1'\bar{\mathbf{A}} - \mathbf{M}_0^{-\top}\mathbf{M}_1^\top\underline{\mathbf{A}}_1 - \mathbf{M}_0^{-\top}\mathbf{M}_2^\top\underline{\mathbf{A}}_2]_2, \\ [\mathbf{L}_2'\bar{\mathbf{A}} - \mathbf{M}_0^{-\top}\mathbf{M}_3^\top\underline{\mathbf{A}}_2]_2, \\ [\mathbf{l}_3' \cdot \bar{\mathbf{A}} - \mathbf{a} \cdot \underline{\mathbf{A}}_1]_T \end{pmatrix}$$

$$\text{HK} = \begin{pmatrix} [\mathbf{L}_1'\bar{\mathbf{A}} - \mathbf{M}_0^{-\top}\mathbf{M}_1^\top\underline{\mathbf{A}}_1 - \mathbf{M}_0^{-\top}\mathbf{M}_2^\top\underline{\mathbf{A}}_2]_2\ \mathbf{s} \\ +\text{TAG}\ [\mathbf{L}_2'\bar{\mathbf{A}} - \mathbf{M}_0^{-\top}\mathbf{M}_3^\top\underline{\mathbf{A}}_2]_2\ \mathbf{s}, \\ [\mathbf{K}_1'\bar{\mathbf{A}} + \underline{\mathbf{A}}_1]_2\ \mathbf{s},\ [\mathbf{K}_2'\bar{\mathbf{A}} + \underline{\mathbf{A}}_2]_2\ \mathbf{s}, \\ [\mathbf{l}_3' \cdot \bar{\mathbf{A}} - \mathbf{a} \cdot \underline{\mathbf{A}}_1]_T\ \mathbf{s}, \end{pmatrix}, \qquad \text{HP} = [\bar{\mathbf{A}}]_2\ \mathbf{s}$$

In Game $\mathbf{G}_2$, we sample fresh random vectors $(\mathbf{s}', \mathbf{s}_1'', \mathbf{s}_2'') \leftarrow \mathbb{Z}_q^k \times \mathbb{Z}_q^l \times \mathbb{Z}_q^{l'}$ and modify the simulated computations as follows:

$$\text{HK} = \begin{pmatrix} [(\mathbf{L}_1' + \text{TAG}\ \mathbf{L}_2')\mathbf{s}' - \mathbf{M}_0^{-\top}\mathbf{M}_1^\top\mathbf{s}_1'' - \mathbf{M}_0^{-\top}(\mathbf{M}_2 + \text{TAG}\ \mathbf{M}_3)^\top\mathbf{s}_2'']_2, \\ [\mathbf{K}_1'\mathbf{s}' + \mathbf{s}_1'']_2,\ [\mathbf{K}_2'\mathbf{s}' + \mathbf{s}_2'']_2, \\ [\mathbf{l}_3' \cdot \mathbf{s}' - \mathbf{a} \cdot \mathbf{s}_1'']_T, \end{pmatrix}, \quad \text{HP} = [\mathbf{s}']_2$$

Given a $\mathcal{D}_{k+n-t,k}$ challenge which is either "real": $([\mathbf{A}]_2, [\bar{\mathbf{A}}\mathbf{s}]_2, [\underline{\mathbf{A}}\mathbf{s}]_2)$ or "fake": $([\mathbf{A}]_2, [\mathbf{s}']_2, [\mathbf{s}_1'', \mathbf{s}_2'']_2)$, we observe that the real tuple can be used to simulate Game $\mathbf{G}_1$, while the fake tuple can be used to simulate Game $\mathbf{G}_2$. Thus the games $\mathbf{G}_1$ and $\mathbf{G}_2$ are indistinguishable by the $\mathcal{D}_{k+l+l',k}$-MDDH assumption, which in turn is implied by the $\mathcal{D}_k$-MDDH assumption by Theorem 1.

Now in Game $\mathbf{G}_2$ we have,

$$\mathsf{privH}(\mathsf{hk}, (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3,\ \text{TAG})) = e\left(\mathbf{y}_1 \cdot (\mathbf{L}_1' + \text{TAG}\mathbf{L}_2') + \mathbf{y}_2 \cdot \mathbf{K}_1' + \mathbf{y}_3 \cdot \mathbf{K}_2' + \mathbf{l}_3'^{\top}, [\mathbf{s}']_2\right)$$
$$+e\left((\mathbf{y}_2 - \mathbf{a})^\top - \mathbf{y}_1 \cdot \mathbf{M}_0^{-\top}\mathbf{M}_1^\top, [\mathbf{s}_1'']_2\right)$$
$$+e\left(\mathbf{y}_3^\top - \mathbf{y}_1 \cdot \mathbf{M}_0^{-\top}(\mathbf{M}_2^\top + \text{TAG}\ \mathbf{M}_3^\top), [\mathbf{s}_2'']_2\right)$$

For the oracle queries with a word in the language, the last two summands would be zero, and hence the simulator can respond with only the first summand. For the adversary supplied word, which is not in the language, at least one of the last two summands would be non-zero and are completely randomized by $\mathbf{s}_1''$ and $\mathbf{s}_2''$, which are independent of everything else given to the adversary, including $\mathbf{s}'$. □

D Proof of UC-APAKE realization

D.1 Main Idea of the UC Simulator

The UC simulator $\mathcal{S}$ works as follows: It simulates the random oracle calls and records all the query response pairs. It will generate the CRS for $\widehat{\mathcal{F}}_{\text{PAKE}}$ using the real world algorithms, except for the Smooth QA-NIZK, for which it uses the simulated CRS generator. It also retains the private hash keys of the SPHF's. The next main difference is in the simulation of the outgoing message of the real world parties: $\mathcal{S}$ uses a dummy message μ instead of the real password which it does not have access to. Further, it postpones computation of W till the session-key generation time. Finally, another difference is in the processing of the incoming message, where $\mathcal{S}$ decrypts the incoming message R_2', S_2' and runs through the list of random oracle queries to search for a pwd$'$, such that the decryption is $\mathbf{b}_{\mathsf{S}}^{\mathcal{RO}(sid,P_i,P_j,\mathrm{pwd}')}$, which it uses to call the ideal functionality's test function. It next generates an sk similar to how it is generated in the real-world. It sends sk to the ideal functionality to be output to the party concerned.

Since the (R_1, S_1) that it sends out is no longer such that $(R_1, S_1, \mathbf{b}_{\mathsf{S}}^{\mathrm{phash}})$ in the language L_{C}, it has to use the private key of the SPHF in order to compute T_1 on $(R_1, S_1, \mathbf{b}_{\mathsf{S}}^{\mathrm{phash}})$ and the QA-NIZK proof simulator to compute W_1.

There are other special steps designed to simulate stealing the password file and then impersonating the server to the client. Specifically, when the password file is stolen, the simulator still may not know pwd. It then preemptively sets phash to a random value and pretends that this is the random oracle response with the correct pwd query. Later on when there is a successful pwd query, which the simulator can find out by the online or offline testpwd ideal functionality calls, it sets the record accordingly.

In case of a stolen password file, the simulator includes a "Client Only Step" which lets it test (modified) server flows for consistency and call the Impersonate functionality if consistency checks out. The server simulation steps do not include such a step to model the security notion that even if the password file is stolen, the adversary should still not be able to impersonate the client.

D.2 Main Idea of the Proof of UC Realization

The proof that the simulator $\mathcal{S}$ described above simulates the Adversary in the real-world protocol, follows essentially from the properties of the Smooth QA-NIZK and smooth$_2$ SPHF, and we give a broad outline here. The proof will describe various experiments between a challenger $\mathcal{C}$ and the adversary, which we will just assume to be the environment $\mathcal{Z}$ (as the adversary $\mathcal{A}$ can be assumed to be just dummy and following $\mathcal{Z}$'s commands). In the first experiment the challenger $\mathcal{C}$ will just be the combination of the code of the simulator $\mathcal{S}$ above and $\widehat{\mathcal{F}}_{\text{PAKE}}$. In particular, after the environment issues a $\mathsf{CltSession}$ request with a password pwd, the challenger gets that password. So, while in the first experiment, the challenger (copying $\mathcal{S}$) does not use pwd directly, from the next experiment onwards, it can use pwd. Thus, the main goal of the ensuing experiments is to modify the fake tuples $\mathbf{g}^{r_1}, \mathbf{g}^{r'}$ by real tuples (as in real-world) $\mathbf{g}^{r_1}, \mathbf{a}_1^{r_1}\mathbf{b}_{\mathsf{C}}^{\mathrm{phash}}$, since the challenger has access to pwd, and hence phash. This is accomplished by a hybrid argument, modifying one instance at a time using DDH assumption in group $\mathbb{G}_1$.

The guarantee that the client cannot be impersonated by the adversary, even when the password file is stolen is established by noting that $\mathbf{b}_{\mathsf{C}}^{\mathrm{phash}}$, which is what the client encrypts in its flows, is hard to compute given the server persistent state $\mathbf{b}_{\mathsf{S}}^{\mathrm{phash}}$. This is formally captured in the proof by using a DDH transition from $(\mathbf{b}_\mathsf{S}, \mathbf{b}_\mathsf{C}, \mathbf{b}_\mathsf{S}^{\text{phash}}, \mathbf{b}_\mathsf{C}^{\text{phash}})$ to $(\mathbf{b}_\mathsf{S}, \mathbf{b}_\mathsf{C}, \mathbf{b}_\mathsf{S}^{\text{phash}}, \mathbf{b}_\mathsf{C}^{z})$, where z is independently random from phash.

Once all the instances are corrected, i.e. R_1, S_1 are generated as $\mathbf{g}^{r_1}, \mathbf{a}_1^{r_1}\mathbf{b}_\mathsf{C}^{\text{phash}}$, the challenger can switch to the real-world because the tuples $R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}$ are now in the language L_C. This implies that the session keys are generated exactly as in the real-world.

D.3 Adaptive Corruption

The UC protocol described above is also UC-secure against adaptive corruption of parties by the Adversary in the erasure model. In the real-world when the adversary corrupts a client (with a Corrupt command), it gets the internal state of the client. Clearly, if the party has already been invoked with a CltSession command then the password pwd is leaked at the minimum, and hence the ideal functionality $\mathcal{F}_{\text{PAKE}}$ leaks the password to the Adversary in the ideal world. In the protocol described above, the Adversary also gets W_1 and HK_1, as this is the only state maintained by each client between sending $R_1, S_1, T_1, \text{HP}_1$, and the final issuance of session-key. Simulation of HK_1 is easy for the simulator $\mathcal{S}$ since $\mathcal{S}$ generates HK_1 exactly as in the real world. For generating W_1, which $\mathcal{S}$ had postponed to computing till it received an incoming message from the adversary, it can now use the pwd which it gets from $\widehat{\mathcal{F}}_{\text{PAKE}}$ by issuing a Corrupt call to $\widehat{\mathcal{F}}_{\text{PAKE}}$. More precisely, it issues the Corrupt call, and gets pwd, and then calls the QA-NIZK simulator with the tuple $(R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}, T_1, i_1)$ to get W_1. Note that this computation of W_1 is identical to the postponed computation of W_1 in the computation of client factor of sk_1 (which is really used in the output to the environment when $\text{pwd}' = \text{pwd}$).

In case of server corruption, the simulator does not get pwd, but is able to set phash which also enables it to compute W_2 using the QA-NIZK simulator on $(R_2, S_2/\mathbf{b}_\mathsf{S}^{\text{phash}}, T_2, i_2)$.

We first define a simulator which interfaces with the ideal functionality and the adversary and then through a series of experiments convert it to just the real world protocol interacting with the same adversary.

D.4 Simulator for the Protocol

We will assume that the adversary $\mathcal{A}$ in the UC protocol is dummy, and essentially passes back and forth commands and messages from the environment $\mathcal{Z}$. Thus, from now on we will use environment $\mathcal{Z}$ as the real adversary, which outputs a single bit. The simulator $\mathcal{S}$ will be the ideal world adversary for $\mathcal{F}_{\mathsf{apwKE}}$. It is a universal simulator that uses $\mathcal{A}$ as a black box. For each instance (session and a party), we will use a prime, to refer to variables received in the message from $\mathcal{Z}$ (i.e. $\mathcal{A}$). We will call a message *legitimate* if it was not altered by $\mathcal{Z}$, and delivered in the correct session and to the correct party.

Functionality $\mathcal{F}_{\mathsf{apwKE}}$

The functionality $\mathcal{F}_{\mathsf{apwKE}}$ is parameterized by a security parameter k. It interacts with an adversary S and a set of parties via the following queries:

Password Storage and Authentication Sessions

Upon receiving a query $(\mathsf{StorePwdFile}, \boldsymbol{sid}, P_i, pw)$ **from party** P_j**:**
If this is the first StorePwdFile query, store password data record $(\mathsf{file}, P_i, P_j, pw)$ and mark it uncompromised.

Upon receiving a query $(\mathsf{CltSession}, \boldsymbol{sid}, \mathsf{ssid}, P_i, P_j, pw)$ **from party** P_i**:**
Send $(\mathsf{CltSession}, sid, \mathsf{ssid}, P_i, P_j)$ to S. In addition, if this is the first CltSession query for ssid, then store session record $(\mathsf{Clt}, \mathsf{ssid}, P_i, P_j, pw)$ and mark this record fresh.

Upon receiving a query $(\mathsf{SvrSession}, \boldsymbol{sid}, \mathsf{ssid})$ **from party** P_j**:**
If there is a password data record $(\mathsf{file}, P_i, P_j, pw)$, then send $(\mathsf{SvrSession}, sid, \mathsf{ssid}, P_j, P_i)$ to S, and if this is the first SvrSession query for ssid, store session record $(\mathsf{Svr}, \mathsf{ssid}, P_j, P_i, pw)$, and mark it fresh.

Stealing Password Data

Upon receiving a query $(\mathsf{StealPwdFile}, \boldsymbol{sid})$ **from adversary** S**:**
If there is no password data record reply to S with 'no password file'. Otherwise, do the folloing: If the password data record $(\mathsf{file}, P_i, P_j, pw)$ is marked uncompromised, mark it compromised. If there is a tuple $(\mathsf{offline}, pw')$ stored with $pw' = pw$ then send pw to S, otherwise reply to S with 'password file stolen'.

Upon receiving a query $(\mathsf{OfflineTestPwd}, \boldsymbol{sid}, pw')$ **from Adversary** S**:**
If there is no password data record, or if there is a password data record $(\mathsf{file}, P_i, P_j, pw)$ that is marked uncompromised, then store $(\mathsf{offline}, pw')$. Otherwise do: if $pw = pw'$, send pw back to S. If $pw \neq pw'$, reply with 'wrong guess'.

Active Session Attacks

Upon receiving a query $(\mathsf{TestPwd}, \boldsymbol{sid}, \mathsf{ssid}, P_i, pw')$ **from the adversary** S**:**
If there is a session record of the form $(\mathsf{role}, \mathsf{ssid}, P_i, P_j, pw)$ which is fresh, then do: If $pw = pw'$, mark the record compromised and reply to S with "correct guess". If $pw \neq pw'$, mark the record interrupted and reply with "wrong guess".

Upon receiving a query $(\mathsf{Impersonate}, \boldsymbol{sid}, \mathsf{ssid})$
If there is a session record of the form $(\mathsf{Clt}, \mathsf{ssid}, P_i, P_j, pw)$ which is fresh, then do: then if there is a password data record file $(\mathsf{file}, P_i, P_j, pw)$ that is marked compromised, mark the session record compromised and reply to S with 'correct guess', else mark the session record interrupted and reply with wrong guess.

Key Generation and Authentication

Upon receiving a query $(\mathsf{NewKey}, \boldsymbol{sid}, \mathsf{ssid}, P_i, sk)$ **from** S**, where** $|sk| = k$**:**
If there is a session record of the form $(\mathsf{role}, \mathsf{ssid}, P_i, P_j, pw)$ that is not marked completed,

- If this record is compromised, or either P_i or P_j is corrupted, then output (sid, ssid, sk) to player P_i.
- If this record is fresh, and there is a session record $(\mathsf{role}, \mathsf{ssid}, P_j, P_i, pw')$ with $pw' = pw$, and a key sk' was sent to P_j, and $(\mathsf{role}, \mathsf{ssid}, P_j, P_i, pw)$ was fresh at the time, then output $(sid, \mathsf{ssid}, sk')$ to P_i.
- In any other case, pick a new random key sk' of length k and send $(sid, \mathsf{ssid}, sk')$ to P_i.

Either way, mark the record (P_i, P_j, pw) as completed.

Upon receiving $(\mathsf{Corrupt}, \boldsymbol{sid}, P)$ **from** S**:** if there is a $(\mathsf{Clt}, sid, P, P', pw)$ recorded, return pw to S, and mark P_i corrupted. If there is a $(\mathsf{Svr}, sid, P, P', pw)$ recorded, then mark P corrupted and (internally) call $(\mathsf{StealPwdFile}, sid)$.

Fig. 5. The password-based key-exchange functionality $\mathcal{F}_{\mathsf{apwKE}}$

Responding to random oracle queries. Let the input be m. If there is a record of the form (m, r), that is, m was queried before and was responded with r, then just return r.

Otherwise, if m is of the form (sid, P_i, P_j, x), for some x and the password file has been stolen then call $\mathsf{OfflineTestPwd}$ with x. If the test succeeds then return phash, which must already have been set (see Stealing Password File below), and record (m, phash).

In all other cases, generate $r \leftarrow \mathbb{Z}_q$, record (m, r) and return r.

Setting the CRS. The simulator $\mathcal{S}$ picks the CRS just as in the real world, except the QA-NIZK CRS-es are generated using the crs-simulators, which also generate simulator trapdoors $\mathsf{trap_C}, \mathsf{trap_S}$. It retains $a_1, a_2, \mathsf{trap_C}, \mathsf{trap_S}, \mathsf{hk_C}, \mathsf{hk_S}$ as trapdoors.

New Client Session: Sending a message to $\mathcal{Z}$. On message ($\mathsf{CltSession}$, sid, ssid, P_i, P_j) from $\mathcal{F}_{\mathsf{apwKE}}$, $\mathcal{S}$ starts simulating a new instance of the protocol for client P_i, server P_j, session identifier ssid, and CRS set as above. We will denote this instance by (P_i, ssid) and call it a *client instance*.

To simulate this instance, $\mathcal{S}$ chooses r_1, r_1', r_1'' at random, and sets $R_1 = \mathbf{g}^{r_1}$, $S_1 = \mathbf{g}_1^{r_1'}$ and $T_1 = \mathbf{g}^{r_1''}$ (note the use of arbitray constant μ instead of phash). Next, $\mathcal{S}$ generates $(\text{HK}_1, \text{HP}_1) \leftarrow \mathsf{ver.hkgen}(\text{CRS}_\mathsf{C})$ and sets $i_1 = \mathcal{H}(sid, \mathsf{ssid}, P_i, P_j, R_1, S_1, \text{HP}_1)$. It retains (i_1, HK_1). It then hands $(R_1, S_1, T_1, \text{HP}_1)$ to $\mathcal{Z}$ on behalf of this instance.

New Server Session: Sending a message to $\mathcal{Z}$. On message ($\mathsf{SvrSession}$, sid, ssid, P_j, P_i) from $\mathcal{F}_{\mathsf{apwKE}}$, $\mathcal{S}$ starts simulating a new instance of the protocol for client P_i, server P_j, session identifier ssid, and CRS set as above. We will denote this instance by (P_j, ssid) and call it a *server instance*.

To simulate this instance, $\mathcal{S}$ chooses r_2, r_2', r_2'' at random, and sets $R_2 = \mathbf{g}^{r_2}$, $S_2 = \mathbf{g}^{r_2'}$ and $T_1 = \mathbf{g}^{r_2''}$ (note the use of arbitrary constant μ instead of phash). Next, $\mathcal{S}$ generates $(\text{HK}_2, \text{HP}_2) \leftarrow \mathsf{ver.hkgen}(\text{CRS}_\mathsf{S})$ and sets $i_2 = \mathcal{H}(sid, \mathsf{ssid}, P_j, P_i, R_2, S_2, \text{HP}_2)$. It retains (i_2, HK_2). It then hands $(R_2, S_2, T_2, \text{HP}_2)$ to $\mathcal{Z}$ on behalf of this instance.

On Receiving a Message from $\mathcal{Z}$. On receiving a message $R_2', S_2', T_2', \hat{\rho}_2'$ from $\mathcal{Z}$ intended for a **client instance** (P, ssid), the simulator S does the following:

1. If any of the the real world protocol checks, namely group membership and non-triviality fail it goes to the step "Other Cases" below.
2. If the message received from $\mathcal{Z}$ is same as message sent by $\mathcal{S}$ on behalf of peer P' in session ssid, then $\mathcal{S}$ just issues a NewKey call for P.
3. ("Client Only Step"): If $\mathsf{StealPwdFile}$ has already taken place then do the following: If $S_2' = R_2'^{a_2} \mathbf{b}_\mathsf{S}^{\text{phash}}$, then $\mathcal{S}$ calls $\mathcal{F}_{\mathsf{apwKE}}$ with ($\mathsf{Impersonate}$, P, sid, ssid) and skips to the "Key Setting" step below, and otherwise go to the step "Other Cases".
4. It searches its random oracle query response pairs $\{(m_k, h_k)\}_k$ and checks whether for some $k = x$ we have $S_2' = R_2'^{a_2} \mathbf{b}_\mathsf{S}^{h_x}$ and m_x is of the form $(sid, P_i, P_j, \text{pwd}')$. If so, then $\mathcal{S}$ calls $\mathcal{F}_{\mathsf{apwKE}}$ with ($\mathsf{TestPwd}$, ssid, P, pwd$'$) else it goes to the step "Other Cases" below. If the test passes, it sets phash =

h_x and goes to the "Key Setting" step below, else it goes to the step "Other Cases" below.

5. ("Key Setting Step"): Compute $i'_2 = \mathcal{H}(sid, \mathsf{ssid}, P_j, P_i, R'_2, S'_2, \hat{\rho}'_2)$.
If $T'_2 \neq \mathsf{sphf_S.privH}(\mathsf{hk}, \langle R'_2, S'_2/\mathbf{b}_\mathsf{S}^{\text{phash}} \rangle, i'_2)$ then goto the step "Other Cases". Else, compute $W_1 = \mathsf{sim}(\text{CRS}_\mathsf{C}, \mathsf{trap_C}, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}, T_1, i_1 \rangle)$. Issue a NewKey call to $\widehat{\mathcal{F}}_{\text{PAKE}}$ with key

$$\mathsf{ver_S.privH}(\text{HK}_1, \langle R'_2, S'_2/\mathbf{b}_\mathsf{S}^{\text{phash}}, T'_2, i'_2 \rangle) \cdot \mathsf{ver_C.pubH}(\text{HP}'_2, W_1)$$

6. ("Other Cases"): $\mathcal{S}$ issues a TestPwd call to $\widehat{\mathcal{F}}_{\text{PAKE}}$ with the dummy password μ, followed by a NewKey call with a random session key, which leads to the functionality issuing a random and independent session key to the party P.

On receiving a message $R'_1, S'_1, T'_1, \text{HP}'_1$ from $\mathcal{Z}$ intended for a **server instance** (P, ssid), the response of the simulator $\mathcal{S}$ is symmetric to the response described above for client instances, except the above step "Client Only Step" is skipped.

Stealing Password File. If there was a successful online TestPwd call by the simulator, before this StealPwdFile call, the corresponding random oracle response h_k was already assigned to the variable phash. Otherwise, the simulator runs through the set of random oracle query response set of the adversary $\{(m_k, h_k)\}_k$, which were not used for an online TestPwd call. For all the m_k's of the form $(sid, P_i, P_j, \text{pwd}')$, it calls $(\mathsf{OfflineTestPwd}, sid, \text{pwd}')$. Next, $\mathcal{S}$ calls StealPwdFile. If StealPwdFile returns pwd then it must equal pwd$'$ in some m_k. Assign to the variable phash the value h_k from the earlier recorded random oracle response to m_k. Otherwise, phash is assigned a fresh random value. The Server Persistent State $\mathbf{b}_\mathsf{S}^{\text{phash}}$ is computed accordingly and given to the adversary.

Client Corruption. On receiving a Corrupt call from $\mathcal{Z}$ for client instance P_i in session ssid, the simulator $\mathcal{S}$ calls the Corrupt routine of $\mathcal{F}_{\mathsf{apwKE}}$ to obtain pwd. If $\mathcal{S}$ had already output a message to $\mathcal{Z}$, and not output sk_1 it computes

$$W_1 = \mathsf{sim_C}(\text{CRS}_\mathsf{C}, \mathsf{trap_C}, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}, T_1, i_1 \rangle).$$

and outputs this W_1 along with pwd, and HK_1 as internal state of P_i. Note that this computation of W_1 is identical to the computation of W_1 in the computation of sk_1 (which is really output to $\mathcal{Z}$ only when pwd$'$ = pwd).

Without loss of generality, we can assume that in the real-world if the Adversary (or Environment $\mathcal{Z}$) corrupts an instance before the session key is output then the instance does not output any session key. This is so because the Adversary (or $\mathcal{Z}$) either sets the key for that session or can compute it from the internal state it broke into.

Server Corruption. On receiving a Corrupt call from $\mathcal{Z}$ for server instance P_j in session ssid, the simulator $\mathcal{S}$ first performs the steps in the section on Stealing Password File above. In particular this sets the value of phash. It then calls the

$\mathsf{Corrupt}$ routine of $\mathcal{F}_{\mathsf{apwKE}}$. If $\mathcal{S}$ had already output a message to $\mathcal{Z}$, and not output sk_1 it computes

$$W_2 = \mathsf{sim}_\mathsf{S}(\textsc{crs}_\mathsf{S}, \mathsf{trap}_\mathsf{S}, \langle R_2, S_2/\mathbf{b}_\mathsf{S}^{\mathrm{phash}}, T_2, i_2\rangle).$$

and outputs this W_2 along with $\textsc{hk}_2$ as internal state of P_j. Note that pwd is not given out.

Complexity of the simulator. Observe that on stealing the password file, the function OfflineTestPwd is only called once for each random oracle input, which was not already tested by calling TestPwd. Hence the number of unique password arguments passed to TestPwd and OfflineTestPwd of $\mathcal{F}_{\mathsf{apwKE}}$ combined in the ideal world is at most the number of random oracle calls in the hybrid model.

Time complexity-wise, most of the simulator steps are $\log q$-time, where q is the security parameter. Due to Step 4 of the simulator code, where for each of the m sessions, in the worst case, it might go through all the n random oracle calls, there is an additive component of $m * n * \log q$ time. So the simulator runs in $O(mn \log q)$-time.

D.5 Proof of Indistinguishability - Series of Experiments.

We now describe a series of experiments between a probabilistic polynomial time challenger $\mathcal{C}$ and the environment $\mathcal{Z}$, starting with Expt_0 which we describe next. We will show that the view of $\mathcal{Z}$ in Expt_0 is same as its view in UC-APAKE ideal-world setting with $\mathcal{Z}$ interacting with $\mathcal{F}_{\mathsf{apwKE}}$ and the UC-APAKE simulator $\mathcal{S}$ described above in Appendix D.4. We end with an experiment which is identical to the real world execution of the protocol in Figure 4. We will show that the environment has negligible advantage in distinguishing between these series of experiments, leading to a proof of realization of $\mathcal{F}_{\mathsf{apwKE}}$ by the protocol Π.

Here is the complete code in Expt_0 (stated as it's overall experiment with $\mathcal{Z}$):

1. Responding to a random oracle query on input m: If there is a record of the form (m, r), then just return r. Otherwise, generate $r \leftarrow \mathbb{Z}_q$, record (m, r) and return r.
2. The challenger $\mathcal{C}$ picks the CRS just as in the real world, except the QA-NIZK CRS-es are generated using the crs-simulators, which also generate simulator trapdoors $\mathsf{trap}_\mathsf{C}, \mathsf{trap}_\mathsf{S}$. It retains $a_1, a_2, \mathsf{trap}_\mathsf{C}, \mathsf{trap}_\mathsf{S}, \mathsf{hk}_\mathsf{C}, \mathsf{hk}_\mathsf{S}$ as trapdoors.
   Next, (on StorePwdFile) the challenger calls the random oracle with query $(sid, P_i, P_j, \mathrm{pwd})$. It sets phash equal to the random oracle response and sets the server persistent state as $\mathbf{b}_\mathsf{S}^{\mathrm{phash}}$.
   Define PHASHISSET to be true after either StealPwdFile has been called or the random oracle has been called with $(sid, P_i, P_j, \mathrm{pwd})$ by the adversary, and false before.
   Define PWDCALLED to be true after the random oracle has been called with $(sid, P_i, P_j, \mathrm{pwd})$ by the adversary, and false before.

3. On receiving ($\mathsf{CltSession}$, sid, ssid, P_i, P_j) from $\mathcal{Z}$, $\mathcal{C}$ generates $(\text{HK}_1, \text{HP}_1) \leftarrow \mathsf{ver_S.hkgen}(\text{CRS}_\mathsf{S})$. Next, $\mathcal{C}$ chooses r_1, r_1', r_1'' at random, and sets $R_1 = \mathbf{g}^{r_1}$, $S_1 = \mathbf{g}^{r_1'}$ and $T_1 = \mathbf{g}^{r_1''}$. It then hands $(R_1, S_1, T_1, \text{HP}_1)$ to $\mathcal{Z}$ on behalf of this instance.
4. On receiving $(R_2', S_2', T_2', \text{HP}_2')$ from $\mathcal{Z}$, intended for client session (P_i, ssid) (and assuming no corruption of this instance):
   (a) If the received elements are either not in their respective groups, or are trivially 1, output $\text{sk}_1 \leftarrow \mathbb{G}_T$.
   (b) If the message received is identical to message sent by $\mathcal{C}$ in the same session (i.e. same ssid) on behalf of the peer, then output $\text{sk}_1 \leftarrow \mathbb{G}_T$ (unless the simulation of peer also received a legitimate message and its key has already been set, in which case the same key is used to output sk_1 here).
   (c) If PhashIsSet is false, then output $\text{sk}_1 \leftarrow \mathbb{G}_T$.
   (d) Compute: $i_2' = \mathcal{H}(sid, \mathsf{ssid}, P_j, P_i, R_2', S_2', \text{HP}_2')$. If $S_2' = R_2'^{a_2}\mathbf{b}_\mathsf{S}^{\text{phash}}$ and $T_2' = \mathsf{sphf_S.privH}(\mathsf{hk}, \langle R_2', S_2'/\mathbf{b}_\mathsf{S}^{\text{phash}}\rangle, i_2')$, compute:

$$W_1 = \mathsf{sim_C}(\text{CRS}_\mathsf{C}, \mathsf{trap_C}, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}, T_1, i_1\rangle).$$

   Output:

$$\text{sk}_1 = \mathsf{ver_S.privH}(\text{HK}_1, \langle R_2', S_2'/\mathbf{b}_\mathsf{S}^{\text{phash}}, T_2', i_2'\rangle) \cdot \mathsf{ver_C.pubH}(\text{HP}_2', W_1)$$

   (e) If the above check failed then output $\text{sk}_1 \leftarrow \mathbb{G}_T$.
5. On a $\mathsf{Corrupt}$ call for client P_i, output pwd. If Step 3 has already happened then also output HK_1 and $W_1 = \mathsf{sim_C}(\text{CRS}_\mathsf{C}, \mathsf{trap_C}, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}, T_1, i_1\rangle)$.
6. On receiving ($\mathsf{SrvSession}$, sid, ssid, P_j, P_i) from $\mathcal{Z}$, follow steps symmetric to Step 4, swapping subscripts and languages accordingly and replacing the condition PhashIsSet by PwdCalled in Step 4c.
7. On a $\mathsf{Corrupt}$ call for server P_j, if Step 3 has already happened then output HK_2, and $W_2 = \mathsf{sim_S}(\text{CRS}_\mathsf{S}, \mathsf{trap_S}, \langle R_2, S_2/\mathbf{b}_\mathsf{S}^{\text{phash}}, T_2, i_2\rangle)$. Finally, execute a $\mathsf{StealPwdFile}$ call, as described below.
8. On a $\mathsf{StealPwdFile}$ call, return $\mathbf{b}_\mathsf{S}^{\text{phash}}$ as the Server Persistent State to the adversary.

All outputs of sk_1 are also accompanied with sid, ssid (but are not mentioned above for ease of exposition).

Note that each instance has two asynchronous phases: a phase in which $\mathcal{C}$ outputs $R_1, S_1, ...$ to $\mathcal{Z}$, and a phase where it receives a message from $\mathcal{Z}$. However, $\mathcal{C}$ cannot output sk_1 until it has completed both phases. These orderings are dictated by $\mathcal{Z}$. We will consider two different kinds of temporal orderings. A temporal ordering of different instances based on the order in which $\mathcal{C}$ outputs sk_1 in an instance will be called **temporal ordering by key output**. A temporal ordering of different instances based on the order in which $\mathcal{C}$ outputs its first message (i.e. $R_1, S_1, ...$) will be called **temporal ordering by message output**. It is easy to see that $\mathcal{C}$ can dynamically compute both these orderings by maintaining a counter (for each ordering).

We now claim that the view of $\mathcal{Z}$ in Expt_0 is statistically indistinguishable from its view in its combined interaction with $\mathcal{F}_{\mathsf{apwKE}}$ and $\mathcal{S}$. The CRS is set identically by both $\mathcal{C}$ and $\mathcal{S}$. While $\mathcal{C}$ has access to pwd from the outset and sets up the random oracle output phash corresponding to $(sid, P_i, P_j, \mathsf{ssid})$ at the beginning, $\mathcal{S}$ doesn't have access to pwd at the beginning and hence defers this step till the point where either (1) a correct online guess has been made, (2) the password file was stolen and a correct offline guess was made, (3) the client was corrupted. In all these three cases the simulator gets to know pwd and has the chance to set phash. At the point when password file is stolen, the correct pwd may not have been guessed, but phash has to be set in order to output the server persistent state. In that case $\mathcal{S}$ generates a random phash, remembers it and assigns it to the correct input when the actual password is queried. At all points, although their algorithms differ, we can see that $\mathcal{C}$ and $\mathcal{S}$ respond to random oracle queries identically.

Both $\mathcal{C}$ and $\mathcal{S}$ generate the client and server flows identically. In particular, observe that the condition PhashIsSet exactly captures the state of $\mathcal{S}$ for a client session where it knows phash and can compute the relevant elements and keys. $\mathcal{C}$ uses the condition PhashIsSet to do the same computations. Similarly for the server sessions with the condition PwdCalled. The stronger condition for the server reflects the absence of the "Client Only Step" in the server sessions simulation. In the steps where a party receives a message from the adversary, both $\mathcal{C}$ and $\mathcal{S}$ end up computing keys identically. While $\mathcal{C}$ directly checks by exponentiation with phash in the case that pwd was guessed correctly, $\mathcal{S}$ goes through the list of random oracle calls to see which response was used for exponentiation as it may not know pwd or phash at this point.

The series of experiments is summarized in the following figures: Figure 6 depicts the algorithms that do not change across the games and Figure 7 shows how steps change with games.

Expt_1 : In this experiment, Step 56 (in Figure 7) is removed from both client and server instances.

For client instances, observe that if the condition PwdCalled does not hold, then phash remains information theoretically unknown to the adversary. Hence the simulator code has statistically negligible chance to reach Step 58.

For server instances, it remains to be proven that even if the adversary steals $\mathbf{b}_{\mathsf{S}}^{\text{phash}}$, there is negligible chance of it passing the condition $S'_1 = R'^{a_1}_1 \mathbf{b}_{\mathsf{C}}^{\text{phash}}$, unless it queries the random oracle with the correct password.

This can be proved by employing DDH on $(\mathbf{b}_{\mathsf{S}}, \mathbf{b}_{\mathsf{C}}, \mathbf{b}_{\mathsf{S}}^{\text{phash}}, \mathbf{b}_{\mathsf{C}}^{\text{phash}})$. Observe that if the random oracle is not called on the correct password, then the whole experiment can be simulated without phash in the clear and just using $(\mathbf{b}_{\mathsf{S}}, \mathbf{b}_{\mathsf{C}}, \mathbf{b}_{\mathsf{S}}^{\text{phash}}, \mathbf{b}_{\mathsf{C}}^{\text{phash}})$. In particular the condition $S'_1 \stackrel{?}{=} R'^{a_1}_1 \mathbf{b}_{\mathsf{C}}^{\text{phash}}$ can be switched by DDH to $S'_1 \stackrel{?}{=} R'^{a_1}_1 \mathbf{b}_{\mathsf{C}}^{z}$, where z is independently random from phash. At this point, we see again that the adversary has statistically negligible chance of making it to Step 58.

$\mathcal{RO}(m)$:

If there is a record of the form (m, r) : (23)

If $m \stackrel{?}{=} (sid, P_i, P_j, \text{pwd})$ then : (24)

PwdCalled := *true*, PHashIsSet := *true* (25)

Return r. (26)

Else (27)

Sample $r \leftarrow \mathbb{Z}_q$, record (m, r) and return r. (28)

StorePwdFile :

phash := $\mathcal{RO}(sid, P_i, P_j, \text{pwd})$. (29)

Server Persistent State := $\mathbf{b}_{\mathsf{S}}^{\text{phash}}$. (30)

StealPwdFile :

PHashIsSet := *true* (31)

Return $\mathbf{b}_{\mathsf{S}}^{\text{phash}}$. (32)

Corrupt(client P_i) :

If Send has already happened then (33)

Output $(\text{pwd}, \text{HK}_1, W_1)$. (34)

Else output pwd. (35)

Corrupt(server P_j) :

Call StealPwdFile. (36)

If Send has already happened then (37)

Output (HK_2, W_2). (38)

Fig. 6. Invariants in the UC-APAKE games.

Setup :

Game 0-3 : $(\textsc{crs}_{\mathsf{C}}, \mathsf{trap}_{\mathsf{C}}) \leftarrow \mathsf{crssim}_{\mathsf{C}}(),\ (\textsc{crs}_{\mathsf{S}}, \mathsf{trap}_{\mathsf{S}}) \leftarrow \mathsf{crssim}_{\mathsf{S}}()$ (39)

Game 4-8 : $\textsc{crs}_{\mathsf{C}} \leftarrow \mathsf{crsgen}_{\mathsf{C}}(),\ \textsc{crs}_{\mathsf{S}} \leftarrow \mathsf{crsgen}_{\mathsf{S}}()$ (40)

PwdCalled $:= false$, PhashIsSet $:= false$ (41)

Call StorePwdFile (42)

Receive(CltSession, sid, ssid, P_i, P_j) :

Sample $(\textsc{hk}_1, \textsc{hp}_1) \leftarrow \mathsf{vers}_{\mathsf{S}}.\mathsf{hkgen}(\textsc{crs}_{\mathsf{S}})$ (43)

Sample $(r_1, r_1', r_1'') \leftarrow \mathbb{Z}_q^3$ (44)

Game 0-2 : Set $(R_1, S_1, T_1) := (\mathbf{g}^{r_1},\ \mathbf{g}^{r_1'},\ \mathbf{g}^{r_1''})$ (45)

Game 3-8 : Set $(R_1, S_1) := (\mathbf{g}^{r_1},\ \mathbf{a}_1^{r_1}\mathbf{b}_{\mathsf{C}}^{\text{phash}})$ (46)

Compute $i_1 := \mathcal{H}(sid, \mathsf{ssid}, P_i, P_j, R_1, S_1, \textsc{hp}_1)$ (47)

Game 3-8 : Set $T_1 := \mathsf{sphf}_{\mathsf{C}}.\mathsf{pubH}(\mathsf{hp}_{\mathsf{C}}, \langle R_1, S_1, \mathbf{b}_{\mathsf{S}}^{\text{phash}}\rangle, i_1; r_1, \text{phash})$ (48)

Game 0-3 : Compute $W_1 = \mathsf{sim}_{\mathsf{C}}(\textsc{crs}_{\mathsf{C}}, \mathsf{trap}_{\mathsf{C}}, \langle R_1, S_1, \mathbf{b}_{\mathsf{S}}^{\text{phash}}, T_1, i_1\rangle)$. (49)

Game 4-8 : Compute $W_1 = \mathsf{prover}_{\mathsf{C}}(\textsc{crs}_{\mathsf{C}}, \langle R_1, S_1, \mathbf{b}_{\mathsf{S}}^{\text{phash}}, T_1, i_1\rangle; r_1, \text{phash})$. (50)

Send $(R_1, S_1, T_1, \textsc{hp}_1)$ to $\mathcal{Z}$. (51)

Receive$(R_2', S_2', T_2', \textsc{hp}_2')$ from $\mathcal{Z}$, intended for client session (P_i, ssid):

If the received elements are invalid, output $\text{sk}_1 \leftarrow \mathbb{G}_T$. (52)

Game 0-7 : If peer sent $(R_2', S_2', T_2', \textsc{hp}_2')$ then: (53)

Game 0-7 : If peer received $(R_1, S_1, T_1, \textsc{hp}_1)$ and outputted sk_1, then output sk_1. (54)

Game 0-4 : Else output $\text{sk}_1 \leftarrow \mathbb{G}_T$. (55)

Game 0 : If PhashIsSet is false, then output $\text{sk}_1 \leftarrow \mathbb{G}_T$. (56)

Compute: $i_2' = \mathcal{H}(sid, \mathsf{ssid}, P_j, P_i, R_2', S_2', \textsc{hp}_2')$. (57)

Game 0-1, 6 : If $S_2' = R_2'^{a_2}\mathbf{b}_{\mathsf{S}}^{\text{phash}}$ and $T_2' = \mathsf{sphf}_{\mathsf{S}}.\mathsf{privH}(\mathsf{hk}, \langle R_2', S_2'/\mathbf{b}_{\mathsf{S}}^{\text{phash}}\rangle, i_2')$ then: (58)

Game 2-5 : If $T_2' = \mathsf{sphf}_{\mathsf{S}}.\mathsf{privH}(\mathsf{hk}, \langle R_2', S_2'/\mathbf{b}_{\mathsf{S}}^{\text{phash}}\rangle, i_2')$ then: (59)

Output $\text{sk}_1 = \mathsf{vers}_{\mathsf{S}}.\mathsf{privH}(\textsc{hk}_1, \langle R_2', S_2'/\mathbf{b}_{\mathsf{S}}^{\text{phash}}, T_2', i_2'\rangle) \cdot \mathsf{ver}_{\mathsf{C}}.\mathsf{pubH}(\textsc{hp}_2', W_1)$ (60)

Game 0-6 : Else output $\text{sk}_1 \leftarrow \mathbb{G}_T$. (61)

For the server side protocol :

Swap subscripts and languages. (62)

Swap PhashIsSet by PwdCalled in Step 56. (63)

Fig. 7. Games for UC-APAKE proof

Once the Step 56 is removed, we switch back to the real DDH tuple, thus reaching $\textsf{\textbf{Expt}}_1$.

$\textsf{\textbf{Expt}}_2$: In this experiment, Step 58 is switched to Step 59, i.e., the disjunct $(S'_2 = R'^{a_2}_2 \mathbf{b}_\mathsf{S}^{\text{phash}})$ is dropped.

First note that T_1 is being computed randomly. The experiment $\textsf{\textbf{Expt}}_3$ is then statistically indistinguishable from $\textsf{\textbf{Expt}}_2$ by smoothness of sphf_S (note that it can be shown that the polynomial number of extra bits of information leaked by the conditions $T'_2 = \mathsf{sphf}_\mathsf{S}.\mathsf{privH}(\mathsf{hk}_\mathsf{S}, \langle R'_2, S'_2/\mathbf{b}_\mathsf{S}^{\text{phash}}\rangle, i'_2)$ themselves have a negligible effect on the smoothness of sphf_S – this argument is employed in the Cramer-Shoup CCA2-encryption scheme [CS02]).

Correcting Message Outputs to use pwd

$\textsf{\textbf{Expt}}_3$: In each instance, S_1 is computed as follows: $\mathbf{a}_1^{r_1}\mathbf{b}_\mathsf{C}^{\text{phash}}$. Further, T_1 is computed as follows: $T_1 = \mathsf{sphf}_\mathsf{C}.\mathsf{pubH}(\mathsf{hp}_\mathsf{C}, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}\rangle, i_1; r_1, \text{phash})$. Symmetrically, for the server instances.

To show that $\textsf{\textbf{Expt}}_2$ is computationally indistinguishable from $\textsf{\textbf{Expt}}_3$, we define several hybrid experiments $\textsf{\textbf{Expt}}_{2,i}$ inductively. Experiment $\textsf{\textbf{Expt}}_{2,0}$ is identical to $\textsf{\textbf{Expt}}_2$. If there are a total of N instances, $\textsf{\textbf{Expt}}_{2,N}$ will be identical to $\textsf{\textbf{Expt}}_3$. Experiment $\textsf{\textbf{Expt}}_{2,i+1}$ differs from experiment $\textsf{\textbf{Expt}}_{2,i}$ in only (temporally ordered by message output) the $(i+1)$-th instance. While in $\textsf{\textbf{Expt}}_{2,i}$, the $(i+1)$-th instance is simulated by $\mathcal{C}$ as in $\textsf{\textbf{Expt}}_2$, in $\textsf{\textbf{Expt}}_{2,i+1}$ this instance is simulated as in $\textsf{\textbf{Expt}}_3$.

Lemma 1. *For all $i : 0 \le i \le N$, the view of $\mathcal{Z}$ in experiment $\textsf{\textbf{Expt}}_{2,i+1}$ is computationally indistinguishable from the view of $\mathcal{Z}$ in $\textsf{\textbf{Expt}}_{2,i}$.*

The lemma is proved in Appendix D.6.

$\textsf{\textbf{Expt}}_4$: In this experiment, the CRS is generated using crsgen instead of the crs-simulator, and W_1 is computed everywhere by prover of the QA-NIZK instead of the proof simulator.

Indistinguishability from the previous experiment follows by zero-knowledge property of the QA-NIZK, noting that all proofs being generated are on language memebers.

Handling Legitimate Messages

$\textsf{\textbf{Expt}}_5$: In this experiment the Step 55 is dropped. So in effect, if the peer's correct message is received and the peer has not outputted a key, then we directly go to Step 57, instead of outputting a random sk_1.

To show that $\textsf{Expt}_5$ is indistinguishable from $\textsf{Expt}_4$ we need to go through several hybrid experiments. In each subsequent hybrid experiment one more instance is modified, and the order in which these instances are handled is determined by temporal order of key output. In the hybrid experiment $\textsf{Expt}_{4,i}$ ($N \geq i \geq 1$), the Step 55 in the i-th temporally ordered instance is dropped as required in the $\textsf{Expt}_5$ description above. Experiment $\textsf{Expt}_{4,0}$ is same as experiment $\textsf{Expt}_4$, and experiment $\textsf{Expt}_{4,N}$ is same as experiment $\textsf{Expt}_5$.

Lemma 2. *For all $i \in [1..N]$, experiment $\textsf{Expt}_{4,i}$ is computationally indistinguishable from $\textsf{Expt}_{4,i-1}$.*

The lemma is proved in Appendix D.6.

Handling Adversarial Messages

$\textsf{Expt}_6$: In this experiment, Step 58 is switched to Step 59. In other words, the disjunct $(S'_2 = R'^{a_2}_2 \mathbf{b}_{\mathsf{S}}^{\text{phash}})$ is introduced.

Indistinguishability follows by the same argument as employed in going from experiment $\textsf{Expt}_1$ to $\textsf{Expt}_2$.

$\textsf{Expt}_7$: In this experiment Step 58 is dropped altogether.

We first show that if the condition:

$$(S'_2 \neq R'^{a_2}_2 \mathbf{b}_{\mathsf{S}}^{\text{phash}}) \textbf{ or } T'_2 \neq \mathsf{sphf_S.privH}(\mathsf{hk_S}, \langle R'_2, S'_2/\mathbf{b}_{\mathsf{S}}^{\text{phash}} \rangle, i'_2)$$

holds, then $(R'_2, S'_2/\mathbf{b}_{\mathsf{S}}^{\text{phash}}, T'_2, i'_2)$ is not in language L^+_{S} (for which the QA-NIZK is defined). Clearly, if the first disjunct does not hold then the tuple is not in the language. So, suppose $(S'_2 = R'^{a_2}_2 \mathbf{b}_{\mathsf{S}}^{\text{phash}})$, with witness r_2 for R'_2. Then, by correctness of the sphf,

$$\mathsf{sphf_S.privH}(\mathsf{hk_S}, \langle R'_2, S'_2/\mathbf{b}_{\mathsf{S}}^{\text{phash}} \rangle, i'_2) = \mathsf{sphf_S.pubH}(\mathsf{hp_S}, \langle R'_2, S'_2/\mathbf{b}^{\text{phash}} \rangle, i'_2; r_2).$$

Therefore, again, the tuple is not in the language.

Thus, $\mathsf{ver_S.privH}(\textsc{hk}_1, \langle R'_2, S'_2/\mathbf{b}_{\mathsf{S}}^{\text{phash}}, T'_2 \rangle, i'_2)$ is random, even when the Adversary is given $\textsc{hp}_1$, by smooth-soundness of the QA-NIZK.

$\textsf{Expt}_8$: In this experiment the Step 55 is dropped. In other words, the challenger code goes straight from Step 52 to Step 57.

Experiments $\textsf{Expt}_8$ and $\textsf{Expt}_7$ produce the same view for $\mathcal{Z}$, since if both peers (of a instance) received legitimate messages forwarded by $\mathcal{Z}$, then Step 58 computes the same instance key in both instances.

Finally, a simple examination shows that the view of $\mathcal{Z}$ in $\textsf{Expt}_8$ is identical to the real world protocol. That completes the proof of Theorem 6. □

D.6 Lemma proofs

Proof (of Lemma 1). We define several hybrid experiments. Experiment $\mathbf{G}_0$ is identical to $\mathsf{Expt}_{2,i}$. We describe the client sessions here - the server sessions are symmetrical.

In $\mathbf{G}_1$, in the $(i+1)$-th instance T_1 is computed differently:

$$T_1 = \mathsf{sphf_C.privH}(\mathsf{hk_C}, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\mathrm{phash}} \rangle, i_1) \tag{64}$$

This is statistically the same as all other T_1 are either randomly computed (in instances greater than $(i+1)$), or are computed using the public hash with hp (in instances less than $(i+1)$). Then the claim follows by smoothness of $\mathsf{sphf_C}$, and noting that $S_1 \neq R_1^{a_1}\mathbf{b}_\mathsf{C}^{\mathrm{phash}}$ in instance $(i+1)$ (by construction of $\mathsf{Expt}_{2,i}$).

In the next experiment $\mathbf{G}_2$, the challenger generates the S_1 in the $(i+1)$-th instance as follows: $S_1 = \mathbf{a}_1^{r_1}\mathbf{b}_\mathsf{C}^{\mathrm{phash}}$. That the view of $\mathcal{Z}$ in experiments $\mathbf{G}_1$ and $\mathbf{G}_2$ are computationally indistinguishable follows from the DDH assumption in group $\mathbb{G}_1$ (note a_1 is not being used by the challenger, now that Step 58 was switched to Step 59).

In the next experiment $\mathbf{G}_3$, change the computation of T_1 in session $(i+1)$ to use the public hash (of $\mathsf{sphf_C}$) and witness r_1. Since, now $(R_1, S_1, \mathbf{b}_\mathsf{S}^{\mathrm{phash}})$ is in the language L_C, indistinguishability from the previous experiment follows by correctness of $\mathsf{sphf_C}$. □

Proof (of Lemma 2). The lemma is proved using several hybrid experiments of its own. The experiment $\mathbf{H}_0$ is same as $\mathsf{Expt}_{4,i-1}$.

In experiment $\mathbf{H}_1$ the CRS is set as in the real world, except that the QA-NIZK CRS_C is set using the crs simulators $\mathsf{crssim_C}$ (the challenger retains the trapdoors $\mathsf{trap_C}$ output by the crs simulator). All proofs W_1 are still computed using $\mathsf{prover_C}$. Experiments $\mathbf{H}_0$ and $\mathbf{H}_1$ are indistinguishable as the QA-NIZK has the property that the simulated CRS and the real-world CRS are statistically identical.

In experiment $\mathbf{H}_2$, in instance i, the value W_1 (in Step 58 or corruption) is generated using the proof simulator using trapdoor trap. Indistinguishability follows by zero-knowledge property of the QA-NIZK as the proof being generated is on a language member.

In experiment $\mathbf{H}_3$, in instance i, the value T_1 is generated using the private hash key $\mathsf{hk_C}$, and the private hash function $\mathsf{sphf_C.privH}$ (thus eliminating the use of witness r_1). Experiments $\mathbf{H}_2$ and $\mathbf{H}_3$ are indistinguishable by the correctness of $\mathsf{sphf_C}$.

In experiment $\mathbf{H}_4$, in instance i, the values R_1, S_1 are generated as $R_1 = \mathbf{g}^{r_1}$, $S_1 = \mathbf{g}^{r_1'}$. where r_1, r_1' are random and independent. This follows by employing DDH on $\mathbf{g}, \mathbf{a}_1, R_1, S_1$.

In experiment $\mathbf{H}_5$, in peer of instance i, Step 59 is switched to Step 58. In other words, the condition $T_1' = \mathsf{sphf_C.privH}(\mathsf{hk_C}, \langle R_1', S_1', \mathbf{b}_\mathsf{S}^{\mathrm{phash}} \rangle, i_1')$ is replaced by $(S_1' = R_1'^{a_1}\mathbf{b}_\mathsf{C}^{\mathrm{phash}})$ **and** $T_1' = \mathsf{sphf_C.privH}(\mathsf{hk_C}, \langle R_1', S_1', \mathbf{b}_\mathsf{S}^{\mathrm{phash}} \rangle, i_1')$. Indistinguishability from experiment $\mathbf{H}_4$ follows by smooth$_2$ property of $\mathsf{sphf_C}$, noting that at most one bad $\mathsf{sphf_C.privH}$ is being output to the Adversary (namely T_1 in instance i).

In experiment $\mathbf{H}_6$, in instance i, change Step 55 as follows: If the message received is identical to message sent by $\mathcal{C}$ in the same instance (i.e. same SSID) on behalf of the peer,

- If simulation of peer also received a legitimate message and its key has already been set, then output that same key here. If peer is corrupted, output the key supplied by the Adversary.
- Else, compute $i'_2 = \mathcal{H}(sid, \mathsf{ssid},\ P_j, P_i,\ R'_2, S'_2, \text{HP}'_2)$, Output

$$\mathsf{ver_S.privH}(\text{HK}_1, \langle R'_2, S'_2/\mathbf{b}_\mathsf{S}^{\text{phash}}, T'_2\rangle, i'_2) \cdot \mathsf{ver_C.privH}(\text{HK}_2, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}, T_1\rangle, i_1)$$

Here HK_2 is the HK output by $\mathsf{ver_S.hkgen}$ in the peer instance of instance i.

The experiments $\mathbf{H}_6$ and $\mathbf{H}_5$ are computationally indistinguishable by noting the following three facts:

1. In the peer of instance of instance i (which generated HK_2), in Step 58 the computation $\mathsf{ver_C.privH}(\text{HK}_2, \cdot)$ is on a language member, as this computation is only reached if the the incoming tuple is in the language.
2. Also, note that only one QA-NIZK proof is being simulated and that is in this same instance, but in a mutually exclusive step (Step 58 or corruption). Moreover, the CRS generated by the crs simulator is statistically identical to the CRS generated by $\mathsf{crsgen_C}$.
3. Then, $\mathsf{ver_C.privH}(\text{HK}_2, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}, T_1\rangle, i_1)$ is random even when the adversary is given HP_2 by smoothness of the QA-NIZK, since $S_1 \neq R_1^{a_1}\mathbf{b}_\mathsf{C}^{\text{phash}}$.

In experiment $\mathbf{H}_7$, in peer of instance i, Step 58 is switched to Step 59. In other words, the condition "$(S'_1 = R_1'^{a_1}\mathbf{b}_\mathsf{C}^{\text{phash}})$ **and** $T'_1 = \mathsf{sphf_C.privH}(\mathsf{hk_C}, \langle R'_1, S'_1, \mathbf{b}_\mathsf{C}^{\text{phash}}\rangle, i'_1)$" is replaced by "$T'_1 \neq \mathsf{sphf_C.privH}(\mathsf{hk_C}, \langle R'_1, S'_1, \mathbf{b}_\mathsf{C}^{\text{phash}}\rangle, i'_1)$". Indistinguishability from experiment $\mathbf{H}_6$ follows by smooth$_2$ property of the $\mathsf{sphf_C}$, noting that at most one bad $\mathsf{sphf_C.privH}$ is being output to the Adversary (namely T_1 in instance i).

In experiment $\mathbf{H}_8$, in instance i, R_1, S_1 are generated as $R_1 = \mathbf{g}^{r_1}$, $S_1 = \mathbf{a}_1^{r_1}\mathbf{b}_\mathsf{C}^{\text{phash}}$, by employing DDH.

In experiment $\mathbf{H}_9$, in instance i, T_1 is generated using the public hash key $\mathsf{hp_C}$, and witness r_1. Indistinguishability follows by correctness of the sphf.

In experiment $\mathbf{H}_{10}$, the QA-NIZK is generated using the real world CRS generator. Moreover, in instance i, in Step 58 and corruption step, W_1 is computed using the real world prover. Indistinguishability follows by zero-knowledge property of the QA-NIZK.

In experiment $\mathbf{H}_{11}$, in Step 55 the key is output as follows:

- Else, compute $i'_2 = \mathcal{H}(sid, \mathsf{ssid},\ P_j, P_i,\ R'_2, S'_2, \text{HP}'_2)$.
  Compute $W_1 = \mathsf{prover_C}(\text{CRS}_\mathsf{C}, \langle R_1, S_1, \mathbf{b}_\mathsf{S}^{\text{phash}}, T_1, i_1\rangle, r_1)$. Output

$$\mathsf{ver_S.privH}(\text{HK}_1, \langle R'_2, S'_2/\mathbf{b}_\mathsf{S}^{\text{phash}}, T'_2\rangle, \iota'_2) \cdot \mathsf{ver_C.pubH}(\text{HP}'_2, W_1)$$

Indistinguishability follows by noting that HP'_2 is exactly the HP_2 computed by the challenger in the peer instance. The claim then follows by completeness of the smooth QA-NIZK.

The induction step is complete now, as the above computation of the session key is same as in Step 58. □

E Single-Round UC Password-Based Key Exchange

The essential elements of the Universal Composability framework can be found in [Can01]. In the following, we adopt the definition for password-based key exchange (UC-PAKE) from Canetti et al [CHK+05].

E.1 UC-PAKE Definition

Functionality $\mathcal{F}_{\mathbf{pake}}$

The functionality $\mathcal{F}_{\text{PAKE}}$ is parameterized by a security parameter k. It interacts with an adversary S and a set of parties via the following queries:

Upon receiving a query ($\mathsf{NewSession}, sid, P_i, P_j, pw, role$) **from party** P_i**:**
Send ($\mathsf{NewSession}, sid, P_i, P_j, role$) to S. In addition, if this is the first $\mathsf{NewSession}$ query, or if this is the second $\mathsf{NewSession}$ query and there is a record (P_j, P_i, pw'), then record (P_i, P_j, pw) and mark this record fresh.
Upon receiving a query ($\mathsf{TestPwd}, sid, P_i, pw'$) **from the adversary** S**:**
If there is a record of the form (P_i, P_j, pw) which is fresh, then do: If $pw = pw'$, mark the record $\mathsf{compromised}$ and reply to S with "correct guess". If $pw \neq pw'$, mark the record $\mathsf{interrupted}$ and reply with "wrong guess".
Upon receiving a query ($\mathsf{NewKey}, sid, P_i, sk$) **from** S**, where** $|sk| = k$**:**
If there is a record of the form (P_i, P_j, pw), and this is the first NewKey query for P_i, then:

- If this record is $\mathsf{compromised}$, or either P_i or P_j is corrupted, then output (sid, sk) to player P_i.
- If this record is fresh, and there is a record (P_j, P_i, pw') with $pw' = pw$, and a key sk' was sent to P_j, and (P_j, P_i, pw) was fresh at the time, then output (sid, sk') to P_i.
- In any other case, pick a new random key sk' of length k and send (sid, sk') to P_i.

Either way, mark the record (P_i, P_j, pw) as completed.
Upon receiving ($\mathsf{Corrupt}, sid, P_i$) **from** S**:** if there is a (P_i, P_j, pw) recorded, return pw to S, and mark P_i corrupted.

Fig. 8. The password-based key-exchange functionality $\mathcal{F}_{\text{PAKE}}$

Just as in the normal key-exchange functionality, if both participating parties are not corrupted, then they receive the same uniformly distributed session key and the adversary learns nothing of the key except that it was generated. However, if one of the parties is corrupted, then the adversary determines the session key. This power to the adversary is *also* given in case it succeeds in guessing the parties' shared password. Participants also detect when the adversary makes an unsuccessful attempt. If the adversary makes a wrong password guess in a given session, then the session is marked $\mathsf{interrupted}$ and the parties are provided random and independent session keys. If however the adversary makes a successful guess, then the session is marked $\mathsf{compromised}$, and the adversary is allowed to set the session key. If a session remains marked fresh, meaning that it is neither interrupted nor compromised. uncorrupted parties conclude with both parties receiving the same, uniformly distributed session key.

The formal description of the UC-PAKE functionality $\mathcal{F}_{\text{PAKE}}$ is given in Figure 8 and the protocol of [JR15] is provided with smooth QA-NIZK abstractions in Figure 9. This protocol is also secure when different sessions use the same common reference string (CRS) To achieve this goal, the *Universal Composability with joint state* (JUC) formalism of Canetti and Rabin [CR03] is considered. This formalism provides a "wrapper layer" that deals with "joint state" among different copies of the protocol. In particular, defining a functionality $\mathcal{F}$ also implicitly defines the multi-session extension of $\mathcal{F}$ (denoted by $\hat{\mathcal{F}}$): $\hat{\mathcal{F}}$ runs multiple independent copies of $\mathcal{F}$, where the copies are distinguished via sub-session IDs ssid. The JUC theorem [CR03] asserts that for any protocol π that uses multiple independent copies of $\mathcal{F}$, composing π instead with a single copy of a protocol that realizes $\hat{\mathcal{F}}$, preserves the security of π.

Party P_i	Network
Generate $\mathbf{g}_1 \leftarrow \mathbb{G}_1, \mathbf{g}_2 \leftarrow \mathbb{G}_2$ and $\mathbf{a} = \mathbf{g}_1^a$ with $a \leftarrow \mathbb{Z}_q$ as DH parameters ρ. Let $\mathcal{H}$ be a CRHF, and sphf be a smooth$_2$ SPHF family for the DH family. $(\mathsf{hp}, \mathsf{hk}) \leftarrow \mathsf{sphf.hkgen}(\rho)$. Let $(\mathsf{pargen}, \mathsf{crsgen}, \mathsf{prover}, \mathsf{ver})$ be a Smooth QA-NIZK for language L, $L = \{R, S, T, l : \exists r, R = \mathbf{g}_1^r, S = \mathbf{a}^r, T = \mathsf{sphf.pubH}(\mathsf{hp}, \langle R, S\rangle, l; r)\}$. $\text{CRS} \leftarrow \mathsf{crsgen}(\rho)$. CRS $:= (\rho, \mathsf{hp}, \text{CRS}, \mathcal{H})$.	
Input $(\mathtt{NewSession}, sid, \mathsf{ssid}, P_i, P_j, \mathrm{pwd}, initiator/responder)$ Choose $r_1 \xleftarrow{\$} \mathbb{Z}_q$, $(\text{HK}_1, \text{HP}_1) \leftarrow \mathsf{ver.hkgen}(\text{CRS})$. Set $R_1 = \mathbf{g}_1^{r_1}$, $S_1 = \mathrm{pwd} \cdot \mathbf{a}^{r_1}$, $T_1 = \mathsf{sphf.pubH}(\mathsf{hp}, \langle R, S_1/\mathrm{pwd}\rangle, i_1; r_1)$, $W_1 = \mathsf{prover}(\text{CRS}, \langle R_1, S_2, T_1, i_1\rangle; r_1)$, where $i_1 = \mathcal{H}(sid, \mathsf{ssid}, P_i, P_j, R_1, S_1, \text{HP}_1)$. Erase r_1, send $(R_1, S_1, T_1, \text{HP}_1)$ and retain (W_1, HK_1).	$\xrightarrow{R_1, S_1, T_1, \text{HP}_1}$ P_j
Receive $R_2', S_2', T_2', \text{HP}_2'$. If any of $R_2', S_2', T_2', \text{HP}_2'$ is not in their respective group or is 1, set $\mathrm{sk}_1 \xleftarrow{\$} \mathbb{G}_T$, else compute $i_2' = \mathcal{H}(sid, \mathsf{ssid}, P_j, P_i, R_2', S_2', \text{HP}_2')$, Compute $\mathrm{sk}_1 = \mathsf{ver.privH}(\text{HK}_1, \langle R_2', S_2'/\mathrm{pwd}, T_2', i_2'\rangle) \cdot \mathsf{ver.pubH}(\text{HP}_2', W_1)$. Output $(sid, \mathsf{ssid}, \mathrm{sk}_1)$.	$\xleftarrow{R_2', S_2', T_2', \text{HP}_2'}$ P_j

Fig. 9. Single-round UC-PAKE protocol under SXDH assumption.