An Algebraic Framework for Diffie-Hellman Assumptions

Alex Escala[1], Gottfried Herold[*2], Eike Kiltz[†2], Carla Ràfols[2], and Jorge Villar[‡1]

[1]Universitat Politècnica de Catalunya, Spain, {alex.escala,jvillar}@ma4.upc.edu
[2]Horst-Görtz Institute for IT Security and Faculty of Mathematics, Ruhr-Universität Bochum, Germany, {gottfried.herold,eike.kiltz,carla.rafols}@rub.de

Abstract

We put forward a new algebraic framework to generalize and analyze Diffie-Hellman like Decisional Assumptions which allows us to argue about security and applications by considering only algebraic properties. Our $\mathcal{D}_{\ell,k}$-MDDH assumption states that it is hard to decide whether a vector in $\mathbb{G}^\ell$ is linearly dependent of the columns of some matrix in $\mathbb{G}^{\ell\times k}$ sampled according to distribution $\mathcal{D}_{\ell,k}$. It covers known assumptions such as DDH, 2-Lin (linear assumption), and k-Lin (the k-linear assumption). Using our algebraic viewpoint, we can relate the generic hardness of our assumptions in m-linear groups to the irreducibility of certain polynomials which describe the output of $\mathcal{D}_{\ell,k}$. We use the hardness results to find new distributions for which the $\mathcal{D}_{\ell,k}$-MDDH-Assumption holds generically in m-linear groups. In particular, our new assumptions 2-SCasc and 2-ILin are generically hard in bilinear groups and, compared to 2-Lin, have shorter description size, which is a relevant parameter for efficiency in many applications. These results support using our new assumptions as natural replacements for the 2-Lin Assumption which was already used in a large number of applications.

To illustrate the conceptual advantages of our algebraic framework, we construct several fundamental primitives based on any MDDH-Assumption. In particular, we can give many instantiations of a primitive in a compact way, including public-key encryption, hash-proof systems, pseudo-random functions, and Groth-Sahai NIZK and NIWI proofs. As an independent contribution we give more efficient NIZK and NIWI proofs for membership in a subgroup of $\mathbb{G}^\ell$. The results imply very significant efficiency improvements for a large number of schemes.

Keywords: Diffie-Hellman Assumption, Generic Hardness, Groth-Sahai proofs, Hash Proof Systems, Public-key Encryption.

[*]Funded by ERC Grant ERC307952 Fast and Sound Cryptography.
[†]Funded by a Sofja Kovalevskaja Award of the Alexander von Humboldt Foundation and the German Federal Ministry for Education and Research.
[‡]Funded by the Spanish research project MTM2013-41426-R.

Contents

1 Introduction **1**
- 1.1 The Matrix Diffie-Hellman Assumption . . . 1
- 1.2 Basic Applications . . . 3

2 Preliminaries **4**
- 2.1 Notation . . . 4
- 2.2 Representing Elements in Groups . . . 4
- 2.3 Standard Diffie-Hellman Assumptions . . . 5
- 2.4 Key Encapsulation Mechanisms . . . 5
- 2.5 Hash Proof Systems . . . 6
- 2.6 Pseudo-Random Functions . . . 6

3 Matrix DH assumptions **6**
- 3.1 Definition . . . 6
- 3.2 Basic Properties . . . 7
- 3.3 Generic Hardness of Matrix DH . . . 8
- 3.4 Examples of $\mathcal{D}_{\ell,k}$-MDDH . . . 9

4 Uniqueness of One-Parameter Matrix DH Problems **10**
- 4.1 Hardness . . . 11
- 4.2 Isomorphic Problems . . . 12

5 Basic Applications **13**
- 5.1 Public-Key Encryption . . . 13
- 5.2 Hash Proof Systems . . . 14
- 5.3 Pseudo-Random Functions . . . 14
- 5.4 Groth-Sahai Non-interactive Zero-Knowledge Proofs . . . 16

6 More Efficient Proofs for Some CRS Dependent Languages **19**
- 6.1 More Efficient Subgroup Membership Proofs . . . 20
- 6.2 Other CRS Dependent Languages . . . 23

A Proof of Theorem 7 **27**

B Proofs for the Generic Hardness results **28**
- B.1 Proof of Theorem 3 . . . 29
- B.2 Proof of Theorem 4 and Generalizations . . . 30

C Proof of Theorem 10 **32**

D Subgroup Membership Proofs for 2-Lin **33**

E Concrete Examples from the k-SCasc Assumption **34**
- E.1 Key Encapsulation . . . 34
- E.2 Pseudo-Random Function . . . 35

1 Introduction

Arguably, one of the most important cryptographic hardness assumptions is the Decisional Diffie-Hellman (DDH) Assumption. For a fixed additive group $\mathbb{G}$ of prime order q and a generator $\mathcal{P}$ of $\mathbb{G}$, we denote by $[a] := a\mathcal{P} \in \mathbb{G}$ the *implicit representation* of an element $a \in \mathbb{Z}_q$. The DDH Assumption states that $([a], [r], [ar]) \approx_c ([a], [r], [z]) \in \mathbb{G}^3$, where a, r, z are uniform elements in $\mathbb{Z}_q$ and $\approx_c$ denotes computationally indistinguishability of the two distributions. It has been used in numerous important applications such as secure encryption [12], key-exchange [20], hash-proof systems [13], pseudo-random functions [37], and many more.

Bilinear Groups and the Linear Assumption. Bilinear groups (i.e., groups $\mathbb{G}, \mathbb{G}_T$ of prime order q equipped with a bilinear map $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$) [4, 24] revolutionized cryptography in recent years and and are the basis for a large number of cryptographic protocols. However, relative to a (symmetric) bilinear map, the DDH Assumption is no longer true in the group $\mathbb{G}$. (This is since $e([a], [r]) = e([1], [ar])$ and hence $[ar]$ is not longer pseudorandom given $[a]$ and $[r]$.) The need for an "alternative" decisional assumption in $\mathbb{G}$ was quickly addressed with the Linear Assumption (2-Lin) introduced by Boneh, Boyen, and Shacham [3]. It states that $([a_1], [a_2], [a_1 r_1], [a_2 r_2], [r_1 + r_2]) \approx_c ([a_1], [a_2], [a_1 r_1], [a_2 r_2], [z]) \in \mathbb{G}^5$, where $a_1, a_2, r_1, r_2, z \leftarrow \mathbb{Z}_q$. 2-Lin holds in generic bilinear groups [3] and it has virtually become the standard decisional assumption in the group $\mathbb{G}$ in the bilinear setting. It has found applications to encryption [5, 7, 29, 38], signatures [3], zero-knowledge proofs [21], pseudorandom functions [6] and many more. More recently, the 2-Lin Assumption was generalized to the $(k\text{-Lin})_{k\in\mathbb{N}}$ Assumption family [23, 45] (1-Lin = DDH), a family of increasingly (strictly) weaker Assumptions which are generically hard in k-linear maps.

Subgroup Membership Problems. Since the work of Cramer and Shoup [13] it has been recognized that it is useful to view the DDH Assumption as a hard subgroup membership problem in $\mathbb{G}^2$. In this formulation, the DDH Assumption states that it is hard to decide whether a given element $([r], [t]) \in \mathbb{G}^2$ is contained in the subgroup generated by $([1], [a])$. Similarly, in this language the 2-Lin Assumption says that it is hard to decide whether a given vector $([r], [s], [t]) \in \mathbb{G}^3$ is in the subgroup generated by the vectors $([a_1], [0], [1]), ([0], [a_2], [1])$. The same holds for the $(k\text{-Lin})_{k\in\mathbb{N}}$ Assumption family: for each k, the k-Lin assumption can be naturally written as a hard subgroup membership problem in $\mathbb{G}^{k+1}$. This alternative formulation has conceptual advantages for some applications, for instance, it allowed to provide more instantiations of the original DDH-based scheme of Cramer and Shoup and it is also the most natural point of view for translating schemes originally constructed in composite order groups into prime order groups [18, 36, 43, 44].

Linear Algebra in Bilinear Groups. In its formulation as subgroup decision membership problem, the k-Lin assumption can be seen as the problem of deciding linear dependence "in the exponent." Recently, a number of works have illustrated the usefulness of a more algebraic point of view on decisional assumptions in bilinear groups, like the Dual Pairing Vector Spaces of Okamoto and Takashima [40] or the Subspace Assumption of Lewko [32]. Although these new decisional assumptions reduce to the 2-Lin Assumption, their flexibility and their algebraic description have proven to be crucial in many works to obtain complex primitives in strong security models previously unrealized in the literature, like Attribute-Based Encryption, Unbounded Inner Product Encryption and many more (see [32, 41, 42], just to name a few).

This Work. Motivated by the success of this algebraic viewpoint of decisional assumptions, in this paper we explore new insights resulting from interpreting the k-Lin decisional assumption as a special case of what we call a Matrix Diffie-Hellman Assumption. The general problem states that it is hard to distinguish whether a given vector in $\mathbb{G}^\ell$ is contained in the space spanned by the columns of a certain matrix $[\mathbf{A}] \in \mathbb{G}^{\ell \times k}$, where $\mathbf{A}$ is sampled according to some distribution $\mathcal{D}_{\ell,k}$. We remark that even though all our results are stated in symmetric bilinear groups, they can be naturally extended to the asymmetric setting.

1.1 The Matrix Diffie-Hellman Assumption

A New Framework for DDH-like Assumptions. For integers $\ell > k$ let $\mathcal{D}_{\ell,k}$ be an (efficiently samplable) distribution over $\mathbb{Z}_q^{\ell \times k}$. We define the $\mathcal{D}_{\ell,k}$-Matrix DH ($\mathcal{D}_{\ell,k}$-MDDH) Assumption as the following subgroup

decision assumption:

$$\mathcal{D}_{\ell,k}\text{-MDDH}: \quad [\mathbf{A}||\mathbf{A}\vec{r}] \approx_c [\mathbf{A}||\vec{u}] \in \mathbb{G}^{\ell\times(k+1)}, \tag{1}$$

where $\mathbf{A} \in \mathbb{Z}_q^{\ell\times k}$ is chosen from distribution $\mathcal{D}_{\ell,k}$, $\vec{r} \leftarrow \mathbb{Z}_q^k$, and $\vec{u} \leftarrow \mathbb{G}^\ell$. The $(k\text{-}\mathsf{Lin})_{k\in\mathbb{N}}$ family corresponds to this problem when $\ell = k+1$, and $\mathcal{D}_{\ell,k}$ is the specific distribution $\mathcal{L}_k$ (formally defined in Example 2).

Generic Hardness. Due to its linearity properties, the $\mathcal{D}_{\ell,k}$-MDDH Assumption does not hold in $(k+1)$-linear groups. In Section 3.3 we give two different theorems which state sufficient conditions for the $\mathcal{D}_{\ell,k}$-MDDH Assumption to hold generically in m-linear groups. Theorem 3 is very similar to the Uber-Assumption [2, 9] that characterizes hardness in bilinear groups (i.e., $m = 2$) in terms of linear independence of polynomials in the inputs. We generalize this to arbitrary m using a more algebraic language. This algebraic formulation has the advantage that one can use additional tools (e.g. Gröbner bases or resultants) to show that a distribution $\mathcal{D}_{\ell,k}$ meets the conditions of Theorem 3, which is specially important for large m. It also allows to prove a completely new result, namely Theorem 4, which states that a matrix assumption with $\ell = k+1$ is generically hard if a certain determinant polynomial is irreducible.

New Assumptions for Bilinear Groups. We propose other families of generically hard decisional assumptions that did not previously appear in the literature, e.g., those associated to $\mathcal{C}_k, \mathcal{SC}_k, \mathcal{IL}_k$ defined below. For the most important parameters $k = 2$ and $\ell = k+1 = 3$, we consider the following examples of distributions:

$$\mathcal{C}_2 : \mathbf{A} = \begin{pmatrix} a_1 & 0 \\ 1 & a_2 \\ 0 & 1 \end{pmatrix} \quad \mathcal{SC}_2 : \mathbf{A} = \begin{pmatrix} a & 0 \\ 1 & a \\ 0 & 1 \end{pmatrix} \quad \mathcal{L}_2 : \mathbf{A} = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \\ 1 & 1 \end{pmatrix} \quad \mathcal{IL}_2 : \mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & a+1 \\ 1 & 1 \end{pmatrix},$$

for uniform $a, a_1, a_2 \in \mathbb{Z}_q$ as well as $\mathcal{U}_{3,2}$, the uniform distribution in $\mathbb{Z}_q^{3\times 2}$ (already considered in [5, 19, 38, 46]). All assumptions are hard in generic bilinear groups. It is easy to verify that $\mathcal{L}_2$-MDDH = 2-Lin. We define 2-Casc := $\mathcal{C}_2$-MDDH (Cascade Assumption), 2-SCasc := $\mathcal{SC}_2$-MDDH (Symmetric Cascade Assumption), and 2-ILin := $\mathcal{IL}_2$-MDDH (Incremental Linear Assumption). In Section 3.4, we show that 2-SCasc $\Rightarrow$ 2-Casc, 2-ILin $\Rightarrow$ 2-Lin and that $\mathcal{U}_{3,2}$-MDDH is the weakest of these assumptions (which extends the results of [18, 19, 46] for 2-Lin). Although originally [16] 2-ILin and 2-SCasc were thought to be incomparable assumptions, in Section 4 we show that 2-SCasc and 2-ILin are indeed equivalent assumptions. The equivalence result, together with the fact that 2-ILin $\Rightarrow$ 2-Lin, imply that 2-SCasc is a stronger assumption than 2-Lin.

Efficiency Improvements. As a measure of efficiency, we define the *representation size* $\mathsf{RE}_{\mathbb{G}}(\mathcal{D}_{\ell,k})$ of an $\mathcal{D}_{\ell,k}$-MDDH assumption as the minimal number of group elements needed to represent $[\mathbf{A}]$ for any $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$. This parameter is important since it affects the performance (typically the size of public/secret parameters) of schemes based on a Matrix Diffie-Hellman Assumption. 2-Lin and 2-Casc have representation size 2 (elements $([a_1], [a_2])$), while 2-SCasc only 1 (element $[a]$). Hence our new assumptions directly translate into shorter parameters for a large number of applications (see the Applications in Section 5). Further, our result points out a tradeoff between efficiency and hardness which questions the role of 2-Lin as the "standard decisional assumption" over a bilinear group $\mathbb{G}$.

New Families of Weaker Assumptions. By defining appropriate distributions $\mathcal{C}_k$, $\mathcal{SC}_k$, $\mathcal{IL}_k$ over $\mathbb{Z}_q^{(k+1)\times k}$, for any $k \in \mathbb{N}$, one can generalize all three new assumptions naturally to k-Casc, k-SCasc, and k-ILin with representation size k, 1, and 1, respectively. Using our results on generic hardness, it is easy to verify that all three assumptions are generically hard in k-linear groups. Actually, in Section 4 we show that k-SCasc, and k-ILin are equivalent for every k. Since all these assumptions are false in $(k+1)$-linear groups this gives us three new families of increasingly strictly weaker assumptions[1]. In particular, the k-SCasc (equivalently, k-ILin) assumption family is of great interest due to its compact representation size of only 1 element.

Relations to Other Standard Assumptions. Surprisingly, the new assumption families can also be related to standard assumptions. The k-Casc Assumption is implied by the $(k+1)$-Party Diffie-Hellman

[1] We actually assume that k and ℓ are considered as constants, i.e. they do not depend on the security parameter. Otherwise, for a general $\mathcal{D}_{\ell,k}$ it is not so easy to solve the $\mathcal{D}_{\ell,k}$-MDDH problem with the only help of a $(k+1)$-linear map, because determinants of size $k+1$ could not be computable in polynomial time.

Assumption ($(k+1)$-PDDH) [7] which states that $([a_1], \ldots, [a_{k+1}], [a_1 \cdots a_{k+1}]) \approx_c ([a_1], \ldots, [a_{k+1}], [z]) \in \mathbb{G}^{k+2}$. Similarly, k-SCasc is implied by the $(k+1)$-Exponent Diffie-Hellman Assumption ($(k+1)$-EDDH) [28] which states that $([a], [a^{k+1}]) \approx_c ([a], [z]) \in \mathbb{G}^2$. Figure 1 on page 11 gives an overview over the relations between the different assumptions.

Uniqueness of One-Parameter Family. The most natural and useful $\mathcal{D}_{\ell,k}$-MDDH assumptions are those with $\ell = k+1$ and the entries of the matrices generated by $\mathcal{D}_{\ell,k}$ are polynomials of degree one in some parameters. Among them, the most compact correpond to the one-parameter distributions. As novel contribution with respect to [16], in Section 4 we show that k-ILin and k-SCasc are tightly equivalent. Moreover, we prove that every $\mathcal{D}_k$-MDDH assumption defined by univariate polynomials of degree one is tightly equivalent to k-SCasc, so we can see k-SCasc as a sort of canonical compact Matrix DH assumption. From the equivalence proof between k-ILin and k-SCasc one can easily construct a reduction from k-SCasc to k-Lin.

1.2 Basic Applications

We believe that all schemes based on 2-Lin can be shown to work for any Matrix Assumption. Consequently, a large class of known schemes can be instantiated more efficiently with the new more compact decisional assumptions, while offering the same generic security guarantees. To support this belief, in Section 5 we show how to construct some fundamental primitives based on any Matrix Assumption. All constructions are purely algebraic and therefore very easy to understand and prove.

- **Public-key Encryption.** We build a key-encapsulation mechanism with security against passive adversaries from any $\mathcal{D}_{\ell,k}$-MDDH Assumption. The public-key is $[\mathbf{A}]$, the ciphertext consists of the first k elements of $[z] = [\mathbf{A}\vec{r}]$, the symmetric key of the last $\ell - k$ elements of $[z]$. Passive security immediately follows from $\mathcal{D}_{\ell,k}$-MDDH.

- **Hash Proof Systems.** We build a smooth projective hash proof system (HPS) from any $\mathcal{D}_{\ell,k}$-MDDH Assumption. It is well-known that HPS imply chosen-ciphertext secure encryption [13], password-authenticated key-exchange [20], zero-knowledge proofs [1], and many other things.

- **Pseudo-Random Functions.** Generalizing the Naor-Reingold PRF [6, 37], we build a pseudo-random function PRF from any $\mathcal{D}_{\ell,k}$-MDDH Assumption. The secret-key consists of *transformation matrices* $\mathbf{T}_1, \ldots, \mathbf{T}_n$ (derived from independent instances $\mathbf{A}_{i,j} \leftarrow \mathcal{D}_{\ell,k}$) plus a vector $\vec{h}$ of group elements. For $x \in \{0,1\}^n$ we define $\mathsf{PRF}_K(x) = \left[\prod_{i:x_i=1} \mathbf{T}_i \cdot \vec{h}\right]$. Using the random self-reducibility of the $\mathcal{D}_{\ell,k}$-MDDH Assumption, we give a tight security proof.

- **Groth-Sahai Non-Interactive Zero-Knowledge Proofs.** Groth and Sahai [21] proposed very elegant and efficient non-interactive zero-knowledge (NIZK) and non-interactive witness-indistinguishable (NIWI) proofs that work directly for a wide class of languages that are relevant in practice. We show how to instantiatiate their proof system based on any $\mathcal{D}_{\ell,k}$-MDDH Assumption. While the size of the proofs depends only on ℓ and k, the CRS and verification depends on the representation size of the Matrix Assumptions. Therefore our new instantiations offer improved efficiency over the 2-Lin-based construction from [21]. This application in particular highlights the usefulness of the Matrix Assumption to describe in a compact way many instantiations of a scheme: instead of having to specify the constructions for the DDH and the 2-Lin assumptions separately [21], we can recover them as a special case of a general construction.

More Efficient Proofs for CRS Dependent Languages. In Section 6 we provide more efficient NIZK proofs for concrete natural languages which are dependent on the common reference string. More specifically, the common reference string of the $\mathcal{D}_{\ell,k}$-MDDH instantiation of Groth-Sahai proofs of Section 5.4 includes as part of the commitment keys the matrix $[\mathbf{A}]$, where $\mathbf{A} \in \mathbb{Z}_q^{\ell \times k} \leftarrow \mathcal{D}_{\ell,k}$. We give more efficient proofs for several languages related to $\mathbf{A}$. Although at first glance the languages considered may

seem quite restricted, they naturally appear in many applications, where typically $\mathbf{A}$ is the public key of some encryption scheme and one wants to prove statements about ciphertexts. More specifically, we obtain improvements for several kinds of statements, namely:

- **Subgroup Membership Proofs**. We give more efficient proofs in the language $\mathcal{L}_{\mathbf{A},\mathbb{G},\mathcal{P}} := \{[\mathbf{A}\vec{r}], \vec{r} \in \mathbb{Z}_q^k\} \subset \mathbb{G}^\ell$. To quantify some concrete improvement, in the 2-Lin case, our proofs of membership are half of the size of a standard Groth-Sahai proof and they require only 6 groups elements. We stress that this improvement is obtained without introducing any new computational assumption. As an example of application, consider for instance the encryption scheme derived from our KEM based on any $\mathcal{D}_{\ell,k}$-MDDH Assumption, where the public key is some matrix $[\mathbf{A}]$, $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$. To see which kind of statements can be proved using our result, note that a ciphertext is a rerandomization of another one only if their difference is in $\mathcal{L}_{\mathbf{A},\mathbb{G},\mathcal{P}}$. The same holds for proving that two commitments with the same key hide the same value or for showing in a publicly verifiable manner that the ciphertext of our encryption scheme opens to some known message $[m]$. This improvement has a significant impact on recent results, like [17, 35], and we think many more examples can be found. Interestingly, in independent work, a number of results ([25, 26, 31, 34]) have constructed even more efficient proofs in linear subspaces by also exploiting the dependency of the common reference string and the matrix which generates the space. We note that although in all these works proofs are shorter, this is at the cost of having only *computationally sound proofs*, while our results retain the perfect soundness inherited from Groth Sahai proofs.

- **Ciphertext Validity and Plaintext Equality.** Similar techniques apply to get more efficient proofs of statements which naturally appear when one wants to prove that a ciphertext is valid and that two ciphertexts encrypted with different public keys open to the same plaintext, e.g., when using Naor-Yung techniques to obtain chosen-ciphertext security [39], like in the encryption schemes of [10, 15, 22, 27].

2 Preliminaries

2.1 Notation

For $n \in \mathbb{N}$, we write 1^n for the string of n ones. Moreover, $|x|$ denotes the length of a bitstring x, while $|S|$ denotes the size of a set S. Further, $s \leftarrow S$ denotes the process of sampling an element s from S uniformly at random. For an algorithm A, we write $z \leftarrow \mathsf{A}(x, y, \ldots)$ to indicate that A is a (probabilistic) algorithm that outputs z on input $(x, y, \ldots)$. If $\mathbf{A}$ is a matrix we denote by a_{ij} the entries and $\vec{a}_i$ the column vectors.

2.2 Representing Elements in Groups

Let Gen be a probabilistic polynomial time (ppt) algorithm that on input 1^λ returns a description $\mathcal{G} = (\mathbb{G}, q, \mathcal{P})$ of a cyclic group $\mathbb{G}$ of order q for a λ-bit prime q and a generator $\mathcal{P}$ of $\mathbb{G}$. More generally, for any fixed $k \geq 1$, let MGen_k be a ppt algorithm that on input 1^λ returns a description $\mathcal{MG}_k = (\mathbb{G}, \mathbb{G}_{T_k}, q, e_k, \mathcal{P})$, where $\mathbb{G}$ and $\mathbb{G}_{T_k}$ are cyclic additive groups of prime-order q, $\mathcal{P}$ a generator of $\mathbb{G}$, and $e_k : \mathbb{G}^k \to \mathbb{G}_{T_k}$ is a (non-degenerated, efficiently computable) k-linear map. For $k = 2$ we define $\mathsf{PGen} := \mathsf{MGen}_2$ to be a generator of a bilinear group $\mathcal{PG} = (\mathbb{G}, \mathbb{G}_T, q, e, \mathcal{P})$.

For an element $a \in \mathbb{Z}_q$ we define $[a] = a\mathcal{P}$ as the implicit representation of a in $\mathbb{G}$. More generally, for a matrix $\mathbf{A} = (a_{ij}) \in \mathbb{Z}_q^{n \times m}$ we define $[\mathbf{A}]$ as the implicit representation of $\mathbf{A}$ in $\mathbb{G}$ and $[\mathbf{A}]_{T_k}$ as the implicit representation of $\mathbf{A}$ in $\mathbb{G}_{T_k}$:

$$[\mathbf{A}] := \begin{pmatrix} a_{11}\mathcal{P} & \ldots & a_{1m}\mathcal{P} \\ & & \\ a_{n1}\mathcal{P} & \ldots & a_{nm}\mathcal{P} \end{pmatrix} \in \mathbb{G}^{n \times m}, \quad [\mathbf{A}]_{T_k} := \begin{pmatrix} a_{11}\mathcal{P}_{T_k} & \ldots & a_{1m}\mathcal{P}_{T_k} \\ & & \\ a_{n1}\mathcal{P}_{T_k} & \ldots & a_{nm}\mathcal{P}_{T_k} \end{pmatrix} \in \mathbb{G}_{T_k}^{n \times m},$$

where $\mathcal{P}_{T_k} = e_k(\mathcal{P}, \ldots, \mathcal{P}) \in \mathbb{G}_{T_k}$.

When talking about elements in $\mathbb{G}$ and $\mathbb{G}_{T_k}$ we will always use this implicit notation, i.e., we let $[a] \in \mathbb{G}$ be an element in $\mathbb{G}$ or $[b]_{T_k}$ be an element in $\mathbb{G}_{T_k}$. Note that from $[a] \in \mathbb{G}$ it is generally hard to compute the value a (discrete logarithm problem in $\mathbb{G}$). Further, from $[b]_{T_k} \in \mathbb{G}_{T_k}$ it is hard to compute the value $b \in \mathbb{Z}_q$ (discrete logarithm problem in $\mathbb{G}_{T_k}$) or the value $[b] \in \mathbb{G}$ (pairing inversion problem). Obviously, given $[a] \in \mathbb{G}$, $[b]_{T_k} \in \mathbb{G}_{T_k}$, and a scalar $x \in \mathbb{Z}_q$, one can efficiently compute $[ax] \in \mathbb{G}$ and $[bx]_{T_k} \in \mathbb{G}_{T_k}$.

Also, all functions and operations acting on $\mathbb{G}$ and $\mathbb{G}_{T_k}$ will be defined implicitly. For example, when evaluating a bilinear pairing $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ in $[a], [b] \in \mathbb{G}$ we will use again our implicit representation and write $[z]_T := e([a], [b])$. Note that $e([a], [b]) = [ab]_T$, for all $a, b \in \mathbb{Z}_q$.

2.3 Standard Diffie-Hellman Assumptions

Let Gen be a ppt algorithm that on input 1^λ returns a description $\mathcal{G} = (\mathbb{G}, q, \mathcal{P})$ of cyclic group $\mathbb{G}$ of prime-order q and a generator $\mathcal{P}$ of $\mathbb{G}$. Similarly, let PGen be a ppt algorithm that returns a description $\mathcal{PG} = (\mathbb{G}, \mathbb{G}_T, q, e, \mathcal{P})$ of a pairing group. We informally recall a number of previously considered Decisional Diffie-Hellman Assumptions.

- **Diffie-Hellman (DDH) Assumption.** It is hard to distinguish $(\mathcal{G}, [x], [y], [xy])$ from $(\mathcal{G}, [x], [y], [z])$, for $\mathcal{G} = (\mathbb{G}, q, \mathcal{P}) \leftarrow \mathsf{Gen}$, $x, y, z \leftarrow \mathbb{Z}_q$.
- **k-Linear (k-Lin) Assumption [3, 23, 45].** It is hard to distinguish $(\mathcal{G}, [x_1], [x_2], \ldots [x_k], [r_1 x_1], [r_2 x_2], \ldots [r_k x_k], [r_1 + \cdots + r_k])$ from $(\mathcal{G}, [x_1], [x_2], \ldots [x_k], [r_1 x_1], [r_2 x_2], \ldots [r_k x_k], [z])$, for $\mathcal{G} \leftarrow \mathsf{Gen}$, $x_1, \ldots, x_k, r_1, \ldots, r_k, z \leftarrow \mathbb{Z}_q$. Clearly, 1-Lin = DDH.
- **Bilinear Diffie-Hellman (BDDH) Assumption [4].** It is hard to distinguish $(\mathcal{PG}, [x], [y], [z], [xyz]_T)$ from $(\mathcal{PG}, [x], [y], [z], [w]_T)$, for $\mathcal{PG} \leftarrow \mathsf{PGen}$, $x, y, z, w \leftarrow \mathbb{Z}_q$.
- **k-Multilinear Diffie-Hellman (k-MLDDH) Assumption [8].** Given k-linear group generator MGen_k it is hard to distinguish $(\mathcal{MG}_k, [x_1], \ldots [x_{k+1}], [x_1 \cdots x_{k+1}]_{T_k})$ from $(\mathcal{MG}_k, [x_1], \ldots [x_{k+1}], [z]_{T_k})$, for $\mathcal{MG}_k \leftarrow \mathsf{MGen}_k$, $x_1, \ldots, x_{k+1}, z \leftarrow \mathbb{Z}_q$. Clearly, 2-MLDDH = BDDH.
- **k-Party Diffie-Hellman (k-PDDH) Assumption.** It is hard to distinguish $(\mathcal{G}, [x_1], [x_2], \ldots [x_k], [x_1 \cdots x_k])$ from $(\mathcal{G}, [x_1], [x_2], \ldots, [x_k], [z])$, for $\mathcal{G} \leftarrow \mathsf{Gen}$, $x_1, \ldots, x_k, z \leftarrow \mathbb{Z}_q$. 2-PDDH = DDH and 3-PDDH was proposed in [7].
- **k-Exponent Diffie-Hellman (k-EDDH) Assumption [28, 47].** It is hard to distinguish $(\mathcal{G}, [x], [x^k])$ from $(\mathcal{G}, [x], [z])$, for $\mathcal{G} \leftarrow \mathsf{Gen}$, $x, z \leftarrow \mathbb{Z}_q$.

2.4 Key Encapsulation Mechanisms

A *key-encapsulation mechanism* $\mathsf{KEM} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ with key-space $\mathcal{K}(\lambda)$ consists of three polynomial-time algorithms (PTAs). Via $(pk, sk) \leftarrow \mathsf{Gen}(1^\lambda)$ the randomized key-generation algorithm produces public/secret keys for security parameter $\lambda \in \mathbb{N}$; via $(K, c) \leftarrow \mathsf{Enc}(pk)$ the randomized encapsulation algorithm creates a uniformly distributed symmetric key $K \in \mathcal{K}(\lambda)$ together with a ciphertext c; via $K \leftarrow \mathsf{Dec}(sk, c)$ the possessor of secret key sk decrypts ciphertext c to get back a key K which is an element in $\mathcal{K}$ or a special rejection symbol $\perp$. For consistency, we require that for all $\lambda \in \mathbb{N}$, and all $(K, c) \leftarrow \mathsf{Enc}(pk)$ we have $\Pr[\mathsf{Dec}(sk, c) = K] = 1$, where the probability is taken over the choice of $(pk, sk) \leftarrow \mathsf{Gen}(1^\lambda)$, and the coins of all the algorithms in the expression above.

For IND-CPA security we require that the distribution $(pk, (c, K))$ is computationally indistinguishable from $(pk, (c, K'))$, where $(pk, sk) \leftarrow \mathsf{Gen}(1^\lambda)$, $(K, c) \leftarrow \mathsf{Enc}(pk)$, and $K' \leftarrow \mathcal{K}(\lambda)$. An IND-CPA secure KEM implies an IND-CPA secure publick-key encryption (PKE) scheme by combining it with a one-time secure symmetric cipher (DEM).

2.5 Hash Proof Systems

We recall the notion of hash proof systems as introduced by Cramer and Shoup [13].

Let $\mathcal{C}, \mathcal{K}$ be sets and $\mathcal{V} \subset \mathcal{C}$ a language. In the context of public-key encryption (and viewing a hash proof system as a key encapsulation mechanism (KEM) [14] with "special algebraic properties") one may think of $\mathcal{C}$ as the set of all *ciphertexts*, $\mathcal{V} \subset \mathcal{C}$ as the set of all *valid (consistent) ciphertexts*, and $\mathcal{K}$ as the set of all *symmetric keys*. Let $\Lambda_{sk} : \mathcal{C} \to \mathcal{K}$ be a hash function indexed with $sk \in \mathcal{SK}$, where $\mathcal{SK}$ is a set. A hash function Λ_{sk} is projective if there exists a projection $\mu : \mathcal{SK} \to \mathcal{PK}$ such that $\mu(sk) \in \mathcal{PK}$ defines the action of Λ_{sk} over the subset $\mathcal{V}$. That is, for every $c \in \mathcal{V}$, the value $K = \Lambda_{sk}(c)$ is uniquely determined by $\mu(sk)$ and c. In contrast, nothing is guaranteed for $c \in \mathcal{C} \setminus \mathcal{V}$, and it may not be possible to compute $\Lambda_{sk}(c)$ from $\mu(sk)$ and c. The projective hash function is (perfectly) universal$_1$ if for all $c \in \mathcal{C} \setminus \mathcal{V}$,

$$(pk, \Lambda_{sk}(c)) \equiv (pk, K) \tag{2}$$

where in the above $pk = \mu(sk)$ for $sk \leftarrow \mathcal{SK}$ and $K \leftarrow \mathcal{K}$, and the symbol $\equiv$ stands for equality of the two distributions.

A hash proof system $\mathsf{HPS} = (\mathsf{Param}, \mathsf{Pub}, \mathsf{Priv})$ consists of three algorithms where the randomized algorithm $\mathsf{Param}(1^\lambda)$ generates instances of $params = (\mathcal{S}, \mathcal{K}, \mathcal{C}, \mathcal{V}, \mathcal{PK}, \mathcal{SK}, \Lambda_{(\cdot)} : \mathcal{C} \to \mathcal{K}, \mu : \mathcal{SK} \to \mathcal{PK})$, where $\mathcal{S}$ may contain some additional structural parameters such as the group description. The deterministic public evaluation algorithm Pub inputs the projection key $pk = \mu(sk)$, $c \in \mathcal{V}$ and a witness w of the fact that $c \in \mathcal{V}$ and returns $K = \Lambda_{sk}(c)$. The deterministic private evaluation algorithm inputs $sk \in \mathcal{SK}$ and returns $\Lambda_{sk}(c)$, without knowing a witness. We further assume there are efficient algorithms given for sampling $sk \in \mathcal{SK}$ and sampling $c \in \mathcal{V}$ uniformly together with a witness w.

As computational problem we require that the *subset membership problem* is hard in HPS which means that the two elements c and c' are computationally indistinguishable, for uniform $c \in \mathcal{V}$ and uniform $c' \in \mathcal{C} \setminus \mathcal{V}$.

2.6 Pseudo-Random Functions

A pseudo-random function $\mathsf{PRF} = (\mathsf{Gen}, \mathsf{F})$ with respect to range $\mathcal{R} = \mathcal{R}(\lambda)$ and message space $\mathcal{M} = \mathcal{M}(\lambda)$ consists of two algorithms, where the randomized algorithm $\mathsf{Gen}(1^\lambda)$ generates a symmetric key K and the deterministic evaluation algorithm $\mathsf{F}_K(x)$ outputs a value in $\mathcal{R}$, for all $x \in \mathcal{M}$. For security we require that an adversary making polynomially many queries to an oracle $\mathcal{O}(\cdot)$, the output of oracle $\mathcal{O}(x) = \mathsf{F}_K(x)$ for a fixed key $K \leftarrow \mathsf{Gen}(1^\lambda)$ is computationally indistinguishable from $\mathcal{O}(x) = f(x)$, where f is chosen uniformly from all functions from mapping $\mathcal{M}$ to $\mathcal{R}$ (i.e., $f(x)$ outputs uniform elements in $\mathcal{R}$).

3 Matrix DH assumptions

3.1 Definition

Definition 1. *Let $\ell, k \in \mathbb{N}$ with $\ell > k$. We call $\mathcal{D}_{\ell,k}$ a matrix distribution if it outputs (in poly time, with overwhelming probability) matrices in $\mathbb{Z}_q^{\ell \times k}$ of full rank k. We define $\mathcal{D}_k := \mathcal{D}_{k+1,k}$.*

For simplicity we will also assume that, wlog, the first k rows of $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$ form an invertible matrix.

We define the $\mathcal{D}_{\ell,k}$-matrix problem as to distinguish the two distributions $([\mathbf{A}], [\mathbf{A}\vec{w}])$ and $([\mathbf{A}], [\vec{u}])$, where $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$, $\vec{w} \leftarrow \mathbb{Z}_q^k$, and $\vec{u} \leftarrow \mathbb{Z}_q^\ell$.

Definition 2 ($\mathcal{D}_{\ell,k}$-Matrix Diffie-Hellman Assumption $\mathcal{D}_{\ell,k}$-MDDH)**.** *Let $\mathcal{D}_{\ell,k}$ be a matrix distribution. We say that the $\mathcal{D}_{\ell,k}$-Matrix Diffie-Hellman ($\mathcal{D}_{\ell,k}$-MDDH) Assumption holds relative to Gen if for all ppt adversaries D,*

$$\mathbf{Adv}_{\mathcal{D}_{\ell,k},\mathsf{Gen}}(\mathsf{D}) = \Pr[\mathsf{D}(\mathcal{G}, [\mathbf{A}], [\mathbf{A}\vec{w}]) = 1] - \Pr[\mathsf{D}(\mathcal{G}, [\mathbf{A}], [\vec{u}]) = 1] = negl(\lambda),$$

where the probability is taken over $\mathcal{G} = (\mathbb{G}, q, \mathcal{P}) \leftarrow \mathsf{Gen}(1^\lambda)$, $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}, \vec{w} \leftarrow \mathbb{Z}_q^k, \vec{u} \leftarrow \mathbb{Z}_q^\ell$ and the coin tosses of adversary D.

Definition 3. *Let $\mathcal{D}_{\ell,k}$ be a matrix distribution. Let $\mathbf{A}_0$ be the first k rows of $\mathbf{A}$ and $\mathbf{A}_1$ be the last $\ell - k$ rows of $\mathbf{A}$. The matrix $\mathbf{T} \in \mathbb{Z}_q^{(\ell-k)\times k}$ defined as $\mathbf{T} = \mathbf{A}_1\mathbf{A}_0^{-1}$ is called the transformation matrix of $\mathbf{A}$.*

We note that using the transformation matrix, one can alternatively define the advantage from Definition 2 as

$$\mathbf{Adv}_{\mathcal{D}_{\ell,k},\mathsf{Gen}}(\mathsf{D}) = \Pr[\mathsf{D}(\mathcal{G}, \begin{bmatrix}\mathbf{A}_0\\ \mathbf{T}\mathbf{A}_0\end{bmatrix}, \begin{bmatrix}\vec{h}\\ \mathbf{T}\vec{h}\end{bmatrix}) = 1] - \Pr[\mathsf{D}(\mathcal{G}, \begin{bmatrix}\mathbf{A}_0\\ \mathbf{T}\mathbf{A}_0\end{bmatrix}, \begin{bmatrix}\vec{h}\\ \vec{u}\end{bmatrix}) = 1],$$

where the probability is taken over $\mathcal{G} = (\mathbb{G}, q, \mathcal{P}) \leftarrow \mathsf{Gen}(1^\lambda)$, $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}, \vec{h} \leftarrow \mathbb{Z}_q^k, \vec{u} \leftarrow \mathbb{Z}_q^{\ell-k}$ and the coin tosses of adversary D.

3.2 Basic Properties

We can generalize Definition 2 to the m-fold $\mathcal{D}_{\ell,k}$-MDDH Assumption as follows. Given $\mathbf{W} \leftarrow \mathbb{Z}_q^{k\times m}$ for some $m \geq 1$, we consider the problem of distinguishing the distributions $([\mathbf{A}], [\mathbf{AW}])$ and $([\mathbf{A}], [\mathbf{U}])$ where $\mathbf{U} \leftarrow \mathbb{Z}_q^{\ell\times m}$ is equivalent to m independent instances of the problem (with the same $\mathbf{A}$ but different $\vec{w}_i$). This can be proved through a hybrid argument with a loss of m in the reduction, or, with a tight reduction (independent of m) via random self-reducibility.

Lemma 1 (Random self reducibility)**.** *For any matrix distribution $\mathcal{D}_{\ell,k}$, $\mathcal{D}_{\ell,k}$-MDDH is random self-reducible. Concretely, for any m,*

$$\mathbf{Adv}^m_{\mathcal{D}_{\ell,k},\mathsf{Gen}}(\mathsf{D}') \leq \begin{cases} m \cdot \mathbf{Adv}_{\mathcal{D}_{\ell,k},\mathsf{Gen}}(\mathsf{D}) & 1 \leq m \leq \ell - k \\ (\ell - k) \cdot \mathbf{Adv}_{\mathcal{D}_{\ell,k},\mathsf{Gen}}(\mathsf{D}) + \dfrac{1}{q-1} & m > \ell - k \end{cases},$$

where

$$\mathbf{Adv}^m_{\mathcal{D}_{\ell,k},\mathsf{Gen}}(\mathsf{D}') = \Pr[\mathsf{D}'(\mathcal{G}, [\mathbf{A}], [\mathbf{AW}]) = 1] - \Pr[\mathsf{D}'(\mathcal{G}, [\mathbf{A}], [\mathbf{U}]) = 1],$$

and the probability is taken over $\mathcal{G} = (\mathbb{G}, q, \mathcal{P}) \leftarrow \mathsf{Gen}(1^\lambda)$, $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}, \mathbf{W} \leftarrow \mathbb{Z}_q^{k\times m}, \mathbf{U} \leftarrow \mathbb{Z}_q^{\ell\times m}$ and the coin tosses of adversary D'.

Proof. The case $1 \leq m \leq \ell - k$ comes from a natural hybrid argument, while the case $m > \ell - k$ is obtained from the inequality

$$\mathbf{Adv}^m_{\mathcal{D}_{\ell,k},\mathsf{Gen}}(\mathsf{D}') \leq \mathbf{Adv}^{\ell-k}_{\mathcal{D}_{\ell,k},\mathsf{Gen}}(\mathsf{D}) + \frac{1}{q-1}.$$

To prove it, we show that there exists an efficient transformation of any instance $([\mathbf{A}], [\mathbf{Z}])$ of the $(\ell-k)$-fold $\mathcal{D}_{\ell,k}$-MDDH problem into another instance $([\mathbf{A}], [\mathbf{Z}'])$ of the m-fold problem, with overwhelming probability.

In particular, we set $\mathbf{Z}' = \mathbf{AR} + \mathbf{ZC}$, for random matrices $\mathbf{R} \leftarrow \mathbb{Z}_q^{k\times m}$ and $\mathbf{C} \leftarrow \mathbb{Z}_q^{(\ell-k)\times m}$. On the one hand, if $\mathbf{Z} = \mathbf{AW}$ then $\mathbf{Z}' = \mathbf{AW}'$ for $\mathbf{W}' = \mathbf{R} + \mathbf{WC}$, which is uniformly distributed in $\mathbb{Z}_q^{k\times m}$. On the other hand, if $\mathbf{Z} = \mathbf{U}$ is uniform then $\mathbf{A}|\mathbf{U}$ is full-rank with probability at least $1 - 1/(q-1)$. In that case, $\mathbf{Z}' = \mathbf{AR} + \mathbf{UC}$ is uniformly distributed in $\mathbb{Z}_q^{\ell\times m}$, which proves the above inequality. □

We remark that, given $[\mathbf{A}], [\vec{z}]$ the above lemma can only be used to re-randomize the value $[\vec{z}]$. In order to re-randomize the matrix $[\mathbf{A}]$ we need that one can sample matrices $\mathbf{L}$ and $\mathbf{R}$ such that $\mathbf{A}' = \mathbf{LAR}$ looks like an independent instance $\mathbf{A}' \leftarrow \mathcal{D}_{\ell,k}$. In all of our example distributions we are able to do this.

Due to its linearity properties, the $\mathcal{D}_{\ell,k}$-MDDH assumption does not hold in $(k+1)$-linear groups, assuming that k is constant, i.e. it does not depend on the security parameter[2].

Lemma 2. *Let $\mathcal{D}_{\ell,k}$ be any matrix distribution. Then the $\mathcal{D}_{\ell,k}$-Matrix Diffie-Hellman Assumption is false in $(k+1)$-linear groups.*

[2]If k grows linearly with the security parameter, computing determinants of size $k+1$ in $\mathbb{G}$ could in general take exponential time. However, for the particular matrices in the forthcoming examples (except for the uniform distribution) the associated determinants are still efficiently computable, and the Matrix DH Assumption is also false in $(k+1)$-linear groups.

Proof. In a $(k+1)$-linear group, the implicit representation of any $r \times r$ determinant for $r \leq k+1$ can be efficiently computed by using the r-linear map given by the Leibnitz formula:

$$\det(\mathbf{M}) = \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) \prod_{i=1}^{r} m_{i,\sigma_i}$$

Using the $(k+1)$-linear map, $[\det(\mathbf{M})]_{T_k}$ can be computed in the target group. Then, given $[\mathbf{B}] := [\mathbf{A}||\vec{z}]$, consider the submatrix $\mathbf{A}_0$ formed by the first k rows of $\mathbf{A}$ and the vector $\vec{z}_0$ formed by the first k elements of $\vec{z}$. If $\det(\mathbf{A}_0) \neq 0$, then define $\mathbf{C}$ as the first $k+1$ rows of $\mathbf{B}$. If $\vec{z}$ is random then $\det(\mathbf{C}) \neq 0$ with overwhelming probability, while if $\vec{z} = \mathbf{A}\vec{w}$ for some vector $\vec{w}$ then $\det(\mathbf{C}) = 0$. Therefore the $\mathcal{D}_{\ell,k}$-Matrix Diffie-Hellman Assumption is false in this case.

Otherwise $\det(\mathbf{A}_0) = 0$. Then $\operatorname{rank}(\mathbf{A}_0||\vec{z}_0) = \operatorname{rank}(\mathbf{A}_0)$ when $\vec{z} = \mathbf{A}\vec{w}$, while $\operatorname{rank}(\mathbf{A}_0||\vec{z}_0) = \operatorname{rank}(\mathbf{A}_0) + 1$ with overwhelming probability if $\vec{z}$ is random. To compute the rank of both matrices the following efficient randomized algorithm can be used. Take random invertible matrices $\mathbf{L}, \mathbf{R} \in \mathbb{Z}_q^{k \times k}$. Then set $[\mathbf{A}'_0] = [\mathbf{L}\mathbf{A}_0\mathbf{R}]$ and $[\vec{z}'_0] = [\mathbf{L}\vec{z}_0]$, which is just a randomized instance of the same problem. Now if $\operatorname{rank}(\mathbf{A}'_0) = r$ then with overwhelming probability its principal $r \times r$ minor is nonzero. Therefore, we can estimate $r = \operatorname{rank}(\mathbf{A}'_0)$ as the size of the largest nonzero principal minor (with negligible error probability). Finally, if the determinant of the submatrix of $\mathbf{A}'_0||\vec{z}'_0$ formed by the first $r+1$ rows and the first r and the last column is nonzero we conclude that $\vec{z}$ is random. □

3.3 Generic Hardness of Matrix DH

Let $\mathcal{D}_{\ell,k}$ be a matrix distribution as in Definition 1, which outputs matrices $\mathbf{A} \in \mathbb{Z}_q^{\ell \times k}$. We call $\mathcal{D}_{\ell,k}$ *polynomial-induced* if the distribution is defined by picking $\vec{t} \in \mathbb{Z}_q^d$ uniformly at random and setting $a_{i,j} := \mathfrak{p}_{i,j}(\vec{t})$ for some polynomials $\mathfrak{p}_{i,j} \in \mathbb{Z}_q[\vec{T}]$ whose degree does not depend on λ. E.g. for 2-Lin from Section 1.1, we have $a_{1,1} = t_1, a_{2,2} = t_2, a_{2,1} = a_{3,2} = 1$ and $a_{1,2} = a_{3,1} = 0$ with t_1, t_2 (called a_1, a_2 in Section 1.1) uniform.

We set $\mathfrak{f}_{i,j} = A_{i,j} - \mathfrak{p}_{i,j}$ and $\mathfrak{g}_i = Z_i - \sum_j \mathfrak{p}_{i,j} W_j$ in the ring $\mathcal{R} = \mathbb{Z}_q[A_{1,1}, \ldots, A_{\ell,k}, \vec{Z}, \vec{T}, \vec{W}]$. Consider the ideal $\mathcal{I}_0$ generated by all $\mathfrak{f}_{i,j}$'s and $\mathfrak{g}_i$'s and the ideal $\mathcal{I}_1$ generated only by the $\mathfrak{f}_{i,j}$'s in $\mathcal{R}$. Let $\mathcal{J}_b := \mathcal{I}_b \cap \mathbb{Z}_q[A_{1,1}, \ldots, A_{\ell,k}, \vec{Z}]$. Note that the equations $\mathfrak{f}_{i,j} = 0$ just encode the definition of the matrix entry $a_{i,j}$ by $\mathfrak{p}_{i,j}(\vec{t})$ and the equation $\mathfrak{g}_i = 0$ encodes the definition of z_i in the case $\vec{z} = \mathbf{A}\vec{\omega}$. So, informally, $\mathcal{I}_0$ encodes the relations between the $a_{i,j}$'s, z_i's, t_i's and w_i's in $([\mathbf{A}], [\vec{z}] = [\mathbf{A}\vec{\omega}])$ and $\mathcal{I}_1$ encodes the relations in $([\mathbf{A}], [\vec{z}] = [\vec{u}])$. For $b = 0$ ($\vec{z} = \mathbf{A}\vec{\omega}$) and $b = 1$ ($\vec{z}$ uniform), $\mathcal{J}_b$ encodes the relations visible by considering only the given data (i.e. the $A_{i,j}$'s and Z_j's).

Theorem 3. *Let $\mathcal{D}_{\ell,k}$ be a polynomial-induced matrix distribution with notation as above. Then the $\mathcal{D}_{\ell,k}$-*MDDH *assumption holds in generic m-linear groups if and only if $(\mathcal{J}_0)_{\leq m} = (\mathcal{J}_1)_{\leq m}$, where the $\leq m$ means restriction to total degree at most m.*

Proof. Note that $\mathcal{J}_{\leq m}$ captures precisely what any adversary can generically compute with polynomially many group and m-linear pairing operations. Formally, this is proven by restating the Uber-Assumption Theorem of [2, 9] and its proof more algebraically. Cf. Appendix B for details. □

For a given matrix distribution, the condition $(\mathcal{J}_0)_{\leq m} = (\mathcal{J}_1)_{\leq m}$ can be verified by direct linear algebra or by elimination theory (using e.g. Gröbner bases).[3] For the special case $\ell = k+1$, we can actually give a criterion that is simple to verify using determinants:

Theorem 4. *Let $\mathcal{D}_k$ be a polynomial-induced matrix distribution, which outputs matrices $a_{i,j} = \mathfrak{p}_{i,j}(\vec{t})$ for uniform $\vec{t} \in \mathbb{Z}_q^d$. Let $\mathfrak{d}$ be the determinant of $(\mathfrak{p}_{i,j}(\vec{T}) \| \vec{Z})$ as a polynomial in $\vec{Z}, \vec{T}$.*

1. *If the matrices output by $\mathcal{D}_k$ always have full rank (not just with overwhelming probability), even for t_i from the algebraic closure $\overline{\mathbb{Z}_q}$, then $\mathfrak{d}$ is irreducible over $\overline{\mathbb{Z}_q}$.*

[3] see Lem. 20 in Appendix B

2. *If all $\mathfrak{p}_{i,j}$ have degree at most one and $\mathfrak{d}$ is irreducible over $\overline{\mathbb{Z}_q}$ and the total degree of $\mathfrak{d}$ is $k+1$, then the $\mathcal{D}_k$-*MDDH *assumption holds in* generic k*-linear groups.*

This theorem and generalizations for non-linear $\mathfrak{p}_{i,j}$ and non-irreducible $\mathfrak{d}$ are proven in Appendix B using tools from algebraic geometry.

3.4 Examples of $\mathcal{D}_{\ell,k}$-MDDH

Let $\mathcal{D}_{\ell,k}$ be a matrix distribution and $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$. Looking ahead to our applications, $[\mathbf{A}]$ will correspond to the public-key (or common reference string) and $[\mathbf{A}\vec{w}] \in \mathbb{G}^\ell$ will correspond to a ciphertext. We define the *representation size* $\mathsf{RE}_{\mathbb{G}}(\mathcal{D}_{\ell,k})$ of a given polynomial-induced matrix distribution $\mathcal{D}_{\ell,k}$ with linear $\mathfrak{p}_{i,j}$'s as the minimal number of group elements it takes to represent $[\mathbf{A}]$ for any $\mathbf{A} \in \mathcal{D}_{\ell,k}$. We will be interested in families of distributions $\mathcal{D}_{\ell,k}$ such that that Matrix Diffie-Hellman Assumption is hard in k-linear groups. By Lemma 2 we obtain a family of strictly weaker assumptions. Our goal is to obtain such a family of assumptions with small (possibly minimal) representation.

Example 1. *Let $\mathcal{U}_{\ell,k}$ be the uniform distribution over $\mathbb{Z}_q^{\ell\times k}$.*

The next lemma says that $\mathcal{U}_{\ell,k}$-MDDH is the weakest possible assumption among all $\mathcal{D}_{\ell,k}$-Matrix Diffie-Hellman Assumptions. However, $\mathcal{U}_{\ell,k}$ has poor representation, i.e., $\mathsf{RE}_{\mathbb{G}}(\mathcal{U}_{\ell,k}) = \ell k$.

Lemma 5. *Let $\mathcal{D}_{\ell,k}$ be any matrix distribution. Then $\mathcal{D}_{\ell,k}$-*MDDH $\Rightarrow \mathcal{U}_{\ell,k}$-MDDH*.*

Proof. Given an instance $([\mathbf{A}], [\mathbf{A}\vec{w}])$ of $\mathcal{D}_{\ell,k}$, if $\mathbf{L} \in \mathbb{Z}_q^{\ell\times\ell}$ and $\mathbf{R} \in \mathbb{Z}_q^{k\times k}$ are two random invertible matrices, it is possible to get a properly distributed instance of the $\mathcal{U}_{\ell,k}$-matrix DH problem as $([\mathbf{LAR}], [\mathbf{LA}\vec{w}])$. Indeed, $\mathbf{LAR}$ has a distribution statistically close to the uniform distribution[4] in $\mathbb{Z}_q^{k\times\ell}$, while $\mathbf{LA}\vec{w} = \mathbf{LAR}\vec{v}$ for $\vec{v} = \mathbf{R}^{-1}\vec{w}$. Clearly, $\vec{v}$ has the uniform distribution in $\mathbb{Z}_q^k$. □

Example 2 (k-Linear Assumption/k-Lin)**.** *We define the distribution $\mathcal{L}_k$ as follows*

$$\mathbf{A} = \begin{pmatrix} a_1 & 0 & \dots & 0 & 0 \\ 0 & a_2 & \dots & 0 & 0 \\ 0 & 0 & & \ddots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & a_k \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} \in \mathbb{Z}_q^{(k+1)\times k},$$

where $a_i \leftarrow \mathbb{Z}_q^*$. The transformation matrix $\mathbf{T} \in \mathbb{Z}_q^{1\times k}$ is given as $\mathbf{T} = (\frac{1}{a_1}, \dots, \frac{1}{a_k})$. Note that the distribution $(\mathbf{A}, \mathbf{A}\vec{w})$ can be compactly written as $(a_1, \dots, a_k, a_1w_1, \dots, a_kw_k, w_1 + \dots + w_k) = (a_1, \dots, a_k, b_1, \dots, b_k, \frac{b_1}{a_1} + \dots + \frac{b_k}{a_k})$ with $a_i \leftarrow \mathbb{Z}_q^*$, $b_i, w_i \leftarrow \mathbb{Z}_q$. Hence the $\mathcal{L}_k$-Matrix Diffie-Hellman Assumption is an equivalent description of the k-linear Assumption [3, 23, 45] with $\mathsf{RE}_{\mathbb{G}}(\mathcal{L}_k) = k$.

It was shown in [45] that k-Lin holds in the generic k-linear group model and hence k-Lin forms a family of increasingly strictly weaker assumptions. Furthermore, in [7] it was shown that 2-Lin $\Rightarrow$ BDDH.

Example 3 (k-Cascade Assumption/k-Casc)**.** *We define the distribution $\mathcal{C}_k$ as follows*

$$\mathbf{A} = \begin{pmatrix} a_1 & 0 & \dots & 0 & 0 \\ 1 & a_2 & \dots & 0 & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & a_k \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

[4] If $\mathbf{A}$ has full-rank (that happens with overwhelming probability) then $\mathbf{LAR}$ is uniformly distributed in the set of full-rank matrices in $\mathbb{Z}_q^{\ell\times k}$, which implies that it is close to uniform in $\mathbb{Z}_q^{\ell\times k}$.

where $a_i \leftarrow \mathbb{Z}_q^*$. *The transformation matrix* $\mathbf{T} \in \mathbb{Z}_q^{1\times k}$ *is given as* $\mathbf{T} = (\pm\frac{1}{a_1\cdots a_k}, \mp\frac{1}{a_2\cdots a_k} \ldots, \frac{1}{a_k})$. *Note that* $(\mathbf{A}, \mathbf{A}\vec{w})$ *can be compactly written as* $(a_1, \ldots, a_k, a_1w_1, w_1+a_2w_2 \ldots, w_{k-1}+a_kw_k, w_k) = (a_1, \ldots, a_k, b_1, \ldots, b_k, \frac{b_k}{a_k} - \frac{b_{k-1}}{a_{k-1}a_k} + \frac{b_{k-2}}{a_{k-2}a_{k-1}a_k} - \ldots \pm \frac{b_1}{a_1\cdots a_k})$. *We have* $\mathsf{RE}_{\mathbb{G}}(\mathcal{C}_k) = k$.

Matrix $\mathbf{A}$ bears resemblance to a cascade which explains the assumption's name. Indeed, in order to compute the right lower entry w_k of matrix $(\mathbf{A}, \mathbf{A}\vec{w})$ from the remaining entries, one has to "descend" the cascade to compute all the other entries w_i $(1 \le i \le k-1)$ one after the other.

A more compact version of $\mathcal{C}_k$ is obtained by setting all $a_i := a$.

Example 4. *(Symmetric k-Cascade Assumption) We define the distribution* $\mathcal{SC}_k$ *as* $\mathcal{C}_k$ *but now* $a_i = a$, *where* $a \leftarrow \mathbb{Z}_q^*$. *Then* $(\mathbf{A}, \mathbf{A}\vec{w})$ *can be compactly written as* $(a, aw_1, w_1 + aw_2, \ldots, w_{k-1} + aw_k, w_k) = (a, b_1, \ldots, b_k, \frac{b_k}{a} - \frac{b_{k-1}}{a^2} + \frac{b_{k-2}}{a^3} - \ldots \pm \frac{b_1}{a^k})$. *We have* $\mathsf{RE}_{\mathbb{G}}(\mathcal{C}_k) = 1$.

Observe that the same trick cannot be applied to the k-Linear assumption k-Lin, as the resulting Symmetric k-Linear assumption does not hold in k-linear groups. However, if we set $a_i := a + i - 1$, we obtain another matrix distribution with compact representation.

Example 5. *(Incremental k-Linear Assumption) We define the distribution* $\mathcal{IL}_k$ *as* $\mathcal{L}_k$ *with* $a_i = a + i - 1$, *for* $a \leftarrow \mathbb{Z}_q^*$. *The transformation matrix* $\mathbf{T} \in \mathbb{Z}_q^{1\times k}$ *is given as* $\mathbf{T} = (\frac{1}{a}, \ldots, \frac{1}{a+k-1})$. $(\mathbf{A}, \mathbf{A}\vec{w})$ *can be compactly written as* $(a, aw_1, (a+1)w_2, \ldots, (a+k-1)w_k, w_1 + \ldots + w_k) = (a, b_1, \ldots, b_k, \frac{b_1}{a} + \frac{b_2}{a+1} + \ldots + \frac{b_k}{a+k-1})$. *We also have* $\mathsf{RE}_{\mathbb{G}}(\mathcal{IL}_k) = 1$.

The last three examples need some work to prove its generic hardness.

Theorem 6. k-Casc, k-SCasc *and* k-ILin *are hard in generic* k*-linear groups.*

Proof. We need to consider the (statistically close) variants with $a_i \in \mathbb{Z}_q$ rather that $\mathbb{Z}_q^*$. The determinant polynomial for $\mathcal{C}_k$ is $\mathfrak{d}(a_1, \ldots, a_k, z_1, \ldots, z_{k+1}) = a_1 \cdots a_k z_{k+1} - a_1 \cdots a_{k-1} z_k + \ldots + (-1)^k z_1$, which has total degree $k+1$. As all matrices in $\mathcal{C}_k$ have rank k, because the determinant of the last k rows in $\mathbf{A}$ is always 1, by Theorem 4 we conclude that k-Casc is hard in k-linear groups. As $\mathcal{SC}_k$ is a particular case of $\mathcal{C}_k$, the determinant polynomial for $\mathcal{SC}_k$ is $\mathfrak{d}(a, z_1, \ldots, z_{k+1}) = a^k z_{k+1} - a^{k-1} z_k + \ldots + (-1)^k z_1$. As before, by Theorem 4, k-SCasc is hard in k-linear groups. Finally, in the case of k-ILin we will show in the next section its equivalence to k-SCasc and therefore it is generically hard in k-linear groups.

□

The previous examples can be related to some known assumptions from Section 2.3. Figure 1 depicts the relations that are also stated in next theorem, except the equivalence of k-ILin and k-SCasc which is addressed in the next section. We stress that this equivalence together with Theorem 7 imply that k-SCasc is a stronger assumption than k-Lin, previously unknown [16].

Theorem 7. *For any* $k \ge 2$, *the following holds:*

$$(k+1)\text{-}\mathsf{PDDH} \Rightarrow k\text{-}\mathsf{Casc}\ ;$$
$$(k+1)\text{-}\mathsf{EDDH} \Rightarrow k\text{-}\mathsf{SCasc} \Rightarrow k\text{-}\mathsf{Casc}\ ; \qquad k\text{-}\mathsf{ILin} \Rightarrow k\text{-}\mathsf{Lin}\ ;$$
$$k\text{-}\mathsf{Casc} \Rightarrow (k+1)\text{-}\mathsf{Casc}\ ; \qquad k\text{-}\mathsf{SCasc} \Rightarrow (k+1)\text{-}\mathsf{SCasc}$$

Further, in k*-linear groups,* k-$\mathsf{Casc} \Rightarrow k$-$\mathsf{MLDDH}$.

Proof. The proof of all implications can be found in Appendix A. □

4 Uniqueness of One-Parameter Matrix DH Problems

Some differently-looking MDDH assumptions can be tightly equivalent, or isomorphic, meaning that there is a very tight generic reduction between the corresponding problems. These reductions are mainly based on the algebraic nature of the MDDH problems.

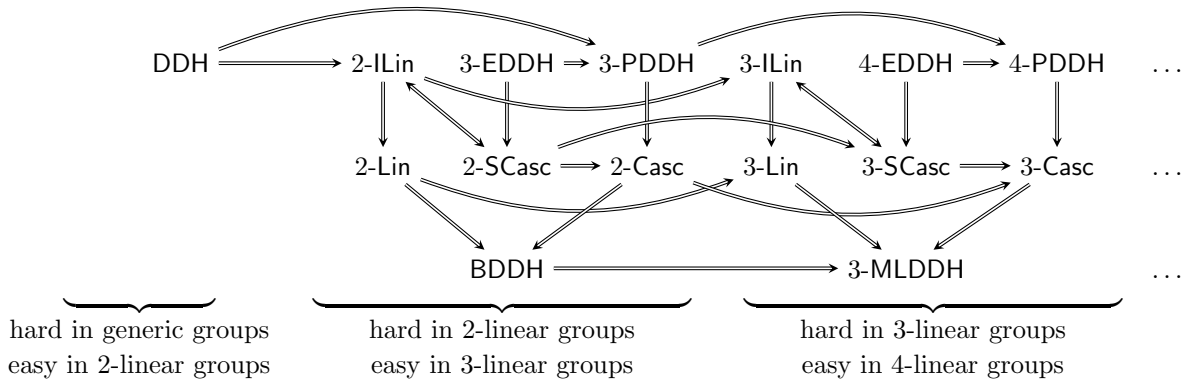

Figure 1: Relation between various assumptions and their generic hardness in k-linear groups.

The simplest and most compact polynomial-induced matrix distributions $\mathcal{D}_k$ are the one-parameter linear ones, where $\mathcal{D}_k$ outputs matrices $\mathbf{A}(t) = \mathbf{A}_0 + \mathbf{A}_1 t$ for a uniformly distributed $t \in \mathbb{Z}_q$, and fixed $\mathbf{A}_0, \mathbf{A}_1 \in \mathbb{Z}_q^{(k+1)\times k}$. The two examples of them given in [16] are $\mathcal{SC}_k$ and $\mathcal{IL}_k$.

A natural question is whether such a tight algebraic reduction exists between $\mathcal{SC}_k$ and $\mathcal{IL}_k$. In this section we prove a much stronger result, which states there exists essentially a single one-parameter linear MDDH problem. Indeed, we show that all one-parameter linear $\mathcal{D}_k$-MDDH problems are isomorphic to $\mathcal{SC}_k$. This result is heavily related to the one-parameter nature of the problems considered, and it seems to be not generalizable to broader families of MDDH problems (e.g., trying to relate $\mathcal{C}_k$ and $\mathcal{L}_k$, or dealing with the case $\ell > k+1$).

4.1 Hardness

Theorem 4 gives an easy-to-check sufficient condition ensuring the $\mathcal{D}_k$-MDDH assumption holds in *generic* k-linear groups for certain matrix distributions $\mathcal{D}_k$, including the one-parameter linear ones. For this particular family, the sufficient condition is that all matrices $\mathbf{A}(t) = \mathbf{A}_0 + \mathbf{A}_1 t$ have full-rank for all $t \in \overline{\mathbb{Z}_q}$, the algebraic closure of the finite field $\mathbb{Z}_q$, and the determinant $\mathfrak{d}$ of $(\mathbf{A}(T)\|\vec{Z})$ as a polynomial in $\vec{Z}, T$ has total degree $k+1$. We first show that indeed it is also a necessary condition for the hardness of the $\mathcal{D}_k$-MDDH problem.

Theorem 8. *Let $\mathcal{D}_k$ be a one-parameter linear matrix distribution, producing matrices $\mathbf{A}(t) = \mathbf{A}_0 + \mathbf{A}_1 t$, such that $\mathcal{D}_k$-MDDH assumption is hard generically in k-linear groups. Then, the determinant $\mathfrak{d}$ of $(\mathbf{A}(T)\|\vec{Z})$ is an irreducible polynomial in $\overline{\mathbb{Z}_q}[\vec{Z}, T]$ with total degree $k+1$, and the rank of $\mathbf{A}_0 + \mathbf{A}_1 t$ is always k, for all $t \in \overline{\mathbb{Z}_q}$.*

Proof. The proof just consists in finding a nonzero polynomial $\mathfrak{h} \in \mathbb{Z}_q[\vec{Z}, T]$ of degree at most k such that $\mathfrak{h}(\mathbf{A}(t)\vec{w}, t) = 0$ for all $t \in \mathbb{Z}_q$ and $\vec{w} \in \mathbb{Z}_q^k$, and then using it to solve the $\mathcal{D}_k$-MDDH problem. If the total degree of $\mathfrak{d}$ is at most k, then we can simply let $\mathfrak{h} = \mathfrak{d}$[5]. Otherwise, assume that the degree of $\mathfrak{d}$ is $k+1$. If $\mathfrak{d}$ is reducible, from Lemma 21 it follows that $\mathfrak{d}$ can be split as $\mathfrak{d} = \mathfrak{c}\mathfrak{d}_0$, where $\mathfrak{c} \in \mathbb{Z}_q[T]$ and $\mathfrak{d}_0 \in \mathbb{Z}_q[\vec{Z}, T]$ are nonconstant. Clearly, if $\mathfrak{c}(t) \neq 0$ then $\mathfrak{d}_0(\mathbf{A}(t)\vec{w}, t) = 0$ for all $\vec{w} \in \mathbb{Z}_q^k$, which means that as a polynomial in $\mathbb{Z}_q[\vec{W}, T]$, $\mathfrak{d}_0(T, \mathbf{A}(T)\vec{W})$ has too many roots, so it is the zero polynomial. Therefore, we are done by taking $\mathfrak{h} = \mathfrak{d}_0$.

Finally, observe that $\mathfrak{d}(\vec{z}, t) = \sum_{i=0}^{k+1} \mathfrak{c}_i(t) z_i$, where $\vec{z} = (z_1, \ldots, z_{k+1})$ and the $\mathfrak{c}_i(t)$ are the (signed) k-minors of $\mathbf{A}(t)$. Therefore, if $\mathbf{A}(t_0)$ has rank less than k for some $t_0 \in \overline{\mathbb{Z}_q}$ then $\mathfrak{d}(\vec{z}, t_0) = 0$ for all $\vec{z} \in \mathbb{Z}_q^{k+1}$,

[5]Actually, it is assumed that $\mathfrak{d} \neq 0$, i.e., some matrices output by $\mathcal{D}_k$ have full-rank. Otherwise, it is not hard finding the polynomial $\mathfrak{h}$ based on a nonzero maximal minor of $\mathbf{A}(t)$, by adding to it an extra row and the column $\vec{Z}$.

which means that $\mathfrak{c}_i(t_0) = 0$ for all i. As a consequence, $T - t_0$ divides all $\mathfrak{c}_i$ and hence it divides $\mathfrak{d}$, that is, $\mathfrak{d}$ is reducible.

Once we have found the polynomial $\mathfrak{h}$ of degree at most k, an efficient distinguisher can use the k-linear map to evaluate $[\mathfrak{h}(\vec{z}, t)]_{T_k}$ from an instance $([\mathbf{A}(t)], [\vec{z}])$ of the $\mathcal{D}_k$-MDDH problem, where $[t]$ can be computed easily from $[\mathbf{A}(t)]$ because $\mathbf{A}_0$ and $\mathbf{A}_1$ are known. If $\vec{z} = \mathbf{A}(t)\vec{w}$ then $\mathfrak{h}(\vec{z}, t) = 0$, while for a randomly chosen $\vec{z}$, $\mathfrak{h}(\vec{z}, t) \neq 0$ with overwhelming probability[6]. Then the distinguisher succeeds with an overwhelming probability. □

4.2 Isomorphic Problems

From now on, we consider in this section a one-parameter linear matrix distribution $\mathcal{D}_k$ such that $\mathcal{D}_k$-MDDH assumption holds in generic k-linear groups. This in particular means that using Theorem 8, the polynomial $\mathfrak{d}$ is irreducible in $\overline{\mathbb{Z}_q}[\vec{Z}, T]$ with total degree $k+1$, and that the rank of $\mathbf{A}_0 + \mathbf{A}_1 t$ is always k, for all $t \in \overline{\mathbb{Z}_q}$. Clearly the rank of $\mathbf{A}_0$ is k, but also $\mathbf{A}_1$ has rank k. Indeed, it is easy to see that the coefficients of the monomials of degree $k+1$ in $\mathfrak{d}$ are exactly the (signed) k-minors of $\mathbf{A}_1$, so they cannot be all zero.

There are some natural families of maps that generically transform MDDH problems into MDDH problems. As mentioned in previous sections, some examples of them are left and right multiplication by an invertible constant matrix. More precisely, let $\mathbf{L} \in GL_{k+1}(\mathbb{Z}_q)$, the set of all invertible matrices in $\mathbb{Z}_q^{(k+1)\times(k+1)}$, and $\mathbf{R} \in GL_k(\mathbb{Z}_q)$. Given some matrix distribution $\mathcal{D}_k$, we write $\mathcal{D}'_k = \mathbf{L}\mathcal{D}_k\mathbf{R}$ to denote the matrix distribution resulting from sampling a matrix from $\mathcal{D}_k$ and multiplying on the left and on the right by $\mathbf{L}$ and $\mathbf{R}$.

This mapping between matrix distrutions can be used to transform any distinguisher for $\mathcal{D}'_k$-MDDH into a distinguisher for $\mathcal{D}_k$-MDDH with the same advantage and essentially the same running time. Indeed, a 'real' instance $([\mathbf{A}], [\mathbf{A}\vec{w}])$ of a MDDH problem can be transformed into a 'real' instance of the other MDDH problem $([\mathbf{A}'], [\mathbf{A}'\vec{w}']) = (\mathbf{L}[\mathbf{A}]\mathbf{R}, \mathbf{L}[\mathbf{A}\vec{w}])$ with the right distribution, because $\mathbf{L}\mathbf{A}\vec{w} = \mathbf{A}'\vec{w}'$, where $\vec{w}' = \mathbf{R}^{-1}\vec{w}$ is uniformly distributed. Similarly, a 'random' instance $([\mathbf{A}], [\vec{z}])$ is transformed into another one $([\mathbf{A}'], [\vec{z}']) = (\mathbf{L}[\mathbf{A}]\mathbf{R}, \mathbf{L}[\vec{z}])$. From an algebraic point of view, we can see the above transformation as changing the bases used to represent certain linear maps as matrices.

In the particular case of one-parameter linear matrix distributions, one can write $\mathbf{A}'(t) = \mathbf{L}\mathbf{A}(t)\mathbf{R} = \mathbf{L}\mathbf{A}_0\mathbf{R} + \mathbf{L}\mathbf{A}_1\mathbf{R}t$, which simply means defining $\mathbf{A}'_0 = \mathbf{L}\mathbf{A}_0\mathbf{R}$ and $\mathbf{A}'_1 = \mathbf{L}\mathbf{A}_1\mathbf{R}$. Consider the injective linear maps $f_0, f_1 : \mathbb{Z}_q^k \to \mathbb{Z}_q^{k+1}$ defined by $f_0(\vec{w}) = \mathbf{A_0}\vec{w}$ and $f_1(\vec{w}) = \mathbf{A_1}\vec{w}$. We need the following technical lemma.

Lemma 9. *If $\mathcal{D}_k$ is generically hard in k-linear groups, no nontrivial subspace $U \subset \mathbb{Z}_q^k$ exists such that $f_0(U) = f_1(U)$.*

Proof. Assume for contradiction a nontrivial subspace U exists such that $f_0(U) = f_1(U)$, and consider the natural automorphism $\phi : U \to U$ defined as $\phi = f_1^{-1} \circ f_0$. It is well defined due to the injectivity of f_0 and f_1. Then, there exists an eigenvector $\vec{v} \neq \vec{0}$ of ϕ for some eigenvalue $\lambda \in \overline{\mathbb{Z}_q}$. The equation $\phi(\vec{v}) = f_1^{-1} \circ f_0(\vec{v}) = \lambda\vec{v}$ implies $(f_0 - \lambda f_1)(\vec{v}) = \vec{0}$. Therefore, $f_0 - \lambda f_1$ is no longer injective and $\mathbf{A}(-\lambda) = \mathbf{A}_0 - \lambda\mathbf{A}_1$ has rank strictly less than k, which contradicts Theorem 8. □

Applying the lemma iteratively one can build special bases for the spaces $\mathbb{Z}_q^k$ and $\mathbb{Z}_q^{k+1}$, and obtain canonical forms simultaneously for $\mathbf{A}_0$ and $\mathbf{A}_1$, as described in the proof of the following theorem, which has some resemblance to the construction of Jordan normal forms of endomorphisms. The proof is rather technical, and it can be found in Appendix C.

Theorem 10. *Let $f_0, f_1 : \mathbb{Z}_q^k \to \mathbb{Z}_q^{k+1}$ two injective linear maps such that $f_0(U) \neq f_1(U)$ for any nontrivial subspace $U \subset \mathbb{Z}_q^k$. There exist bases of $\mathbb{Z}_q^k$ and $\mathbb{Z}_q^{k+1}$ such that f_0 and f_1 are represented in those bases*

[6] As a polynomial of total degree at most k it vanishes with probability at most k/q at a uniformly distributed point.

respectively by the matrices

$$\mathbf{J}_0 = \begin{pmatrix} 0 & \cdots & 0 \\ 1 & \ddots & \vdots \\ \vdots & \ddots & 0 \\ 0 & \cdots & 1 \end{pmatrix} \qquad \mathbf{J}_1 = \begin{pmatrix} 1 & \cdots & 0 \\ 0 & \ddots & \vdots \\ \vdots & \ddots & 1 \\ 0 & \cdots & 0 \end{pmatrix}$$

Corollary 1. *All one-parameter linear hard $\mathcal{D}_k$-MDDH problems are isomorphic to the $\mathcal{SC}_k$-MDDH problem, i.e., there exist invertible matrices $\mathbf{L} \in GL_{k+1}(\mathbb{Z}_q)$ and $\mathbf{R} \in GL_k(\mathbb{Z}_q)$ such that $\mathcal{D}_k = \mathbf{L}\mathcal{SC}_k\mathbf{R}$.*

Proof. Combining the previous results, the maps f_0, f_1 defined from the hard $\mathcal{D}_k$-MDDH problem are injective and they can be represented in the bases given in Theorem 10. In terms of matrices this means that there exist $\mathbf{L} \in GL_{k+1}(\mathbb{Z}_q)$ and $\mathbf{R} \in GL_k(\mathbb{Z}_q)$ such that $\mathbf{A}_0 = \mathbf{L}\mathbf{J}_0\mathbf{R}$ and $\mathbf{A}_1 = \mathbf{L}\mathbf{J}_1\mathbf{R}$, that is,

$$\mathbf{A}(t) = \mathbf{L} \begin{pmatrix} t & \cdots & 0 \\ 1 & \ddots & \vdots \\ \vdots & \ddots & t \\ 0 & \cdots & 1 \end{pmatrix} \mathbf{R}$$

which concludes the proof. □

As an example, we show an explicit isomorphism between $\mathcal{SC}_2$-MDDH and $\mathcal{IL}_2$-MDDH problems.

$$\begin{pmatrix} t & 0 \\ 0 & t+1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t & 0 \\ 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}$$

We stress that 'isomorphic' does not mean 'identical', and it is still useful having at hand different representations of essentially the same computational problem, as it would help finding applications.

5 Basic Applications

5.1 Public-Key Encryption

Let Gen be a group generating algorithm and $\mathcal{D}_{\ell,k}$ be a matrix distribution that outputs a matrix over $\mathbb{Z}_q^{\ell\times k}$ such that the first k-rows form an invertible matrix with overwhelming probability. We define the following key-encapsulation mechanism $\mathsf{KEM}_{\mathsf{Gen},\mathcal{D}_{\ell,k}} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ with key-space $\mathcal{K} = \mathbb{G}^{\ell-k}$.

- $\mathsf{Gen}(1^\lambda)$ runs $\mathcal{G} \leftarrow \mathsf{Gen}(1^\lambda)$ and $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$. Let $\mathbf{A}_0$ be the first k rows of $\mathbf{A}$ and $\mathbf{A}_1$ be the last $\ell - k$ rows of $\mathbf{A}$. Define $\mathbf{T} \in \mathbb{Z}_q^{(\ell-k)\times k}$ as the transformation matrix $\mathbf{T} = \mathbf{A}_1\mathbf{A}_0^{-1}$. The public/secret-key is

$$pk = (\mathcal{G}, [\mathbf{A}] \in \mathbb{G}^{\ell\times k}), \quad sk = (pk, \mathbf{T} \in \mathbb{Z}_q^{(\ell-k)\times k})$$

- Enc_{pk} picks $\vec{w} \leftarrow \mathbb{Z}_q^k$. The ciphertext/key pair is

$$[\vec{c}] = [\mathbf{A}_0\vec{w}] \in \mathbb{G}^k, \quad [K] = [\mathbf{A}_1\vec{w}] \in \mathbb{G}^{\ell-k}$$

- $\mathsf{Dec}_{sk}([\vec{c}] \in \mathbb{G}^k)$ recomputes the key as $[K] = [\mathbf{T}\vec{c}] \in \mathbb{G}^{\ell-k}$.

Correctness follows by the equation $\mathbf{T} \cdot \vec{c} = \mathbf{T} \cdot \mathbf{A}_0\vec{w} = \mathbf{A}_1\vec{w}$. The public key contains $\mathsf{RE}_{\mathbb{G}}(\mathcal{D}_{\ell,k})$ and the ciphertext k group elements. An example scheme from the k-SCasc Assumption is given in Appendix E.1.

Theorem 11. *Under the $\mathcal{D}_{\ell,k}$-MDDH Assumption $\mathsf{KEM}_{\mathsf{Gen},\mathcal{D}_{\ell,k}}$ is IND-CPA secure.*

Proof. By the $\mathcal{D}_{\ell,k}$ Matrix Diffie-Hellman Assumption, the distribution of $(pk, [\vec{c}], [K]) = ((\mathcal{G}, [\mathbf{A}]), [\mathbf{A}\vec{w}])$ is computationally indistinguishable from $((\mathcal{G}, [\mathbf{A}]), [\vec{u}])$, where $\vec{u} \leftarrow \mathbb{Z}_q^\ell$. □

5.2 Hash Proof Systems

Let $\mathcal{D}_{\ell,k}$ be a matrix distribution. We build a universal$_1$ hash proof system $\mathsf{HPS} = (\mathsf{Param}, \mathsf{Pub}, \mathsf{Priv})$, whose hard subset membership problem is based on the $\mathcal{D}_{\ell,k}$ Matrix Diffie-Hellman Assumption.

- $\mathsf{Param}(1^\lambda)$ runs $\mathcal{G} \leftarrow \mathsf{Gen}(1^\lambda)$ and picks $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$. Define the language

$$\mathcal{V} = \mathcal{V}_{\mathbf{A}} = \{[\vec{c}] = [\mathbf{A}\vec{w}] \in \mathbb{G}^\ell \ : \ \vec{w} \in \mathbb{Z}_q^k\} \subseteq \mathcal{C} = \mathbb{G}^\ell.$$

  The value $\vec{w} \in \mathbb{Z}_q^k$ is a witness of $[\vec{c}] \in \mathcal{V}$. Let $\mathcal{SK} = \mathbb{Z}_q^\ell$, $\mathcal{PK} = \mathbb{G}^k$, and $\mathcal{K} = \mathbb{G}$. For $sk = \vec{x} \in \mathbb{Z}_q^\ell$, define the projection $\mu(sk) = [\vec{x}^\top \mathbf{A}] \in \mathbb{G}^k$. For $[\vec{c}] \in \mathcal{C}$ and $sk \in \mathcal{SK}$ we define

$$\Lambda_{sk}([\vec{c}]) := [\vec{x}^\top \cdot \vec{c}] \ . \tag{3}$$

  The output of Param is $params = \big(\mathcal{S} = (\mathcal{G}, [\mathbf{A}]), \mathcal{K}, \mathcal{C}, \mathcal{V}, \mathcal{PK}, \mathcal{SK}, \Lambda_{(\cdot)}(\cdot), \mu(\cdot)\big)$.

- $\mathsf{Priv}(sk, [\vec{c}])$ computes $[K] = \Lambda_{sk}([\vec{c}])$.

- $\mathsf{Pub}(pk, [\vec{c}], \vec{w})$. Given $pk = \mu(sk) = [\vec{x}^\top \mathbf{A}]$, $[\vec{c}] \in \mathcal{V}$ and a witness $\vec{w} \in \mathbb{Z}_q^k$ such that $[\vec{c}] = [\mathbf{A} \cdot \vec{w}]$ the public evaluation algorithm $\mathsf{Pub}(pk, [\vec{c}], \vec{w})$ computes $[K] = \Lambda_{sk}([\vec{c}])$ as

$$[K] = [(\vec{x}^\top \cdot \mathbf{A}) \cdot \vec{w}] \ .$$

Correctness follows by (3) and the definition of μ. Clearly, under the $\mathcal{D}_{\ell,k}$-Matrix Diffie-Hellman Assumption, the subset membership problem is hard in HPS.

We now show that Λ is a universal$_1$ projective hash function. Let $[\vec{c}] \in \mathcal{C} \setminus \mathcal{V}$ be an element outside of the language. Then the matrix $(\mathbf{A}||\vec{c}) \in \mathbb{Z}_q^{\ell \times (k+1)}$ is of full rank $k+1$ and consequently $(\vec{x}^\top \cdot \mathbf{A} || \vec{x}^\top \cdot \vec{c}) \equiv (\vec{x}^\top \mathbf{A} || u)$ for $\vec{x} \leftarrow \mathbb{Z}_q^k$ and $u \leftarrow \mathbb{Z}_q$. Hence, $(pk, \Lambda_{sk}([\vec{c}]) = ([\vec{x}^\top \mathbf{A}], [\vec{x}^\top \vec{c}]) \equiv ([\vec{x}^\top \mathbf{A}], [u]) = ([\vec{x}^\top \mathbf{A}], [K])$.

We remark that Λ can be transformed into a universal$_2$ projective hash function by applying a four-wise independent hash function [30]. Alternatively, one can construct a computational version of a universal$_2$ projective hash function as follows. Let $\mathcal{SK} = (\mathbb{Z}_q^\ell)^2$, $\mathcal{PK} = (\mathbb{G}^k)^2$, and $\mathcal{K} = \mathbb{G}$. For $sk = (\vec{x}_1, \vec{x}_2) \in (\mathbb{Z}_q^\ell)^2$, define the projection $\mu(sk) = [\vec{x}_1^\top \mathbf{A}, \vec{x}_2^\top \mathbf{A}] \in (\mathbb{G}^k)^2$. For $[\vec{c}] \in \mathcal{C}$ and $sk \in \mathcal{SK}$, define $\Lambda_{sk}([\vec{c}]) := [(t\vec{x}_1^\top + \vec{x}_2^\top) \cdot \vec{c}]$, where $t = H(\vec{c})$ and $H : \mathcal{C} \to \mathbb{Z}_q$ is a collision-resistant hash function. The corresponding Priv and Pub algorithms are adapted accordingly. It is easy to verify that for all values $[\vec{c}_1], [\vec{c}_2] \in \mathcal{C} \setminus \mathcal{V}$ with $H(\vec{c}_1) \neq H(\vec{c}_2)$, we have $(pk, \Lambda_{sk}([\vec{c}_1], \Lambda_{sk}([\vec{c}_2]) \equiv (pk, [K_1], [K_2])$, for $K_1, K_2 \leftarrow \mathbb{Z}_q$.

5.3 Pseudo-Random Functions

Let Gen be a group generating algorithm and $\mathcal{D}_{\ell,k}$ be a matrix distribution that outputs a matrix over $\mathbb{Z}_q^{\ell \times k}$ such that the first k-rows form an invertible matrix with overwhelming probability. We define the following pseudo-random function $\mathsf{PRF}_{\mathsf{Gen},\mathcal{D}_{\ell,k}} = (\mathsf{Gen}, \mathsf{F})$ with message space $\mathcal{M} = \{0,1\}^n$ and range $\mathcal{R} = \mathbb{G}^k$. For simplicity we assume that $\ell - k$ divides k.

- $\mathsf{Gen}(1^\lambda)$ runs $\mathcal{G} \leftarrow \mathsf{Gen}(1^\lambda)$, $\vec{h} \in \mathbb{Z}_q^k$, and $\mathbf{A}_{i,j} \leftarrow \mathcal{D}_{\ell,k}$ for $i = 1, \ldots, n$ and $j = 1, \ldots, t := k/(\ell - k)$ and computes the transformation matrices $\mathbf{T}_{i,j} \in \mathbb{Z}_q^{(\ell-k)\times k}$ of $\mathbf{A}_{i,j} \in \mathbb{Z}_q^{\ell \times k}$ (cf. Definition 3). For $i = 1, \ldots, n$ define the aggregated transformation matrices

$$\mathbf{T}_i = \begin{pmatrix} \mathbf{T}_{i,1} \\ \vdots \\ \mathbf{T}_{i,t} \end{pmatrix} \in \mathbb{Z}_q^{k \times k}$$

  The key is defined as

$$K = (\mathcal{G}, \vec{h}, \mathbf{T}_1, \ldots, \mathbf{T}_n).$$

- $\mathsf{F}_K(x)$ computes

$$\mathsf{F}_K(x) = \left[\prod_{i:x_i=1} \mathbf{T}_i \cdot \vec{h} \right] \in \mathbb{G}^k.$$

$\mathsf{PRF}_{\mathsf{Gen},\mathcal{L}_k}$ (i.e., setting $\mathcal{D}_{\ell,k} = \mathcal{L}_k$) is the PRF from Lewko and Waters [33]. A more efficient PRF from the k-SCasc Assumption is given in Appendix E.2.

Note that the elements $\mathbf{T}_1, \ldots, \mathbf{T}_t$ of the secret-key consist of the transformation matrices of independently sampled matrices $\mathbf{A}_{i,j}$. Interestingly, for a number of distributions $\mathcal{D}_{\ell,k}$ the distribution of the transformation matrix $\mathbf{T}$ is the same. For example, the transformation matrix for $\mathcal{L}_k$ consists of a uniform row vector, so does the transformation matrix for $\mathcal{C}_k$ and for $\mathcal{U}_{k+1,k}$. Consequently, $\mathsf{PRF}_{\mathsf{Gen},\mathcal{C}_k} = \mathsf{PRF}_{\mathsf{Gen},\mathcal{L}_k} = \mathsf{PRF}_{\mathsf{Gen},\mathcal{U}_{k+1,k}}$ and in light of the theorem below, $\mathsf{PRF}_{\mathsf{Gen},\mathcal{L}_k}$ proposed by Lewko and Waters can also be proved on the $\mathcal{U}_{k+1,k}$-MDDH assumption, the weakest among all MDDH assumptions of matching dimensions.

Theorem 12. *Under the $\mathcal{D}_{\ell,k}$-MDDH Assumption $\mathsf{PRF}_{\mathsf{Gen},\mathcal{D}_{\ell,k}}$ is a secure pseudo-random function.*

The proof is based on the augmented cascade construction of Boneh et al. [6]. Here we give a direct self-contained proof. We first state and prove the following lemma.

Lemma 13. *Let Q be a polynomial. Under the $\mathcal{D}_{\ell,k}$-MDDH Assumption,*

$$\left[\begin{pmatrix} \vec{h}^1 \\ \hat{\mathbf{T}}\vec{h}^1 \end{pmatrix}, \ldots, \begin{pmatrix} \vec{h}^Q \\ \hat{\mathbf{T}}\vec{h}^Q \end{pmatrix} \right] \in \mathbb{G}^{2k \times Q}$$

is computationally indistinguishable from a uniform $[\mathbf{H}] \in \mathbb{G}^{2k \times Q}$, where $\vec{h}^i \leftarrow \mathbb{Z}_q^k$,

$$\hat{\mathbf{T}} = \begin{pmatrix} \hat{\mathbf{T}}_1 \\ \vdots \\ \hat{\mathbf{T}}_t \end{pmatrix} \in \mathbb{Z}_q^{k \times k},$$

and $\hat{\mathbf{T}}_j$ $(1 \le j \le t)$ are the transformation matrices of $\mathbf{A}_j \leftarrow \mathcal{D}_{\ell,k}$.

Proof. By a hybrid argument over $j = 1, \ldots, t$ it is sufficient to show that

$$\left[\begin{pmatrix} \vec{h}^1 \\ \hat{\mathbf{T}}_1\vec{h}^1 \end{pmatrix}, \ldots, \begin{pmatrix} \vec{h}^Q \\ \hat{\mathbf{T}}_1\vec{h}^Q \end{pmatrix} \right] \in \mathbb{G}^{\ell \times Q}$$

is computationally indistinguishable from a uniform $[\mathbf{H}_1] \leftarrow \mathbb{G}^{\ell \times Q}$, i.e., for one single transformation matrix $\hat{\mathbf{T}}_1$ of $\mathbf{A}_1 \leftarrow \mathcal{D}_{\ell,k}$. This in turn follows directly by Lemma 1 (random self-reducibility of $\mathcal{D}_{\ell,k}$-MDDH). Note that the overall loss in the security reduction is $k = t \cdot (\ell - k)$, where the factor t stems from the hybrid argument and the factor $\ell - k$ stems from Lemma 1. □

Proof of Theorem 12. For $x \in \{0,1\}^n$ and $0 \le \mu \le n$, define $\mathsf{suffix}^\mu(x)$ as the μ-th suffix of x, i.e., $\mathsf{suffix}^\mu(x) := (x_{n-\mu+1}, \ldots, x_n)$. We make the convention that $\mathsf{suffix}^0(x) = \varepsilon$, the empty string.

We will use a hybrid argument over n, the bitlength of x. In Hybrid μ $(0 \le \mu \le n)$, let $\mathsf{RF}^\mu : \{0,1\}^\mu \to \mathbb{Z}_q^k$ be a truly random function and define the oracle

$$\mathcal{O}^\mu(x) = \left[\prod_{\substack{1 \le i \le n-\mu \\ i:x_i=1}} \mathbf{T}_i \cdot \mathsf{RF}^\mu(\mathsf{suffix}^\mu(x)) \right] \in \mathbb{G}^k,$$

where the $\mathbf{T}_i$ are defined as in the real scheme. With this definition we have that $\mathcal{O}^0(x) = \mathsf{F}_K(x)$ (by defining $\mathsf{RF}(\varepsilon) := \vec{h}$) and $\mathcal{O}^n(x)$ is a truly random function. It leaves to show that the output of oracle

$\mathcal{O}^\mu(\cdot)$ is computationally indistinguishable from $\mathcal{O}^{\mu+1}(\cdot)$. For the reduction we use Lemma 13, where Q is the maximal number of queries to oracle $\mathcal{O}$ made by the PRF adversary. It inputs

$$\left[\begin{pmatrix}\vec{h}_0^1\\ \vec{h}_1^1\end{pmatrix},\ldots,\begin{pmatrix}\vec{h}_0^Q\\ \vec{h}_1^Q\end{pmatrix}\right],$$

where $\vec{h}_1^j = \hat{\mathbf{T}}\vec{h_0^j}$ or uniformly random. Next, it picks $\mathbf{T}_i$ $(1 \le i \le n-\mu)$ and implicitly defines $\mathbf{T}_{n-\mu} = \hat{\mathbf{T}}$. On the j-th query $x^j = (x_1^j,\ldots,x_n^j)$ $(1 \le j \le Q$ and wlog all queries are distinct) to oracle $\mathcal{O}$, it returns

$$\mathcal{O}(x^j) = \left[\prod_{\substack{1\le i\le n-\mu-1\\ i:x_i^j=1}} \mathbf{T}_i \cdot \vec{h}^j_{x^j_{n-\mu}}\right].$$

If $\vec{h}_1^j = \hat{\mathbf{T}}\vec{h_0^j}$, then

$$\mathsf{RF}^\mu(\mathsf{suffix}^\mu(x^j)) = \vec{h}_0^j$$

is a random function on μ bits and

$$\mathcal{O}(x^j) = \left[\prod_{\substack{1\le i\le n-\mu\\ i:x_i^j=1}} \mathbf{T}_i \cdot \vec{h}_0^j\right] = \left[\prod_{\substack{1\le i\le n-\mu\\ i:x_i^j=1}} \mathbf{T}_i \cdot \mathsf{RF}^\mu(\mathsf{suffix}^\mu(x^j))\right]$$

perfectly simulates oracle $\mathcal{O}^\mu$ from Hybrid μ.

If $\vec{h}_1^j$ is uniform and independent from $\vec{h}_0^j$, then

$$\mathsf{RF}^{\mu+1}(\mathsf{suffix}^{\mu+1}(x^j)) = \vec{h}^j_{x^j_{n-\mu}}$$

is a random function on $\mu+1$ bits and

$$\mathcal{O}(x^j) = \left[\prod_{\substack{1\le i\le n-\mu-1\\ i:x_i^j=1}} \mathbf{T}_i \cdot \mathsf{RF}^{\mu+1}(\mathsf{suffix}^{\mu+1}(x^j))\right]$$

perfectly simulates oracle $\mathcal{O}^{\mu+1}$ from Hybrid $\mu+1$.

We remark that the loss in the reduction is independent of the number of queries Q to oracle $\mathcal{O}$, i.e., the reduction loses a factor of nk, where the factor n stems from the above hybrid argument, and the factor k from Lemma 13. □

5.4 Groth-Sahai Non-interactive Zero-Knowledge Proofs

Groth and Sahai gave a method to construct non-interactive witness-indistinguishable (NIWI) and non-interactive zero-knowledge (NIZK) proofs for satisfiability of a set of equations in a bilinear group $\mathcal{PG}$. (For formal definitions of NIWI and NIZK proofs we refer to [21].) The equations in the set can be of different types, but they can be written in a unified way as

$$\sum_{j=1}^{n} f(a_j,\mathsf{y}_j) + \sum_{i=1}^{m} f(\mathsf{x}_i,b_i) + \sum_{i=1}^{m}\sum_{j=1}^{n} f(\mathsf{x}_i,\gamma_{ij}\mathsf{y}_j) = t, \tag{4}$$

where A_1, A_2, A_T are $\mathbb{Z}_q$-modules, $\vec{\mathsf{x}} \in A_1^m$, $\vec{\mathsf{y}} \in A_2^n$ are the variables, $\vec{a} \in A_1^n$, $\vec{b} \in A_2^m$, $\mathbf{\Gamma} = (\gamma_{ij}) \in \mathbb{Z}_q^{m\times n}$, $t \in A_T$ are the constants and $f : A_1 \times A_2 \to A_T$ is a bilinear map. More specifically, considering only symmetric bilinear groups, equations are of one of these types:

i) Pairing product equations, with $A_1 = A_2 = \mathbb{G}$, $A_T = \mathbb{G}_T$, $f([\mathsf{x}], [\mathsf{y}]) = [\mathsf{xy}]_T \in \mathbb{G}_T$.

ii) Multi-scalar multiplication equations, with $A_1 = \mathbb{Z}_q$, $A_2 = A_T = \mathbb{G}$, $f(\mathsf{x}, [\mathsf{y}]) = [\mathsf{xy}] \in \mathbb{G}$.

iii) Quadratic equations in $\mathbb{Z}_q$, with $A_1 = A_2 = A_T = \mathbb{Z}_q$, $f(\mathsf{x}, \mathsf{y}) = \mathsf{xy} \in \mathbb{Z}_q$.

Overview. The GS proof system allows to construct NIWI and NIZK proofs for satisfiability of a set of equations of the type (4), i.e., proofs that there is a choice of variables — the witness — satisfying all equations simultaneously. The prover gives to the verifier a commitment to each element of the witness and some additional information, the proof. Commitments and proof satisfy some related set of equations computable by the verifier because of their algebraic properties. We stress that to compute the proof, the prover needs the randomness which it used to create the commitments. To give new instantiations of GS proofs we need to specify the distribution of the common reference string, which includes the commitment keys and some maps whose purpose is roughly to give some algebraic structure to the commitment space.

Commitments. We will now construct commitments to elements in $\mathbb{Z}_q$ and $\mathbb{G}$. The commitment key $[\mathbf{U}] = ([\vec{u}_1], \ldots, [\vec{u}_{k+1}]) \in \mathbb{G}^{\ell \times (k+1)}$ is of the form

$$[\mathbf{U}] = \begin{cases} [\mathbf{A} || \mathbf{A}\vec{w}] & \text{binding key (soundness setting)} \\ [\mathbf{A} || \mathbf{A}\vec{w} - \vec{z}] & \text{hiding key (WI setting)} \end{cases},$$

where $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$, $\vec{w} \leftarrow \mathbb{Z}_q^k$, and $\vec{z} \in \mathbb{Z}_q^\ell$, $\vec{z} \notin \mathrm{Im}(\mathbf{A})$ is a fixed, public vector. The two types of commitment keys are computationally indistinguishable based on the $\mathcal{D}_{\ell,k}$-MDDH Assumption.

To commit to $[\mathsf{y}] \in \mathbb{G}$ using randomness $\vec{r} \leftarrow \mathbb{Z}_q^{k+1}$ we define maps $\iota : \mathbb{G} \to \mathbb{Z}_q^\ell$ and $p : \mathbb{G}^\ell \to \mathbb{Z}_q$ as

$$\iota([\mathsf{y}]) = \mathsf{y} \cdot \vec{z}, \quad p([\vec{c}]) = \vec{\xi}^\top \cdot \vec{c}, \quad \text{defining } \mathsf{com}_{[\mathbf{U}],\vec{z}}([\mathsf{y}]; \vec{r}) := [\iota([\mathsf{y}]) + \mathbf{U}\vec{r}] \in \mathbb{G}^\ell,$$

where $\vec{\xi} \in \mathbb{Z}_q^\ell$ is an arbitrary vector such that $\vec{\xi}^\top \mathbf{A} = \vec{0}$ and $\vec{\xi}^\top \cdot \vec{z} = 1$. Note that, given $[\mathsf{y}]$, $\iota([\mathsf{y}])$ is not efficiently computable, but $[\iota([\mathsf{y}])]$ is, and this suffices to compute the commitment. On a binding key (soundness setting) we have that $p([\iota([\mathsf{y}])]) = \mathsf{y}$ for all $[\mathsf{y}] \in \mathbb{G}$ and that $p([\vec{u}_i]) = 0$ for all $i = 1 \ldots k+1$. So $p(\mathsf{com}_{[\mathbf{U}],\vec{z}}([\mathsf{y}]; \vec{r})) = \vec{\xi}^\top(\vec{z}\mathsf{y} + \mathbf{U}\vec{r}) = \vec{\xi}^\top \vec{z}\mathsf{y} + \vec{\xi}^\top(\mathbf{A}||\mathbf{A}\vec{w})\vec{r} = \mathsf{y}$ and the commitment is perfectly binding. On a hiding key (WI setting), $\iota([\mathsf{y}]) \in \mathrm{Span}(\vec{u}_1, \ldots, \vec{u}_{k+1})$ for all $[\mathsf{y}] \in \mathbb{G}$ which implies that the commitments are perfectly hiding.

To commit to a scalar $\mathsf{x} \in \mathbb{Z}_q$ using randomness $\vec{s} \leftarrow \mathbb{Z}_q^k$ we define the maps $\iota' : \mathbb{Z}_q \to \mathbb{Z}_q^\ell$ and $p' : \mathbb{G}^\ell \to \mathbb{Z}_q$ as

$$\iota'(\mathsf{x}) = \mathsf{x} \cdot (\vec{u}_{k+1} + \vec{z}), \quad p'([\vec{c}]) = \vec{\xi}^\top \vec{c}, \quad \text{defining } \mathsf{com}'_{[\mathbf{U}],\vec{z}}(\mathsf{x}; \vec{s}) := [\iota'(\mathsf{x}) + \mathbf{A}\vec{s}] \in \mathbb{G}^\ell.$$

where $\vec{\xi}$ is defined as above. Note that, given x, $\iota(\mathsf{x})$ is not efficiently computable, but $[\iota(\mathsf{x})]$ is, and this suffices to compute the commitment. On a binding key (soundness setting) we have that $p'([\iota'(\mathsf{x})]) = \mathsf{x}$ for all $\mathsf{x} \in \mathbb{Z}_q$ and $p'([\vec{u}_i]) = 0$ for all $i = 1 \ldots k$ so the commitment is perfectly binding. On a hiding key (WI setting), $\iota'(x) \in \mathrm{Span}(\vec{u}_1, \ldots, \vec{u}_k)$ for all $x \in \mathbb{Z}_q$, which implies that the commitment is perfectly hiding.

It will also be convenient to define a vector of commitments as $\mathsf{com}_{[\mathbf{U}],\vec{z}}([\vec{\mathsf{y}}]; \mathbf{R}) = [\iota([\vec{\mathsf{y}}^\top]) + \mathbf{UR}]$ and $\mathsf{com}'_{[\mathbf{U}],\vec{z}}(\vec{\mathsf{x}}; \mathbf{S}) = [\iota'(\vec{\mathsf{x}}^\top) + \mathbf{AS}]$, where $[\vec{\mathsf{y}}] \in \mathbb{G}^m$, $\vec{\mathsf{x}} \in \mathbb{Z}_q^n$, $\mathbf{R} \leftarrow \mathbb{Z}_q^{(k+1)\times m}$, $\mathbf{S} \leftarrow \mathbb{Z}_q^{k \times n}$ and the inclusion maps are defined component-wise.

Inclusion and Projection Maps. As we have seen, commitments are elements of $\mathbb{G}^\ell$. The main idea of GS NIWI and NIZK proofs is to give some algebraic structure to the commitment space (in this case, $\mathbb{G}^\ell$) so that the commitments to a solution in A_1, A_2 of a certain set of equations satisfy a related set of equations in some larger modules. For this purpose, if $[\vec{\mathsf{x}}] \in \mathbb{G}^\ell$ and $[\vec{\mathsf{y}}] \in \mathbb{G}^\ell$, we define the bilinear map $\tilde{F} : \mathbb{G}^\ell \times \mathbb{G}^\ell \to \mathbb{Z}_q^{\ell \times \ell}$ defined implicitly as:

$$\tilde{F}([\vec{\mathsf{x}}], [\vec{\mathsf{y}}]) = \vec{\mathsf{x}} \cdot \vec{\mathsf{y}}^\top,$$

as well as its symmetric variant $F([\vec{\mathsf{x}}],[\vec{\mathsf{y}}]) = \frac{1}{2}\tilde{F}([\vec{\mathsf{x}}],[\vec{\mathsf{y}}]) + \frac{1}{2}\tilde{F}([\vec{\mathsf{y}}],[\vec{\mathsf{x}}])$. Additionally, for any two row vectors of elements of $\mathbb{G}^\ell$ of equal length r $[\mathsf{X}] = [\vec{\mathsf{x}}_1, \ldots, \vec{\mathsf{x}}_r]$ and $[\mathsf{Y}] = [\vec{\mathsf{y}}_1, \ldots, \vec{\mathsf{y}}_r]$, we define the maps $\tilde{\bullet}$, $\bullet$ associated with $\tilde{F}$ and F as $[\mathsf{X}] \tilde{\bullet} [\mathsf{Y}] = [\sum_{i=1}^r \tilde{F}([\vec{\mathsf{x}}_i],[\vec{\mathsf{y}}_i])]_T$ and $[\mathsf{X}] \bullet [\mathsf{Y}] = [\sum_{i=1}^r F([\vec{\mathsf{x}}_i],[\vec{\mathsf{y}}_i])]_T$. To complete the details of the new instantiation, we must specify for each type of equation, for both $F' = F$ and $F' = \tilde{F}$:

a) some maps ι_T and p_T such that for all $\mathsf{x} \in A_1, \mathsf{y} \in A_2, [\vec{\mathsf{x}}] \in \mathbb{G}^\ell, [\vec{\mathsf{y}}] \in \mathbb{G}^\ell$,

$$F'([\iota_1(\mathsf{x})],[\iota_2(\mathsf{y})]) = \iota_T(f(\mathsf{x},\mathsf{y})) \qquad \text{and} \qquad p_T([F'([\vec{\mathsf{x}}],[\vec{\mathsf{y}}])]_T) = f(p_1([\vec{\mathsf{x}}]), p_2([\vec{\mathsf{y}}])),$$

where ι_1, ι_2 are either ι or ι' and p_1, p_2 either $[p]$ or p', according to the appropriate A_1, A_2 for each equation,

b) matrices $\mathbf{H}_1, \ldots, \mathbf{H}_\eta \in \mathbb{Z}_q^{k_1 \times k_2}$, where k_1, k_2 are the number of columns of $\mathbf{U_1}, \mathbf{U_2}$ respectively and which, in the witness indistinguishability setting, are a basis of all the matrices which are a solution of the equation $[\mathbf{U}_1\mathbf{H}] \bullet [\mathbf{U}_2] = [\mathbf{0}]_T$ if $F' = F$ or $[\mathbf{U}_1\mathbf{H}] \tilde{\bullet} [\mathbf{U}_2] = [\mathbf{0}]_T$ if $F' = \tilde{F}$, where $\mathbf{U}_1, \mathbf{U}_2$ are either $\mathbf{U}$ or $\mathbf{A}$, depending on the modules A_1, A_2. These matrices are necessary to randomize the NIWI and NIZK proofs.

To present the instantiations in concise form, in the following $\mathbf{H}^{r,s,m,n} = (h_{ij}) \in \mathbb{Z}_q^{m \times n}$ denotes the matrix such that $h_{rs} = -1$, $h_{sr} = 1$ and $h_{ij} = 0$ for $(i,j) \notin \{(r,s),(s,r)\}$. In summary, the elements which must be defined are:

- **Pairing product equations**. In this case, $A_1 = A_2 = \mathbb{G}$, $A_T = \mathbb{G}_T$, $\iota_1 = \iota_2 = \iota$, $p_1 = p_2 = [p]$, $\mathbf{U}_1 = \mathbf{U}_2 = \mathbf{U}$ and both for $F' = F$ and $F' = \tilde{F}$,

$$\iota_T([\mathsf{z}]_T) = \mathsf{z} \cdot \vec{z} \cdot \vec{z}^\top \in \mathbb{Z}_q^{\ell \times \ell} \qquad p_T([\mathsf{Z}]_T) = [\vec{\xi}^\top \mathsf{Z} \vec{\xi}]_T,$$

where $\mathsf{Z} = (\mathsf{Z}_{ij})_{1 \le i,j \le \ell} \in \mathbb{Z}_q^{\ell \times \ell}$. The equation $[\mathbf{UH}] \tilde{\bullet} [\mathbf{U}] = [\mathbf{0}]_T$ admits no solution, while all the solutions to $[\mathbf{UH}] \bullet [\mathbf{U}] = [\mathbf{0}]_T$ are generated by $\left\{\mathbf{H}^{r,s,k+1,k+1}\right\}_{1 \le r < s \le k+1}$.

- **Multi-scalar multiplication equations**. In this case, $A_1 = \mathbb{Z}_q$, $A_2 = A_T = \mathbb{G}$, $\iota_1 = \iota', \iota_2 = \iota$, $p_1 = p'$, $p_2 = [p]$, $\mathbf{U_1} = \mathbf{A}$, $\mathbf{U_2} = \mathbf{U}$ and for both $F' = \tilde{F}$ and $F' = F$,

$$\iota_T([\mathsf{z}]) = F'([\iota'(1)],[\iota([\mathsf{z}])]) \qquad p_T([\mathsf{Z}]_T) = [\vec{\xi}^\top \mathsf{Z} \vec{\xi}].$$

The equation $[\mathbf{AH}] \tilde{\bullet} [\mathbf{U}] = [\mathbf{0}]_T$ admits no solution, while all the solutions to $[\mathbf{AH}] \bullet [\mathbf{U}] = [\mathbf{0}]_T$ are generated by $\left\{\mathbf{H}^{r,s,k,k+1}\right\}_{1 \le r < s \le k}$.

- **Quadratic equations**. In this case, $A_1 = A_2 = A_T = \mathbb{Z}_q$, $\iota_1 = \iota_2 = \iota'$, $p_1 = p_2 = p'$ and $\mathbf{U_1} = \mathbf{U_2} = \mathbf{A}$, for both $F' = \tilde{F}$ and $F' = F$, we define

$$\iota_T(z) = F'([\iota'(1)],[\iota'(z)]) \qquad p_T([\mathsf{Z}]_T) = \vec{\xi}^\top \mathsf{Z} \vec{\xi}.$$

The equation $[\mathbf{AH}] \tilde{\bullet} [\mathbf{A}] = [\mathbf{0}]_T$ admits no solution, while all the solutions to $[\mathbf{AH}] \bullet [\mathbf{A}] = [\mathbf{0}]_T$ are generated by $\left\{\mathbf{H}^{r,s,k,k}\right\}_{1 \le r < s \le k}$.

To argue that the equation $[\mathbf{U}_1\mathbf{H}] \tilde{\bullet} [\mathbf{U}_2] = [\mathbf{0}]_T$ admits no solution, for each of the cases above, it is sufficient to argue that the vectors $\tilde{F}([\vec{u}_i],[\vec{u}_j])$ are linearly independent. This holds regardless of the matrix distribution $\mathcal{D}_{\ell,k}$ from basic linear algebra, since $\tilde{F}([\vec{u}_i],[\vec{u}_j])$ was defined as the implicit representation of the outer product of $\vec{u}_i$ and $\vec{u}_j$ and $\vec{u}_1, \ldots, \vec{u}_{k+1}$ are linearly independent.

Proof and Verification. For completeness, we now describe how do the prover and the verifier proceed. Define k_1, k_2 as the number of columns of $\mathbf{U}_1, \mathbf{U}_2$ respectively. On input $\mathcal{PG}$, $[\mathbf{U}], \vec{z}$, a set of equations and a set of witnesses $\vec{\mathsf{x}} \in A_1^m, \vec{\mathsf{y}} \in A_2^n$ the prover proceeds as follows:

$\mathcal{D}_{\ell,k}$-MDDH instantiation	elements of $\mathbb{G}$	elements of $\mathbb{Z}_q$
Commitment to a Variable	ℓ	0
Pairing product equation	$\ell(k+1)$	0
- Linear equation:	$k+1$	0
Multi-scalar multiplication equation	$\ell(k+1)$	0
- Linear equation with variables in $\mathbb{G}$	0	$k+1$
- Linear equation with variables in $\mathbb{Z}_q$	k	0
Quadratic equation	ℓk	0
- Linear equation	0	k

Table 1: Size of the proofs based on the $\mathcal{D}_{\ell,k}$-MDDH Assumption.

1. Commit to $\vec{\mathsf{x}}$ and $\vec{\mathsf{y}}$ as
$$[\mathbf{C}] = [\iota_1(\vec{\mathsf{x}}^\top) + \mathbf{U}_1\mathbf{R}], \qquad [\mathbf{D}] = [\iota_2(\vec{\mathsf{y}}^\top) + \mathbf{U}_2\mathbf{S}]$$
where $\mathbf{R} \leftarrow \mathbb{Z}_q^{k_1 \times m}$, $\mathbf{S} \leftarrow \mathbb{Z}_q^{k_2 \times n}$.

2. For each equation of the type (4), pick $\mathbf{T} \leftarrow \mathbb{Z}_q^{k_1 \times k_2}$, $r_i \leftarrow \mathbb{Z}_q$ and output $([\mathbf{\Pi}], [\mathbf{\Theta}])$, defined as:

$$\begin{aligned}
[\mathbf{\Pi}] &:= [\iota_2(\vec{b}^\top)\mathbf{R}^\top + \iota_2(\vec{\mathsf{y}}^\top)\mathbf{\Gamma}^\top\mathbf{R}^\top + \mathbf{U}_2\mathbf{S}\mathbf{\Gamma}^\top\mathbf{R}^\top - \mathbf{U}_2\mathbf{T}^\top + \sum_{1\le i\le \eta} r_i\mathbf{U}_2\mathbf{H}_i^\top] \\
[\mathbf{\Theta}] &:= [\iota_1(\vec{a}^\top)\mathbf{S}^\top + \iota_1(\vec{\mathsf{x}}^\top)\mathbf{\Gamma}\mathbf{S}^\top + \mathbf{U}_1\mathbf{T}]
\end{aligned}$$

The proof described above is for a general equation, the same optimizations for special types of equation as in the full version of [21] apply. In particular, when the map used is the symmetric map F, the size of the proof can be reduced. In addition, the size of the proof can also be reduced when all the elements in either A_1 or A_2 are constants. Taking these optimizations into account, we give the size of the commitments and the proof for the different types of equations in Table 1.

To verify a proof, on input the commitments $[\mathbf{C}]$, $[\mathbf{D}]$ and a proof $([\mathbf{\Pi}], [\mathbf{\Theta}])$, the verifier checks if

$$[\iota_1(\vec{a}^\top)] \bullet' [\mathbf{D}] + [\mathbf{C}] \bullet' [\iota_2(\vec{b}^\top)] + [\mathbf{C}] \bullet' [\mathbf{D}\mathbf{\Gamma}^\top] = [\iota_T(t)]_T + [\mathbf{U}_1] \bullet' [\mathbf{\Pi}] + [\mathbf{\Theta}] \bullet' [\mathbf{U}_2],$$

where $\bullet'$ is either $\bullet$ or $\tilde{\bullet}$, depending on whether F' is F or $\tilde{F}$. If the equation is satisfied, the verifier accepts the proof for this equation and rejects otherwise. In general, the verification cost depends on ℓ and k, though a bit might be gained in pairing computations when using batch verification techniques and if some components of the commitment keys are trivial or are repeated, i.e. if the $\mathcal{D}_{\ell,k}$ admits short representation.

Efficiency. We emphasize that for $\mathcal{D}_{\ell,k} = \mathcal{L}_2$ and $\vec{z} = (0,0,1)^\top$ and for $\mathcal{D}_{\ell,k} = \mathsf{DDH}$ and $\vec{z} = (0,1)^\top$ (in the natural extension to asymmetric bilinear groups), we recover the 2-Lin and the SXDH instantiations of [21]. While the size of the proofs depends only on ℓ and k, both the size of the CRS and the cost of verification increase with $\mathsf{RE}_{\mathbb{G}}(\mathcal{D}_{\ell,k})$. In particular, in terms of efficiency, the $\mathcal{SC}_2$ Assumption is preferable to the 2-Lin Assumption but the main reason to consider more instantiations of GS proofs is to obtain more efficient proofs for a large class of languages in Section 6.

6 More Efficient Proofs for Some CRS Dependent Languages

Let $[\mathbf{U}]$ be the commitment key defined in last section as part of a $\mathcal{D}_{\ell,k}$-MDDH instantiation, for some $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$. In this section, we show how to obtain shorter proofs of some languages related to $\mathbf{A}$. The common idea of all the improvements is to exploit the special structure of the homomorphic commitments used in Groth-Sahai proofs.

6.1 More Efficient Subgroup Membership Proofs

We first show how to obtain shorter proofs of membership in the language $\mathcal{L}_{\mathbf{A},\mathcal{PG}} := \{[\mathbf{A}\vec{r}], \vec{r} \in \mathbb{Z}_q^k\} \subset \mathbb{G}^\ell$.

Intuition. Our proofs implicitly use the GS framework, although we have preferred to give the proofs without using the GS notation. Indeed, the idea behind our improvement is to exploit the special algebraic structure of commitments in GS proofs, namely the observation that if $[\vec{\Phi}] = [\mathbf{A}\vec{r}] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}}$ then $[\vec{\Phi}] = \mathsf{com}'_{[\mathbf{U}],\vec{z}}(0;\vec{r})$. Therefore, to prove that $[\vec{\Phi}] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}}$, we proceed as if we were giving a GS proof of satisfability of the equation $\mathsf{x} = 0$ where the randomness used for the commitment to x is $\vec{r}$. In particular, no commitments have to be given in the proof, which results in shorter proofs. To prove zero-knowledge we rewrite the equation $\mathsf{x} = 0$ as $\mathsf{x} \cdot \delta = 0$. The real proof is just a standard GS proof with the commitment to $\delta = 1$ being $\iota'(1) = \mathsf{com}_{[\mathbf{U}]}(1;\vec{0})$, while in the simulated proof the trapdoor allows to open $\iota'(1)$ as a commitment to 0, so we can proceed as if the equation was the trivial one $\mathsf{x} \cdot 0 = 0$, for which it is easy to give a proof of satisfiability.

Related Work. It is interesting to compare in detail with a recent line of work aiming at obtaining very efficient arguments of membership in linear subspaces ([25, 26, 31, 34]) which also exploits the dependency of the common reference string and the space where one wants to prove membership in. More specifically, these works construct NIZK arguments of membership in the space generated by $[\mathbf{A}] \in \mathbb{G}^{\ell\times k}$, with perfect zero-knowledge and computational soundness. We compare our results with [31], who give two different constructions which generalize and simplify previous results. In their work, computational soundness is based on any $\mathcal{D}_m$-MDDH Assumption[7]. In the first construction, the proof size is $m+1$, the common reference string must include $m\ell + (m+1)k + \mathsf{RE}_{\mathbb{G}}(\mathcal{D}_m)$ group elements and a description of $[\mathbf{A}]$. In the second construction, which assumes that $[\mathbf{A}]$ is drawn from a witness samplable distribution, the proof size is m and the common reference string must include $m\ell + mk + \mathsf{RE}_{\mathbb{G}}(\overline{\mathcal{D}}_m)$ group elements, where $\overline{\mathcal{D}}_m$ denotes the distribution of the first m rows of the matrices sampled according to $\mathcal{D}_m$, and a description of $[\mathbf{A}]$. Our proof, on the other hand, has perfect soundness, composable zero-knowledge under the $\mathcal{D}_{\ell,k}$-MDDH Assumption, proof of size ℓk and apart from a description of $[\mathbf{A}]$, the common reference string consists of only ℓ elements of $\mathbb{G}$.

6.1.1 Construction

Define $\mathcal{H} := \{\mathbf{H} \in \mathbb{Z}_q^{k\times k} : \mathbf{H} + \mathbf{H}^\top = \mathbf{0}\}$. Following the intuition given above, the actual construction looks as follows:

Setup. At the setup stage, some group $\mathcal{PG} = (\mathbb{G}, \mathbb{G}_T, q, e, \mathcal{P}) \leftarrow \mathsf{PGen}(1^\lambda)$ is specified.
Common reference string. We define $[\mathbf{U}] = ([\vec{u}_1], \ldots, [\vec{u}_{k+1}])$ as $[\mathbf{A}||\mathbf{A}\vec{w} + \vec{z}]$ in the soundness setting and $[\mathbf{A}||\mathbf{A}\vec{w}]$ in the witness indistinguishability setting, where $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$, $\vec{w} \leftarrow \mathbb{Z}_q^k$, and $\vec{z} \in \mathbb{Z}_q^\ell$, $\vec{z} \notin \mathrm{Im}(\mathbf{A})$. The common reference string is $\sigma := (\mathcal{PG}, [\mathbf{U}], \vec{z})$.
Simulation trapdoor. The simulation trapdoor τ is the vector $\vec{w} \in \mathbb{Z}_q^k$.
Prover. On input σ, a vector $[\vec{\Phi}] = [\mathbf{A}\vec{r}] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}}$ and the witness $\vec{r} \in \mathbb{Z}_q^k$, the prover chooses a matrix $\mathbf{H} \leftarrow \mathcal{H}$ and computes

$$[\mathbf{\Pi}] = [\vec{u}_{k+1}\vec{r}^\top + \mathbf{A}\mathbf{H}].$$

Verifier. On input $\sigma, [\vec{\Phi}], [\mathbf{\Pi}]$, the verifier checks if $[\vec{\Phi}\vec{u}_{k+1}^\top + \vec{u}_{k+1}\vec{\Phi}^\top]_T = [\mathbf{\Pi}\mathbf{A}^\top + \mathbf{A}\mathbf{\Pi}^\top]_T$.
Simulator. On input $\sigma, [\vec{\Phi}], \tau$ the simulator picks a matrix $\mathbf{H}' \leftarrow \mathcal{H}$ and computes

$$[\mathbf{\Pi}_{\mathrm{sim}}] = [\vec{\Phi}\vec{w}^\top + \mathbf{A}\mathbf{H}'].$$

Theorem 14. *Let $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$, where $\mathcal{D}_{\ell,k}$ is a matrix distribution. There exists a Non-Interactive Zero-Knowledge Proof for the language $\mathcal{L}_{\mathbf{A},\mathcal{PG}}$, with perfect completeness, perfect soundness and composable zero-knowledge of $k\ell$ group elements based on the $\mathcal{D}_{\ell,k}$-MDDH Assumption.*

[7] Actually, to be precise, soundness is based on a computational variant of the $\mathcal{D}_m$-MDDH Assumption.

The proof follows directly by implicitly reconstructing the same arguments which prove the same properties for the GS proof system.

Proof. First, it is clear that under the $\mathcal{D}_{\ell,k}$-MDDH Assumption, the soundness and the WI setting are computationally indistinguishable.

Completeness. To see completeness, we see that a real proof satisfies the verification equation. Indeed, in the soundness setting, the left term of the verification equation is:

$$\begin{aligned}[\vec{\Phi}\vec{u}_{k+1}^{\top} + \vec{u}_{k+1}\vec{\Phi}^{\top}]_T &= [\mathbf{A}\vec{r}(\mathbf{A}\vec{w}+\vec{z})^{\top} + (\mathbf{A}\vec{w}+\vec{z})(\mathbf{A}\vec{r})^{\top}]_T \\ &= [\mathbf{A}(\vec{r}\vec{w}^{\top} + \vec{w}\vec{r}^{\top})\mathbf{A}^{\top} + \mathbf{A}\vec{r}\vec{z}^{\top} + \vec{z}\vec{r}^{\top}\mathbf{A}^{\top}]_T\end{aligned}$$

while the right term in the real proof is:

$$\begin{aligned}[\mathbf{\Pi}\mathbf{A}^{\top} + \mathbf{A}\mathbf{\Pi}^{\top}]_T &= [\mathbf{A}(\vec{w}\vec{r}^{\top} + \vec{w}\vec{r}^{\top})\mathbf{A}^{\top} + \mathbf{A}(\mathbf{H}+\mathbf{H}^{\top})\mathbf{A}^{\top} + \mathbf{A}\vec{r}\vec{z}^{\top} + \vec{z}\vec{r}^{\top}\mathbf{A}^{\top}]_T &(5)\\ &= [\mathbf{A}(\vec{r}\vec{w}^{\top} + \vec{w}\vec{r}^{\top})\mathbf{A}^{\top} + \mathbf{A}\vec{r}\vec{z}^{\top} + \vec{z}\vec{r}^{\top}\mathbf{A}^{\top}]_T. &(6)\end{aligned}$$

This proves perfect completeness.

Soundness. Let $\vec{\xi} \in \mathbb{Z}_q^{\ell}$ be any vector such that $\vec{\xi}^{\top}\mathbf{A} = \vec{0}$, $\vec{\xi}^{\top}\vec{z} = 1$. This implies that in the soundness setting, $\vec{\xi}^{\top}\vec{u}_{k+1} = 1$. Therefore, if $[\mathbf{\Pi}]$ is any proof that satisfies the verification equation, multiplying on the left by $\vec{\xi}^{\top}$ and the right by $\vec{\xi}$,

$$\vec{\xi}^{\top}[\vec{\Phi}\vec{u}_{k+1}^{\top} + \vec{u}_{k+1}\vec{\Phi}^{\top}]_T\vec{\xi} = \vec{\xi}^{\top}[\mathbf{\Pi}\mathbf{A}^{\top} + \mathbf{A}\mathbf{\Pi}^{\top}]_T\vec{\xi},$$

we obtain

$$[\vec{\xi}^{\top}\vec{\Phi} + \vec{\Phi}^{\top}\vec{\xi}]_T = [0]_T. \tag{7}$$

Since $[\vec{\xi}^{\top}\vec{\Phi} + \vec{\Phi}^{\top}\vec{\xi}]_T = 2[\vec{\xi}^{\top}\vec{\Phi}]_T$, from this last equation it follows that $[\vec{\xi}^{\top}\vec{\Phi}]_T = [0]_T$. This holds for any vector $\vec{\xi}$ such that $\vec{\xi}^{\top}\mathbf{A} = \vec{0}$ and $\vec{\xi}^{\top}\vec{z} = 1$, which implies that $[\Phi] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}}$, which proves perfect soundness.

Composable Zero-Knowledge. We will now see that, in the witness indistinguishability setting, both a real proof and a simulated proof have the same distribution when $[\Phi] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}}$. We first note that they both satisfy the verification equation. Indeed, the left term of the verification equation in the WI setting is

$$[\vec{\Phi}\vec{u}_{k+1}^{\top} + \vec{u}_{k+1}\vec{\Phi}^{\top}]_T = [\mathbf{A}(\vec{r}\vec{w}^{\top} + \vec{w}\vec{r}^{\top})\mathbf{A}^{\top}]_T,$$

which is obviously equal to the right term of the verification equation for the real proof (rewrite equation (5) in the WI setting). On the other hand, if $[\Phi] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}}$, the right term of the verification equation for a simulated proof is:

$$\begin{aligned}[\mathbf{\Pi}_{\text{sim}}\mathbf{A}^{\top} + \mathbf{A}\mathbf{\Pi}_{\text{sim}}^{\top}]_T &= [\mathbf{A}(\vec{r}\vec{w}^{\top} + \vec{w}\vec{r}^{\top})\mathbf{A}^{\top} + \mathbf{A}(\mathbf{H}' + (\mathbf{H}')^{\top})\mathbf{A}^{\top}]_T \\ &= [\mathbf{A}(\vec{r}\vec{w}^{\top} + \vec{w}\vec{r}^{\top})\mathbf{A}^{\top}]_T,\end{aligned}$$

for some $\mathbf{H}' \in \mathcal{H}$.

We now argue that an honestly generated proof $[\mathbf{\Pi}]$ and a simulated proof $[\mathbf{\Pi}_{\text{sim}}]$ have the same distribution. By construction, there exist some matrices $\mathbf{\Theta}$ and $\mathbf{\Theta}'$ such that $[\mathbf{\Pi}] = [\mathbf{A}\mathbf{\Theta}]$ and $[\mathbf{\Pi}_{\text{sim}}] = [\mathbf{A}\mathbf{\Theta}']$. Now, if $[\mathbf{\Pi}_1] = [\mathbf{A}\mathbf{\Theta}_1]$ and $[\mathbf{\Pi}_2] = [\mathbf{A}\mathbf{\Theta}_2]$ are two proofs, real or simulated, which satisfy the verification equation, then necessarily $[(\mathbf{\Pi}_1 - \mathbf{\Pi}_2)\mathbf{A}^{\top} + \mathbf{A}(\mathbf{\Pi}_1 - \mathbf{\Pi}_2)]_T = [\mathbf{A}((\mathbf{\Theta}_1 - \mathbf{\Theta}_2) + (\mathbf{\Theta}_1 - \mathbf{\Theta}_2)^{\top})\mathbf{A}^{\top}]_T = 0$.

Since with overwhelming probability, $\mathbf{A}$ has rank k, it must hold that $(\mathbf{\Theta}_1 - \mathbf{\Theta}_2) + (\mathbf{\Theta}_1 - \mathbf{\Theta}_2)^{\top} = 0$, that is, it must hold that $(\mathbf{\Theta}_1 - \mathbf{\Theta}_2) \in \mathcal{H}$. By construction, both for honestly generated proofs $[\mathbf{\Pi}]$ and simulated proofs these difference is uniformly distributed in $\mathcal{H}$. □

6.1.2 Efficiency Comparison and Applications

For the 2-Lin Assumption, ($\ell = 3, k = 2$) our proof consists of only 6 group elements, whereas without using our technique the proof consists of 12 elements.[8] More generally, To prove that $[\vec{\Phi}] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}}$, for some $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$ with a GS instantiation based on a (possibly unrelated) $\mathcal{D}_{\ell',k'}$-matrix DH problem using standard GS proofs, one would prove that the following equation is satisfiable for all $i = 1 \dots \ell$:

$$r_1[u_{1,i}] + \ldots + r_k[u_{k,i}] = [\Phi_i], \tag{8}$$

that is, one needs to prove that ℓ linear equations with k variables are satisfied. Therefore, according to Table 1, the verifier must be given $k\ell'$ elements of $\mathbb{G}$ for the commitments and $\ell k'$ elements of $\mathbb{G}$ for the proof. On the other hand, proving $[\vec{\Phi}] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}}$ using our approach requires ℓk elements of $\mathbb{G}$, corresponding to the size of the proof of one quadratic equation.

Applications. For a typical application scenario of Theorem 14, think of $[\mathbf{A}]$ as part of the public parameters of the hash proof system of Section 5.2. Proving that a ciphertext is well-formed is proving membership in $\mathcal{L}_{\mathbf{A},\mathcal{PG}}$. Another application is to show that two ciphertexts encrypt the same message under the same public key, a common problem in electronic voting or anonymous credentials. There are many other settings in which subgroup membership problems naturally appear, for instance the problem of certifying public keys or given some plaintext m, the problem of proving that a certain ciphertext is an encryption of $[m]$. We stress that in our construction the setup of the CRS can be built on top of the encryption key so that proofs can be simulated without the decryption key, which is essential for many of these applications. More concretely, below we give two application examples.

Application Example 1. The standard proof of membership in $\mathcal{L}_{\mathbf{A},\mathcal{PG}}$, when $\mathbf{A} \leftarrow$ 2-Lin based on the same assumption (with $\ell = \ell' = 3$, $k = k' = 2$), requires 12 group elements, while with our approach only 6 elements are required[9]. This reduces the ciphertext size of one of the instantiations of [35] from 15 to 9 group elements.

Application Example 2. With our results we can also give a more efficient proof of correct opening of the Cramer Shoup ciphertext. We briefly recall the CS encryption scheme based on the 2-Lin-Assumption ([23, 45]). The public key consists of the description of some group $\mathcal{G}$ and a tuple $[a_1, a_2, X_1, X_2, X_3, X_4, X_5, X_6] \in \mathbb{G}^8$. Given a message $[m] \in \mathbb{G}$, a ciphertext is constructed by picking random $r, s \in \mathbb{Z}_q$ and setting

$$C := [r(a_1, 0, 1, X_5, X_1 + \alpha X_3) + s(0, a_2, 1, X_6, X_2 + \alpha X_4) + (0, 0, m, 0, 0)],$$

where α is the hash of some components of the ciphertext and possibly some label. To prove that a ciphertext opens to a (known) message $[m]$, substract $[m]$ from the third component of the ciphertext and prove membership in $\mathcal{L}_{\mathbf{A}_\alpha,\mathcal{PG}}$, where $\mathbf{A}_\alpha$ is defined as:

$$\mathbf{A}_\alpha := \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \\ 1 & 1 \\ X_5 & X_6 \\ X_1 + \alpha X_3 & X_2 + \alpha X_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \alpha \end{pmatrix} \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \\ 1 & 1 \\ X_5 & X_6 \\ X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}.$$

Denote $\mathbf{M}_\alpha, \mathbf{C}$, the two matrices of the right term of the previous equation such that $\mathbf{A}_\alpha = \mathbf{M}_\alpha \mathbf{C}$. The matrix $\mathbf{A}_\alpha$ depends on α and is different for each ciphertext, so it cannot be included in the CRS. Instead, we include the matrix $[\mathbf{U}_C] := [\mathbf{C} || \mathbf{C}\vec{w} + \vec{z}_C]$ in the soundness setting and $[\mathbf{U}_C] := [\mathbf{C} || \mathbf{C}\vec{w}]$ in the WI setting, for $\vec{z}_C \notin \mathrm{Im}(\mathbf{C})$, for instance $\vec{z}_C^\top := (0, 0, 0, 0, 1, 0)$. To prove membership in $\mathcal{L}_{\mathbf{A}_\alpha,\mathcal{PG}}$ as we explained, we would make the proof with respect to the CRS $[\mathbf{U}_\alpha] := [\mathbf{M}_\alpha \mathbf{U}_C]$. Clearly, if $\vec{z}^\top := (0, 0, 0, 0, 1)$, $[\mathbf{U}_\alpha] = [\mathbf{A}_\alpha || \mathbf{A}_\alpha \vec{w} + \vec{z}]$ in the soundness setting and $[\mathbf{U}_\alpha] = [\mathbf{A}_\alpha || \mathbf{A}_\alpha \vec{w}]$ in the WI, as required. The resulting proof consists of 10 group elements, as opposed to 16 using standard GS proofs. This applies to the result of [17], Section 3.

[8] For completeness, a detailed comparison for the 2-Lin case can be found in appendix D.

[9] A detailed comparison for 2-Lin case is given in Appendix D. The same results hold for the Symmetric 2-cascade assumption.

6.2 Other CRS Dependent Languages

The techniques of the previous section can be extended to other languages, namely:

- A proof of validity of a ciphertext, that is, given $[\mathbf{A}]$, $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$, and some vector $\vec{z} \in \mathbb{Z}_q^\ell$, $\vec{z} \notin \mathrm{Im}(\mathbf{A})$, one can use the same techniques to give a more efficient proof of membership in the space:

$$\mathcal{L}_{\mathbf{A},\vec{z},\mathcal{PG}} = \{[\vec{c}] : \vec{c} = \mathbf{A}\vec{r} + m\vec{z}\} \subset \mathbb{G}^\ell,$$

  where $(\vec{r}, [m]) \in \mathbb{Z}_q^k \times \mathbb{G}$ is the witness. This is also a proof of membership in the subspace of $\mathbb{G}^\ell$ spanned by the columns of $[\mathbf{A}]$ and the vector $\vec{z}$, but part of the witness, $[m]$, is in the group $\mathbb{G}$ and not in $\mathbb{Z}_q$, while part of the matrix generating the subspace is in $\mathbb{Z}_q$. However, it is not hard to modify the subgroup membership proofs as described in Section 6.1 to account for this. In particular, since GS are non-interactive zero-knowledge proofs of knowledge when the witnesses are group elements, the proof guarantees both that $[\vec{c}]$ is well-formed and that the prover knows $[m]$. In a typical application, $[\vec{c}]$ will be the ciphertext of some encryption scheme, in which case $\vec{r}$ will be the ciphertext randomness and $[m]$ the message.

- A proof of plaintext equality. The encryption scheme derived from the KEM given in Section 5.1 corresponds to a commitment in GS proofs — except that the commitment is always binding. That is, if $pk_A = (\mathcal{G}, [\mathbf{A}] \in \mathbb{G}^{\ell \times k})$, for some $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}$, given $\vec{r} \in \mathbb{Z}_q^k$,

$$\mathsf{Enc}_{pk_A}([m]; \vec{r}) = [\vec{c}] = [\mathbf{A}\vec{r} + (0, \ldots, 0, m)^\top] = [\mathbf{A}\vec{r} + m \cdot \vec{z}] = \mathsf{com}_{[\mathbf{A}||\mathbf{A}\vec{w}]}([m]; \vec{s}),$$

  where $\vec{s}^\top := (\vec{r}^\top, 0)$ and $\vec{z} := (0, \ldots, 0, 1)^\top$. Therefore, given two (potentially distinct) matrix distributions $\mathcal{D}_{\ell_1,k_1}$, $\mathcal{D}'_{\ell_2,k_2}$ and $\mathbf{A} \leftarrow \mathcal{D}_{\ell_1,k_1}, \mathbf{B} \leftarrow \mathcal{D}'_{\ell_2,k_2}$, proving equality of plaintexts of two ciphertexts encrypted under pk_A, pk_B, corresponds to proving that two commitments under different keys open to the same value. One can gain in efficiency with respect to the standard use of GS proofs because one does not need to give any commitments as part of the proof, since the ciphertexts themselves play this role. More specifically, given $[\vec{c}_A] = \mathsf{Enc}_{pk_A}([m])$ and $[\vec{c}_B] = \mathsf{Enc}_{pk_B}([m])$, one can treat $[\vec{c}_A]$ as a commitment to the variable $[\mathsf{x}] \in A_1 = \mathbb{G}$ and $[\vec{c}_B]$ as a commitment to the variable $[\mathsf{y}] \in A_2 = \mathbb{G}$ and prove that the quadratic equation $e([\mathsf{x}], [1]) \cdot e([-1], [\mathsf{y}]) = [0]_T$ is satisfied. The problem is only how to construct the simulator of the NIZK proof system, since commitments are always binding. For this, one uses a similar trick as in the membership proofs, namely to let the zero-knowledge simulator open $\iota_1([1])$, $\iota_2([-1])$ as commitments to the $[0]$ variable and simulate a proof for the equation $e([\mathsf{x}], [0]) \cdot e([0], [\mathsf{y}]) = [0]_T$, which is trivially satisfiable and can be simulated. In [27], we reduce the size of the proof by 4 group elements from 18 to 22, while in [22] we save 9 elements although their proof is quite inefficient altogether. We note that even if both papers give a proof that two ciphertexts under two different 2-Lin public keys correspond to the same value, the proof in [22] is more inefficient because it must use GS proofs for pairing product equations instead of multi-scalar multiplication equations. Other examples include [10, 15].

Acknowledgements

We thank Duong Hieu Phan for pointing out a small mistake in a previous draft.

References

[1] O. Blazy, D. Pointcheval, and D. Vergnaud. Round-optimal privacy-preserving protocols with smooth projective hash functions. In R. Cramer, editor, *TCC 2012*, volume 7194 of *LNCS*, pages 94–111, Taormina, Sicily, Italy, Mar. 19–21, 2012. Springer, Berlin, Germany. 3

[2] D. Boneh, X. Boyen, and E.-J. Goh. Hierarchical identity based encryption with constant size ciphertext. In R. Cramer, editor, *EUROCRYPT 2005*, volume 3494 of *LNCS*, pages 440–456, Aarhus, Denmark, May 22–26, 2005. Springer, Berlin, Germany. 2, 8, 29

[3] D. Boneh, X. Boyen, and H. Shacham. Short group signatures. In M. Franklin, editor, *CRYPTO 2004*, volume 3152 of *LNCS*, pages 41–55, Santa Barbara, CA, USA, Aug. 15–19, 2004. Springer, Berlin, Germany. 1, 5, 9

[4] D. Boneh and M. K. Franklin. Identity-based encryption from the Weil pairing. In J. Kilian, editor, *CRYPTO 2001*, volume 2139 of *LNCS*, pages 213–229, Santa Barbara, CA, USA, Aug. 19–23, 2001. Springer, Berlin, Germany. 1, 5

[5] D. Boneh, S. Halevi, M. Hamburg, and R. Ostrovsky. Circular-secure encryption from decision Diffie-Hellman. In D. Wagner, editor, *CRYPTO 2008*, volume 5157 of *LNCS*, pages 108–125, Santa Barbara, CA, USA, Aug. 17–21, 2008. Springer, Berlin, Germany. 1, 2

[6] D. Boneh, H. W. Montgomery, and A. Raghunathan. Algebraic pseudorandom functions with improved efficiency from the augmented cascade. In E. Al-Shaer, A. D. Keromytis, and V. Shmatikov, editors, *ACM CCS 10*, pages 131–140, Chicago, Illinois, USA, Oct. 4–8, 2010. ACM Press. 1, 3, 15, 35

[7] D. Boneh, A. Sahai, and B. Waters. Fully collusion resistant traitor tracing with short ciphertexts and private keys. In S. Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 573–592, St. Petersburg, Russia, May 28 – June 1, 2006. Springer, Berlin, Germany. 1, 3, 5, 9

[8] D. Boneh and A. Silverberg. Applications of multilinear forms to cryptography. *Contemporary Mathematics*, 324:71–90, 2003. 5

[9] X. Boyen. The uber-assumption family (invited talk). In S. D. Galbraith and K. G. Paterson, editors, *PAIRING 2008*, volume 5209 of *LNCS*, pages 39–56, Egham, UK, Sept. 1–3, 2008. Springer, Berlin, Germany. 2, 8, 29

[10] J. Camenisch, N. Chandran, and V. Shoup. A public key encryption scheme secure against key dependent chosen plaintext and adaptive chosen ciphertext attacks. In A. Joux, editor, *EUROCRYPT 2009*, volume 5479 of *LNCS*, pages 351–368, Cologne, Germany, Apr. 26–30, 2009. Springer, Berlin, Germany. 4, 23

[11] D. Cox, J. Little, and D. O'Shea. *Ideal, Varieties and Algorithms*. Springer, second edition, 1996. 28, 29, 30, 31

[12] R. Cramer and V. Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In H. Krawczyk, editor, *CRYPTO'98*, volume 1462 of *LNCS*, pages 13–25, Santa Barbara, CA, USA, Aug. 23–27, 1998. Springer, Berlin, Germany. 1

[13] R. Cramer and V. Shoup. Universal hash proofs and a paradigm for adaptive chosen ciphertext secure public-key encryption. In L. R. Knudsen, editor, *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 45–64, Amsterdam, The Netherlands, Apr. 28 – May 2, 2002. Springer, Berlin, Germany. 1, 3, 6

[14] R. Cramer and V. Shoup. Design and analysis of practical public-key encryption schemes secure against adaptive chosen ciphertext attack. *SIAM Journal on Computing*, 33(1):167–226, 2003. 6

[15] Y. Dodis, K. Haralambiev, A. López-Alt, and D. Wichs. Cryptography against continuous memory attacks. In *51st FOCS*, pages 511–520, Las Vegas, Nevada, USA, Oct. 23–26, 2010. IEEE Computer Society Press. 4, 23

[16] A. Escala, G. Herold, E. Kiltz, C. Ràfols, and J. Villar. An algebraic framework for Diffie-Hellman assumptions. In R. Canetti and J. A. Garay, editors, *CRYPTO 2013, Part II*, volume 8043 of *LNCS*, pages 129–147, Santa Barbara, CA, USA, Aug. 18–22, 2013. Springer, Berlin, Germany. 2, 3, 10, 11

[17] M. Fischlin, B. Libert, and M. Manulis. Non-interactive and re-usable universally composable string commitments with adaptive security. In D. H. Lee and X. Wang, editors, *ASIACRYPT 2011*, volume 7073 of *LNCS*, pages 468–485, Seoul, South Korea, Dec. 4–8, 2011. Springer, Berlin, Germany. 4, 22

[18] D. M. Freeman. Converting pairing-based cryptosystems from composite-order groups to prime-order groups. In H. Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 44–61, French Riviera, May 30 – June 3, 2010. Springer, Berlin, Germany. 1, 2

[19] D. Galindo, J. Herranz, and J. L. Villar. Identity-based encryption with master key-dependent message security and leakage-resilience. In S. Foresti, M. Yung, and F. Martinelli, editors, *ESORICS 2012*, volume 7459 of *LNCS*, pages 627–642, Pisa, Italy, Sept. 10–12, 2012. Springer, Berlin, Germany. 2

[20] R. Gennaro and Y. Lindell. A framework for password-based authenticated key exchange. In E. Biham, editor, *EUROCRYPT 2003*, volume 2656 of *LNCS*, pages 524–543, Warsaw, Poland, May 4–8, 2003. Springer, Berlin, Germany. `http://eprint.iacr.org/2003/032.ps.gz`. 1, 3

[21] J. Groth and A. Sahai. Efficient noninteractive proof systems for bilinear groups. *SIAM J. Comput.*, 41(5):1193–1232, 2012. 1, 3, 16, 19

[22] D. Hofheinz and T. Jager. Tightly secure signatures and public-key encryption. In R. Safavi-Naini and R. Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 590–607, Santa Barbara, CA, USA, Aug. 19–23, 2012. Springer, Berlin, Germany. 4, 23

[23] D. Hofheinz and E. Kiltz. Secure hybrid encryption from weakened key encapsulation. In A. Menezes, editor, *CRYPTO 2007*, volume 4622 of *LNCS*, pages 553–571, Santa Barbara, CA, USA, Aug. 19–23, 2007. Springer, Berlin, Germany. 1, 5, 9, 22, 34

[24] A. Joux. A one round protocol for tripartite Diffie-Hellman. *Journal of Cryptology*, 17(4):263–276, Sept. 2004. 1

[25] C. S. Jutla and A. Roy. Shorter quasi-adaptive NIZK proofs for linear subspaces. In K. Sako and P. Sarkar, editors, *ASIACRYPT 2013, Part I*, volume 8269 of *LNCS*, pages 1–20, Bengalore, India, Dec. 1–5, 2013. Springer, Berlin, Germany. 4, 20

[26] C. S. Jutla and A. Roy. Switching lemma for bilinear tests and constant-size NIZK proofs for linear subspaces. In J. A. Garay and R. Gennaro, editors, *CRYPTO 2014, Part II*, volume 8617 of *LNCS*, pages 295–312, Santa Barbara, CA, USA, Aug. 17–21, 2014. Springer, Berlin, Germany. 4, 20

[27] J. Katz and V. Vaikuntanathan. Round-optimal password-based authenticated key exchange. In Y. Ishai, editor, *TCC 2011*, volume 6597 of *LNCS*, pages 293–310, Providence, RI, USA, Mar. 28–30, 2011. Springer, Berlin, Germany. 4, 23

[28] E. Kiltz. A tool box of cryptographic functions related to the Diffie-Hellman function. In C. P. Rangan and C. Ding, editors, *INDOCRYPT 2001*, volume 2247 of *LNCS*, pages 339–350, Chennai, India, Dec. 16–20, 2001. Springer, Berlin, Germany. 3, 5

[29] E. Kiltz. Chosen-ciphertext security from tag-based encryption. In S. Halevi and T. Rabin, editors, *TCC 2006*, volume 3876 of *LNCS*, pages 581–600, New York, NY, USA, Mar. 4–7, 2006. Springer, Berlin, Germany. 1

[30] E. Kiltz, K. Pietrzak, M. Stam, and M. Yung. A new randomness extraction paradigm for hybrid encryption. In A. Joux, editor, *EUROCRYPT 2009*, volume 5479 of *LNCS*, pages 590–609, Cologne, Germany, Apr. 26–30, 2009. Springer, Berlin, Germany. 14

[31] E. Kiltz and H. Wee. Quasi-adaptive NIZK for linear subspaces revisited. In E. Oswald and M. Fischlin, editors, *EUROCRYPT 2015, Part II*, volume 9057 of *LNCS*, pages 101–128, Sofia, Bulgaria, Apr. 26–30, 2015. Springer, Berlin, Germany. 4, 20

[32] A. B. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters. Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption. In H. Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 62–91, French Riviera, May 30 – June 3, 2010. Springer, Berlin, Germany. 1

[33] A. B. Lewko and B. Waters. Efficient pseudorandom functions from the decisional linear assumption and weaker variants. In E. Al-Shaer, S. Jha, and A. D. Keromytis, editors, *ACM CCS 09*, pages 112–120, Chicago, Illinois, USA, Nov. 9–13, 2009. ACM Press. 15

[34] B. Libert, T. Peters, M. Joye, and M. Yung. Non-malleability from malleability: Simulation-sound quasi-adaptive NIZK proofs and CCA2-secure encryption from homomorphic signatures. In P. Q. Nguyen and E. Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 514–532, Copenhagen, Denmark, May 11–15, 2014. Springer, Berlin, Germany. 4, 20

[35] B. Libert and M. Yung. Non-interactive CCA-secure threshold cryptosystems with adaptive security: New framework and constructions. In R. Cramer, editor, *TCC 2012*, volume 7194 of *LNCS*, pages 75–93, Taormina, Sicily, Italy, Mar. 19–21, 2012. Springer, Berlin, Germany. 4, 22, 33

[36] S. Meiklejohn, H. Shacham, and D. M. Freeman. Limitations on transformations from composite-order to prime-order groups: The case of round-optimal blind signatures. In M. Abe, editor, *ASIACRYPT 2010*, volume 6477 of *LNCS*, pages 519–538, Singapore, Dec. 5–9, 2010. Springer, Berlin, Germany. 1

[37] M. Naor and O. Reingold. Number-theoretic constructions of efficient pseudo-random functions. In *38th FOCS*, pages 458–467, Miami Beach, Florida, Oct. 19–22, 1997. IEEE Computer Society Press. 1, 3

[38] M. Naor and G. Segev. Public-key cryptosystems resilient to key leakage. In S. Halevi, editor, *CRYPTO 2009*, volume 5677 of *LNCS*, pages 18–35, Santa Barbara, CA, USA, Aug. 16–20, 2009. Springer, Berlin, Germany. 1, 2

[39] M. Naor and M. Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *22nd ACM STOC*, pages 427–437, Baltimore, Maryland, USA, May 14–16, 1990. ACM Press. 4

[40] T. Okamoto and K. Takashima. Fully secure functional encryption with general relations from the decisional linear assumption. In T. Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 191–208, Santa Barbara, CA, USA, Aug. 15–19, 2010. Springer, Berlin, Germany. 1

[41] T. Okamoto and K. Takashima. Achieving short ciphertexts or short secret-keys for adaptively secure general inner-product encryption. In D. Lin, G. Tsudik, and X. Wang, editors, *CANS 11*, volume 7092 of *LNCS*, pages 138–159, Sanya, China, Dec. 10–12, 2011. Springer, Berlin, Germany. 1

[42] T. Okamoto and K. Takashima. Fully secure unbounded inner-product and attribute-based encryption. In X. Wang and K. Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 349–366, Beijing, China, Dec. 2–6, 2012. Springer, Berlin, Germany. 1

[43] J. H. Seo. On the (im)possibility of projecting property in prime-order setting. In X. Wang and K. Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 61–79, Beijing, China, Dec. 2–6, 2012. Springer, Berlin, Germany. 1

[44] J. H. Seo and J. H. Cheon. Beyond the limitation of prime-order bilinear groups, and round optimal blind signatures. In R. Cramer, editor, *TCC 2012*, volume 7194 of *LNCS*, pages 133–150, Taormina, Sicily, Italy, Mar. 19–21, 2012. Springer, Berlin, Germany. 1

[45] H. Shacham. A cramer-shoup encryption scheme from the linear assumption and from progressively weaker linear variants. Cryptology ePrint Archive, Report 2007/074, 2007. `http://eprint.iacr.org/`. 1, 5, 9, 22, 34

[46] J. L. Villar. Optimal reductions of some decisional problems to the rank problem. In X. Wang and K. Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 80–97, Beijing, China, Dec. 2–6, 2012. Springer, Berlin, Germany. 2

[47] S. Wolf. *Information-Theoretically and Computionally Secure Key Agreement in Cryptography.* PhD thesis, ETH Zuerich, 1999. 5

A Proof of Theorem 7

We split the theorem in several lemmas.

Lemma 15. $(k+1)$-PDDH $\Rightarrow$ k-Casc.

Proof. The idea of the proof is that an instance of the $(k+1)$-PDDH problem can be viewed as an instance of the $\mathcal{C}$-MDDH problem with a non-uniform distribution of $\vec{w}$. A suitable re-randomization of $\vec{w}$ yields the result. Let $(\mathcal{G}, [x_1], \ldots, [x_{k+1}], [z])$ be a $(k+1)$-PDDH instance with either $z \in \mathbb{Z}_q$ uniform or $z = x_1 \cdots x_{k+1}$. We will construct a k-Casc instance from that, setting $[\mathbf{A}]$ as follows:

$$[\mathbf{A}] = \begin{pmatrix} [x_1] & [0] & \ldots & [0] & [0] \\ [1] & [x_2] & \ldots & [0] & [0] \\ [0] & [1] & \ddots & & [0] \\ \vdots & & \ddots & & \vdots \\ [0] & [0] & \ldots & [1] & [x_k] \\ [0] & [0] & \ldots & 0 & [1] \end{pmatrix},$$

Let $[\vec{b}^\top] := \big((-1)^{k+1}[z], [0], [0], \ldots, [0], [x_{k+1}]\big)^T$. Since $\mathbf{A}$ has full rank, $\vec{b}$ is in the span of the columns of $\mathbf{A}$ if and only if $\det(\mathbf{A} \| \vec{b}) = 0$. Since $\det(\mathbf{A} \| \vec{b}) = x_1 \cdots x_k - z$, this depends on the distribution of z as desired. To obtain a properly distributed k-Casc instance $(\mathcal{G}, [\mathbf{A}], [\vec{b}'])$, we set $[\vec{b}'] = [\vec{b}] + \sum_i w_i [\vec{a}_i]$ for uniform $w_i \in \mathbb{Z}_q$. Clearly, if $\vec{b}$ is in the span of the columns of $\mathbf{A}$, $\vec{b}'$ will be a uniform element in the span of the columns of $\mathbf{A}$, whereas if it is not, $\vec{b}'$ will be uniform in all of $\mathbb{Z}_q^{k+1}$. □

Lemma 16. $(k+1)$-EDDH $\Rightarrow$ k-SCasc.

Proof. The proof is analogous to the proof of the preceding Lemma 15. Let $(\mathcal{G}, [x], [z])$ be a $(k+1)$-EDDH instance with either $z \in \mathbb{Z}_q$ uniform or $z = x^{k+1}$. We will construct a k-SCasc instance from that, defining $[\mathbf{A}]$ as the following $k \times (k+1)$-matrix:

$$[\mathbf{A}] = \begin{pmatrix} [x] & [0] & \ldots & [0] & [0] \\ [1] & [x] & \ldots & [0] & [0] \\ [0] & [1] & \ddots & & [0] \\ \vdots & & \ddots & & \vdots \\ [0] & [0] & \ldots & [1] & [x] \\ [0] & [0] & \ldots & [0] & [1] \end{pmatrix},$$

Set $[\vec{b}^\top] := \big((-1)^{k+1}[z], [0], [0], \ldots, [0], [x]\big)$. As above, $\vec{b}$ is in the span of the columns of $\mathbf{A}$ if and only if $z = x^{k+1}$. To obtain a properly distributed k-SCasc instance $(\mathcal{G}, [\mathbf{A}], [\vec{b}'])$, we set $[\vec{b}'] = [\vec{b}] + \sum_i w_i [\vec{a}_i]$ for uniform $w_i \in \mathbb{Z}_q$. □

Lemma 17. *In k-linear groups,* k-Casc $\Rightarrow$ k-MLDDH.

Proof. Assume for the purpose of contradiction that k-MLDDH does not hold. To break the k-Casc problem, we are given an instance $[\mathbf{A}], [\vec{z}]$, where $\mathbf{A} \leftarrow \mathcal{C}_k$ and we have to distinguish between $\vec{z} = \mathbf{A}\vec{w}$ for uniform $\vec{w}$ and uniform $\vec{z}$. Or, equivalently, we have to test if the determinant of matrix $\mathbf{B} = \mathbf{A}||\vec{z} \in \mathbb{Z}_q^{(k+1)\times(k+1)}$ is zero. But $\det \mathbf{B}$ is just the determinant polynomial of k-Casc defined in Section 3.3 and explicitly computed in the proof of Theorem 6. Namely,

$$\begin{aligned} \det \mathbf{B} &= \mathfrak{d}(a_1, \ldots, a_k, z_1, \ldots, z_{k+1}) = a_1 \cdots a_k z_{k+1} - a_1 \cdots a_{k-1} z_k + \ldots + (-1)^k z_1 = \\ &= a_1 \cdots a_k z_{k+1} + R_k(a_1, \ldots, a_k, z_1, \ldots, z_{k+1}), \end{aligned}$$

where R_k is a polynomial of degree k.

Hence, to test whether $\det(\mathbf{B}) = 0$, we compute $[b]_{T_k} = [-R_k(a_1, \ldots, a_k, z_1, \ldots, z_{k+1})]_{T_k}$ using the k-linear map, and then we use the oracle k-MLDDH$([a_1], \ldots, [a_k], [z_{k+1}], [b]_{T_k})$ to check if $a_1 \cdots a_k z_{k+1} = -b$. □

Lemma 18. *k-SCasc $\Rightarrow$ k-Casc, k-ILin $\Rightarrow$ k-Lin*

Proof. Both implications follow by simple rerandomization arguments. A k-SCasc instance $([a_1], \ldots, [a_k], [a_1 w_1], [w_1 + a_2 w_2], \ldots, [w_{k-1} + a_k w_k], [w_k])$ can be transformed into a k-Casc instance by picking $\alpha_1, \alpha_2, \ldots, \alpha_k \leftarrow \mathbb{Z}_q^*$ and computing $([a\alpha_1], [a\alpha_2], \ldots, [a\alpha_k], [aw_1], [\frac{w_1 + aw_2}{\alpha_1}], \ldots, [\frac{w_{k-1} + a_k w_k}{\alpha_1 \cdots \alpha_{k-1}}], [\frac{w_k}{\alpha_1 \cdots \alpha_k}])$. Similarly, a k-ILin instance $([a], [aw_1], [(a+1)w_2], \ldots, [(a+k-1)w_k], [w_1 + \ldots + w_k])$ can be transformed into a k-Lin instance by picking random $\alpha_1, \alpha_2, \ldots, \alpha_k \leftarrow \mathbb{Z}_q^*$ and computing $([a\alpha_1], [(a+1)\alpha_2], \ldots, [(a+k-1)\alpha_k], [aw_1\alpha_1], [(a+1)w_2\alpha_2], \ldots, [(a+k-1)w_k\alpha_k], [w_1 + \ldots + w_k])$. □

Lemma 19. *k-Casc $\Rightarrow$ $(k+1)$-Casc, k-SCasc $\Rightarrow$ $(k+1)$-SCasc.*

Proof. To show the first implication, we transform a given instance of the k-Casc problem $\mathcal{D}_1 = ([a_1], \ldots, [a_k], [a_1 w_1], [w_1 + a_2 w_2], \ldots, [w_{k-1} + a_k w_k], [w_k])$ into an instance of the $(k+1)$-Casc problem by picking uniform $w_{k+1} \leftarrow \mathbb{Z}_q$ and $[a_{k+1}] \leftarrow \mathbb{G}$ and computing $\mathcal{D}_2 = ([a_1], \ldots, [a_{k+1}], [a_1 w_1], [w_1 + a_2 w_2], \ldots, [w_{k-1} + a_k w_k], [w_k + a_{k+1} w_{k+1}], [w_{k+1}])$. Note that $\mathcal{D}_2$ is pseudorandom if and only if $\mathcal{D}_1$ is pseudorandom. The same reduction also works in the symmetric case. □

B Proofs for the Generic Hardness results

In this section, we give the remaining proofs for the results on the $\mathcal{D}_{\ell,k}$-MDDH assumption in generic m-linear groups from Section 3.3. We refer to reader to e.g. [11] for necessary background on the algebraic material such as polynomial rings, ideals, Gröbner bases, varieties and irreducibility used in this section. Note that in this paper irreducibility is *not* implicit in the definition of a variety.

Recall that our setup is that $\mathcal{D}_{\ell,k}$ is a matrix distribution which outputs $a_{i,j} = \mathfrak{p}_{i,j}(\vec{t})$ for uniform $\vec{t} \in \mathbb{Z}_q^d$ and possibly multivariate *polynomials* $\mathfrak{p}_{i,j}$, whose degree does not depend on λ and hence not on q. The distributions $([\mathbf{A}], [\vec{z}] = [\mathbf{A}\vec{\omega}])$ respectively $([\mathbf{A}], [\vec{z}] = [\vec{u}])$ for $\mathbf{A} \leftarrow \mathcal{D}_{\ell,k}, \vec{\omega} \leftarrow \mathbb{Z}_q^k, \vec{u} \leftarrow \mathbb{Z}_q^\ell$ are denoted by $\mathcal{D}^0$ respectively $\mathcal{D}^1$. In order to describe all of these data, we consider the polynomial ring $\mathcal{R} = \mathbb{Z}_q[\vec{A}, \vec{Z}, \vec{T}, \vec{W}]$, introducing formal variables $\vec{A} = A_{1,1}, \ldots, A_{\ell,k}$ to describe the matrix $\mathbf{A}$, $\vec{Z} = Z_1, \ldots, Z_\ell$ to describe the vector $\vec{z}$, $\vec{T} = T_1, \ldots T_d$ for some d to describe the underlying t's used to sample the $a_{i,j}$'s via $a_{i,j} = \mathfrak{p}_{i,j}(\vec{t})$, and formal variables $\vec{W} = W_1, \ldots, W_k$ to describe $\vec{\omega}$ (which only appears in $\mathcal{D}^0$). Note that we shorthand write $\vec{A}$ for the collection of all $A_{i,j}$'s if the structure as a matrix is not crucial. Furthermore, we write $\mathbf{A} = \mathfrak{p}(\vec{t})$ or $\vec{a} = \vec{\mathfrak{p}}(\vec{t})$, meaning that $a_{i,j} = \mathfrak{p}_{i,j}(\vec{t})$. We further consider the polynomial subring $\mathcal{S} = \mathbb{Z}_q[\vec{A}, \vec{Z}] \subset \mathcal{R}$ to describe the publicly known expressions. We can now encode our distributions $\mathcal{D}^0$ and $\mathcal{D}^1$ by polynomials in the following way: let $\mathfrak{f}_{i,j} = A_{i,j} - \mathfrak{p}_{i,j}(\vec{T})$ and $\mathfrak{g}_i = Z_i - \sum_j \mathfrak{p}_{i,j}(\vec{T}) W_j$. Let G_0 be the set of all $\mathfrak{f}$'s and $\mathfrak{g}$'s, whereas G_1 only consists of the $\mathfrak{f}$'s, but not the $\mathfrak{g}$'s. The generators G_b span the ideals $\mathcal{I}_b$ over $\mathcal{R}$, which encode all the relations in $\mathcal{D}^b$ for $b \in \{0, 1\}$. Of course, $\mathcal{I}_1 \subset \mathcal{I}_0$.

We consider $\mathcal{J}_b = \mathcal{I}_b \cap \mathcal{S}$, which are ideals in $\mathcal{S}$ encoding the relations between the known data. We will show that $(\mathcal{J}_b)_{\leq m}$, where $_{\leq m}$ denotes restriction to total degree at most m, captures exactly what can

be generically computed by an adversary performing only polynomially many group and m-linear pairing operations:

B.1 Proof of Theorem 3

Let $\mathcal{D}_{\ell,k}$ be a matrix distribution with polynomial defining equations and $\mathcal{I}_0, \mathcal{I}_1$ be as above. Then the $\mathcal{D}_{\ell,k}$-MDDH assumption holds in generic m-linear groups if and only if $\mathcal{J}_0$ and $\mathcal{J}_1$ are equal up to total degree m, i.e. $(\mathcal{J}_0)_{\leq m} = (\mathcal{J}_1)_{\leq m}$.

Proof. The proof is analogous to the one from [2, 9], apart from being stated more algebraically. Let D be a ppt distinguisher with input from $\mathcal{D}^b$ for either $b = 0$ or $b = 1$. Let $\kappa = \mathrm{poly}(\lambda)$ be an upper bound on the number of D's oracle queries and initial input group elements. We will replace the oracles D has access to, show that this replacement can only be detected with negligible probability and show that D's advantage with the replaced oracles is zero.

Our replacement of D's oracles is as follows: We replace (the random representation of) $\mathbb{G}$ and its associated oracles by (a random representation[10] of) the quotient $Q = \mathcal{R}/\mathcal{I}_b$. Similarly $\mathbb{G}_T$ is replaced by an isomorphic copy Q' of $\mathcal{R}/\mathcal{I}_b$ (with another random representation independent from the one for $\mathbb{G}$). The oracle for e is replaced by an oracle computing the product in Q and outputting the (representation of the) associated element in Q'. The initial elements $[a_{i,j}]$ respectively $[z_i]$ are replaced by $\pi(A_{i,j}) \in Q$ respectively $\pi(Z_i) \in Q$, where π respectively π' denotes the projection $\pi\colon \mathcal{R} \to Q$ respectively $\pi'\colon \mathcal{R} \to Q'$. The generators g and g_T are replaced by $\pi(1) \in Q$ and $\pi'(1) \in Q'$. The representations of Q and Q' are as usual defined on demand by keeping a list of all elements queried so far and choosing random representations for new elements; queries with representations as input that have not been previously defined produce an invalid answer $\perp$, as do queries using the wrong isomorphic copy and/or mixing them. Note that we assume here that in the random group model the representations are sufficiently long, say a generous $\geq 5 \log q$, such that representations are hard to guess and the sets of representations for G and G_T are disjoint with overwhelming probability.

By Buchberger's First Criterion [11], the given generating set G_b is actually a Gröbner basis with respect to any lexicographic ordering, where any Z_i's are larger than any $A_{i,j}$'s and both are larger than any T_i's or W_i's. We identify elements from $\mathcal{R}/\mathcal{I}_b$ by their remainders modulo G_b. Note that computing this remainder just means replacing any occurrence of $A_{i,j}$ by $\mathfrak{p}_{i,j}$ and, if $b = 0$, additionally replacing Z_i by $\sum_j \mathfrak{p}_{i,j} W_j$.

After D has run, we sample $\vec{t} \leftarrow \mathbb{Z}_q^d, \vec{\omega} \leftarrow \mathbb{Z}_q^k, \vec{u} \leftarrow \mathbb{Z}_q^\ell$. For any remainder $\mathfrak{h} \in Q$, define $\mathrm{ev}(\mathfrak{h})$ as $\mathrm{ev}(\mathfrak{h}) = [\mathfrak{h}(0, \vec{u}, \vec{t}, \vec{\omega})] \in \mathbb{G}$, where we plug in $\vec{u}$ for $\vec{Z}$, $\vec{t}$ for $\vec{T}$ and $\vec{\omega}$ for $\vec{W}$. Note that there are no $A_{i,j}$'s in $\mathfrak{h}$ and in the case $b = 0$ no Z_i's occur either. For $\mathfrak{h}' \in Q'$ we define $\mathrm{ev}(\mathfrak{h}') \in \mathbb{G}_T$ analogously.

Since D can only apply e in Q, but not in Q', any element seen in Q by D can be written as a sum of elements initially presented to D. Elements seen in Q' can be written as sums of m-fold products of such elements. So let $\mathfrak{k}_1, \ldots, \mathfrak{k}_r \in \mathcal{S}_{\leq 1}$ and $\mathfrak{k}'_1, \ldots, \mathfrak{k}'_{r'} \in \mathcal{S}_{\leq m}$ with $r + r' \leq \kappa$ be the elements constructed by D. Let $\mathfrak{h}_i := \mathfrak{k}_i \bmod \mathcal{I}_b \in Q$ and $\mathfrak{h}'_i := \mathfrak{k}'_i \bmod \mathcal{I}_b \in Q'$. The distinct elements among the $\mathfrak{h}_i$ and $\mathfrak{h}'_i$ are exactly the distinct elements from Q respectively Q' seen by D, whereas the $\mathfrak{k}_i$ and $\mathfrak{k}'_i$ keep track of how D constructed those. Note that the $\bmod \mathcal{I}_b$ map need not be injective on $\mathcal{S}_{\leq m}$.

Since computing $\bmod \mathcal{I}_b$ is just a replacement of each $A_{i,j}$ and possibly Z_i by a polynomial of degree at most $\deg + 1$, the total degree of all remainders $\mathfrak{h}_i$ and $\mathfrak{h}'_i$ is bounded by the constant $(\deg + 1)^m$, where deg is the upper bound on the total degree of the $\mathfrak{p}_{i,j}$, which is independent of the security parameter λ by assumption. Let $\mathtt{Good}$ denote the event that for all $\mathfrak{h}_i \neq \mathfrak{h}_j$ we have $\mathrm{ev}(\mathfrak{h}_i) \neq \mathrm{ev}(\mathfrak{h}_j)$ and for all $\mathfrak{h}'_i \neq \mathfrak{h}'_j$ we have $\mathrm{ev}(\mathfrak{h}'_i) \neq \mathrm{ev}(\mathfrak{h}'_j)$. By construction, if $\mathtt{Good}$ occurs, the view of D with the replaced oracles is identical to the view if D would have had access to the original oracles. Since each such equality $\mathrm{ev}(\mathfrak{h}_i) = \mathrm{ev}(\mathfrak{h}_j)$ or $\mathrm{ev}(\mathfrak{h}'_i) = \mathrm{ev}(\mathfrak{h}'_j)$ is a non-zero polynomial equation of total degree at most $(\deg + 1)^m$ in uniformly chosen unknowns from $\mathbb{Z}_q$, each one holds only with probability at most $\frac{(\deg+1)^m}{q} = \mathrm{negl}(\lambda)$. Since there are only polynomially many pairs $i \neq j$, $\mathtt{Good}$ occurs with overwhelming probability of at least $1 - \frac{\kappa(\kappa-1)(\deg+1)^m}{2q}$.

[10] Strictly speaking, only those polynomially many elements ever appearing even have a well-defined representation. Note that Q is infinite.

Furthermore, D's view can only depend on b if we have $\mathfrak{k}_i - \mathfrak{k}_j \equiv 0 \bmod \mathcal{I}_0$ but $\mathfrak{k}_i - \mathfrak{k}_j \not\equiv 0 \bmod \mathcal{I}_1$ (or the analogous in Q') for some elements $\mathfrak{k}_i, \mathfrak{k}_j$ constructed by D. We know that any $\mathfrak{k}_i$ or $\mathfrak{k}'_i$ is in $\mathcal{S}_{\leq m}$. So, since $\mathcal{I}_0 \cap \mathcal{S}_{\leq m} = (\mathcal{J}_0)_{\leq m} = (\mathcal{J}_1)_{\leq m} = \mathcal{I}_1 \cap \mathcal{S}_{\leq m}$, D's view (with the replaced oracles) does not depend on b.

For the other direction of the theorem, note that if there exists $\mathfrak{k} \in (\mathcal{J}_0)_{\leq m} \setminus (\mathcal{J}_1)_{\leq m}$ then it is easy to construct a ppt distinguisher D that computes $h = [\mathfrak{k}(a_{i,j}, z_i)]_T \in \mathbb{G}_T$. If $b = 0$, we always have $h = [0]_T$ whereas if $b = 1$, we have $h = [0]_T$ only with probability at most $\frac{(\deg+1)^m}{q} = \mathrm{negl}(\lambda)$. □

The ideals $\mathcal{J}_0$ and $\mathcal{J}_1$ can be computed from $\mathcal{I}_0$ and $\mathcal{I}_1$ using elimination theory. If we use Gröbner bases for that, the condition $(\mathcal{J}_0)_{\leq m} = (\mathcal{J}_1)_{\leq m}$ can be rephrased as follows:

Lemma 20. *Let notation be as before and $m > 0$. Let $<$ be an elimination order on the monomials of $\mathcal{R}$ such that any monomial containing any T_i or W_i is larger than any monomial from $\mathcal{S}$. Further assume that, restricted to the monomials of $\mathcal{S}$, $<$ sorts by total degree first. Let H_0 respectively H_1 be* reduced *Gröbner bases for $\mathcal{I}_0$ respectively $\mathcal{I}_1$ w.r.t. $<$. Then the following are equivalent:*

1. $(\mathcal{J}_0)_{\leq m} = (\mathcal{J}_1)_{\leq m}$
2. $H_0 \cap \mathcal{S}_{\leq m} = H_1 \cap \mathcal{S}_{\leq m}$
3. $H_0 \cap \mathcal{S}_{\leq m}$ *does not involve any Z_i's.*
4. *There exists a not necessarily reduced Gröbner basis H'_0 for $\mathcal{I}_0$ such that $H'_0 \cap \mathcal{S}_{\leq m}$ does not involve any Z_i's.*

Proof. First, note that by the elimination theorem of Gröbner bases [11], $\mathcal{J}_b$ is an ideal over $\mathcal{S}$ with reduced Gröbner basis $H_b \cap \mathcal{S}$.

(1) ⇒ (2) : Assume $(\mathcal{J}_0)_{\leq m} = (\mathcal{J}_1)_{\leq m}$. Let $\mathfrak{h} \in H_0 \cap \mathcal{S}_{\leq m}$, but assume towards a contradiction $\mathfrak{h} \notin H_1 \cap \mathcal{S}_{\leq m}$. Since $\mathfrak{h} \in \mathcal{I}_1 \cap \mathcal{S}_{\leq m}$, there must be some $\mathfrak{k} \in H_1 \cap \mathcal{S}, \mathfrak{k} \neq \mathfrak{h}$ such that the leading term of $\mathfrak{k}$ divides the leading term of $\mathfrak{h}$. By assumption, $<$ sorts by total degree first, so the total degree of $\mathfrak{k}$ is at most m. Hence $\mathfrak{k} \in \mathcal{I}_0 \cap \mathcal{S}_{\leq m}$ with leading term diving that of $\mathfrak{h}$, contradicting the reducedness of $H_0 \cap \mathcal{S}$. The other inclusion $H_1 \cap \mathcal{S}_{\leq m} \subset H_0 \cap \mathcal{S}_{\leq m}$ is analogous.

(2) ⇒ (3) : H_1 does not involve any Z_i's, since the generating set G_1 does not.

(3) ⇒ (4) : Obvious.

(4) ⇒ (1) : Assume $H'_0 \cap \mathcal{S}_{\leq m}$ does not involve any Z_i. We first show that for any $\mathfrak{h} \in H'_0 \cap \mathcal{S}_{\leq m}$ we have $\mathfrak{h} \in \mathcal{I}_1$. To see this, write $\mathfrak{h} = \sum_{i,j} \mathfrak{c}_{i,j}\mathfrak{f}_{i,j} + \sum_i \mathfrak{d}_i\mathfrak{g}_i$ as a linear combination in our original generators G_0 with polynomial coefficients $\mathfrak{c}_{i,j}, \mathfrak{d}_i \in \mathcal{R}$. Plugging in 0 for all W_i's and Z_i's into this equation does not affect $\mathfrak{h}$ by assumption and eliminates all $\mathfrak{g}_i$, so we obtain $\mathfrak{h} = \sum_{i,j} \mathfrak{c}'_{i,j}\mathfrak{f}_{i,j}$ for some $\mathfrak{c}'_{i,j}$ showing $\mathfrak{h} \in \mathcal{I}_1$.
Now let $\mathfrak{k} \in \mathcal{I}_0 \cap \mathcal{S}_{\leq m} = (\mathcal{J}_0)_{\leq m}$ be arbitrary. Since $H'_0 \cap \mathcal{S}$ is a Gröbner basis w.r.t to $<$, which sorts by total degree first, we have $\mathfrak{k} = \sum_i \mathfrak{e}_i\mathfrak{h}_i$ for some $\mathfrak{e}_i \in \mathcal{S}$ and $\mathfrak{h}_i \in H'_0 \cap \mathcal{S}_{\leq \deg \mathfrak{k}}$. Since we have shown that all the $\mathfrak{h}_i$ that appear here are in $\mathcal{I}_1$, we have $\mathfrak{k} \in \mathcal{I}_1$, showing $(\mathcal{J}_0)_{\leq m} \subset (\mathcal{J}_1)_{\leq m}$. The other inclusion is trivial. □

B.2 Proof of Theorem 4 and Generalizations

Theorem 4 will follow as a corollary from the following lemma, which is a generalization to non-linear $\mathfrak{p}_{i,j}$ and non-irreducible $\mathfrak{d}$:

Lemma 21. *Let notation be as before. We assume that $\ell = k + 1$ and $\mathbf{A}$ can be full rank for some values of $\vec{t}$. Let $\mathfrak{d}$ be the determinant of $(\mathfrak{p}(\vec{T})\|\vec{Z})$ as a polynomial in $\vec{Z}, \vec{T}$ and consider the ideal $\mathcal{J} := \mathcal{I}_0 \cap \mathbb{Z}_q[\vec{A}, \vec{Z}, \vec{T}]$ over $\mathbb{Z}_q[\vec{A}, \vec{Z}, \vec{T}]$. Then there exists a unique (up to scalar) decomposition $\mathfrak{d} = \mathfrak{c} \cdot \mathfrak{d}_0$ over $\mathbb{Z}_q$, where $\mathfrak{c}$ only involves the $\vec{T}$ and $\mathfrak{d}_0$ is irreducible over the algebraic closure $\overline{\mathbb{Z}_q}$. Furthermore, $\mathcal{J}$ is generated by G_1 and $\mathfrak{d}_0$.*

Proof. Since $\mathbf{A}$ can be full rank, there exists some $\vec{z}, \vec{t}$ with $\mathfrak{d}(\vec{z}, \vec{t}) \neq 0$, so $\mathfrak{d}$ is not the zero polynomial. For the existence and uniqueness of $\mathfrak{c}$ and $\mathfrak{d}_0$, consider the (up to scalar) unique decomposition $\mathfrak{d} = \mathfrak{c}_1^{e_1}\mathfrak{c}_2^{e_2} \cdots \mathfrak{c}_s^{e_s}$ of

$\mathfrak{d}$ into distinct irreducible polynomials $\mathfrak{c}_i$ in $\overline{\mathbb{Z}_q}[\vec{Z},\vec{T}]$. Since $\mathfrak{d}$ is linear in the Z_i's, only one factor, w.l.o.g. $\mathfrak{c}_s$ with $e_s = 1$, can contain any of the Z_i's. Note that this implies that $\mathfrak{c}_s$ is linear in the Z_i's as well. So we have the up to scalar unique decomposition $\mathfrak{d}(\vec{Z},\vec{T}) = \mathfrak{c}(\vec{T})\mathfrak{d}_0(\vec{Z},\vec{T})$ with $\mathfrak{d}_0 = \mathfrak{c}_s$ and $\mathfrak{c} = \mathfrak{c}_1^{e_1}\cdots\mathfrak{c}_{s-1}^{e_{s-1}}$, which has the desired properties, provided that $\mathfrak{d}_0$ and $\mathfrak{c}$ actually have coefficients in the base field $\mathbb{Z}_q$ rather than $\overline{\mathbb{Z}_q}$.

To show the latter, write $\mathfrak{d} = \sum_i \mathfrak{a}_i Z_i$ with $\mathfrak{a}_i \in \mathbb{Z}_q[\vec{T}]$. By construction, $\mathfrak{c}$ divides $\mathfrak{d}$ and $\mathfrak{c}$ involves no $\vec{Z}$. Plugging in $Z_i = 1$ for $i = i_0$ and $Z_i = 0$ for $i \neq i_0$ into $\mathfrak{d} = \mathfrak{c}\cdot\mathfrak{d}_0$ shows that $\mathfrak{c}$, and consequently $\mathfrak{c}_j^{e_j}$, divides $\mathfrak{a}_{i_0}$. So, for all $1 \leq i \leq \ell, 1 \leq j \leq s-1$ we have $\mathfrak{a}_i = \mathfrak{c}_j^{e_j}\cdot\mathfrak{b}_{i,j}$ for some $\mathfrak{b}_{i,j} \in \overline{\mathbb{Z}_q}[\vec{T}]$ and indeed $\mathfrak{c}$ is nothing but the gcd of the $\mathfrak{a}_i$. Since $\mathfrak{a}_i \in \mathbb{Z}_q[\vec{T}]$, it follows that $\sigma(\mathfrak{a}_i) = \mathfrak{a}_i = \sigma(\mathfrak{c}_j)^{e_j}\cdot\sigma(\mathfrak{b}_{i,j})$, where σ is the (coefficient-wise) Frobenius. So $\sigma(\mathfrak{c}_j)^{e_j}$ divides each $\mathfrak{a}_i$, hence every Frobenius-conjugate must appear (up to scalar) in the decomposition $\mathfrak{c} = \mathfrak{c}_1^{e_1}\cdots\mathfrak{c}_{s-1}^{e_{s-1}}$ with the same multiplicity. This shows that we can choose $\mathfrak{c} \in \mathbb{Z}_q[\vec{T}]$ after adjusting scalars. It follows that $\mathfrak{d}_0 = \frac{\mathfrak{d}}{\mathfrak{c}}$ is also in the base field.

For the second part of the lemma, we first observe that both ideals $\mathcal{I}_0$ and $\mathcal{I}_1$ are radical: Since they can be generated by polynomials of the form $A_{i,j} - \mathfrak{p}_{i,j}(\vec{T}), Z_i - \mathfrak{q}_i(\vec{T},\vec{W})$ expressing one set of variables as functions of another disjoint set of variables, the quotient $\mathcal{R}/\mathcal{I}_0$ respectively $\mathcal{R}/\mathcal{I}_1$ is isomorphic to $\mathbb{Z}_q[\vec{T},\vec{W}]$ respectively $\mathbb{Z}_q[\vec{Z},\vec{T},\vec{W}]$. Since these quotients have no nilpotent elements, the ideals $\mathcal{I}_0, \mathcal{I}_1$ are radical. It follows that $\mathcal{J}$ is radical, since intersection with a polynomial subring preserves being radical. Since $\mathfrak{d}_0$ is irreducible, the quotient $\mathbb{Z}_q[\vec{A},\vec{Z},\vec{T}]/(G_1,\mathfrak{d}_0)$, which is isomorphic to $\mathbb{Z}_q[\vec{Z},\vec{T}]/(\mathfrak{d}_0)$, contains no nilpotent elements, hence the ideal generated by $\mathcal{I}_1$ and $\mathfrak{d}_0$ in $\mathbb{Z}_q[\vec{A},\vec{Z},\vec{T}]$ is radical. It thus suffices to consider the corresponding varieties (all varieties are over the algebraic closure $\overline{\mathbb{Z}_q}$) $V(G_1,\mathfrak{d}_0)$ and $V(\mathcal{J})$ by the Nullstellensatz. Let $V(\mathcal{I}_1)$ be the variety associated to $\mathcal{I}_1$. By the Closure Theorem [11], the variety $V(\mathcal{J})$ associated to $\mathcal{J}$ is given by the Zariski closure of $\{(\vec{a},\vec{z},\vec{t}) \in V(\mathcal{I}_1) \mid \exists\vec{\omega}, \text{ s.t. } z_i = \sum_j \omega_j a_{i,j}\}$. Let us start by showing $V(G_1,\mathfrak{d}_0) \subset V(\mathcal{J})$:

If for some value of $\vec{t}$, $\mathfrak{c}(\vec{t}) = 0$, then $\det(\mathfrak{p}(\vec{t})\|\vec{z}) = 0$ for all values of $\vec{z}$, hence $\mathfrak{p}(\vec{t})$ has rank $< k$. Consider the variety V_{bad} of all $(\vec{a},\vec{z},\vec{t}) \in V(\mathcal{I}_1)$ such that $\mathbf{A} = (a_{i,j})$ has rank $< k$, which is indeed an algebraic set (consider $\det(\mathbf{A}\|\vec{e}_i) = 0$ for canonical basis vectors $\vec{e}_i$) and $V_{\text{bad}} \supset V(\mathfrak{c},\mathcal{I}_1)$. Outside of this bad set, $\mathbf{A} = \mathfrak{p}(\vec{t})$ has full rank k and hence there exists $\vec{\omega}$ such that $\vec{z} = \mathbf{A}\cdot\vec{\omega}$ if and only if $\det(\mathbf{A}\|\vec{z}) = 0$, or equivalently, since $\mathfrak{c}(\vec{t}) \neq 0$, $\mathfrak{d}_0(\vec{z},\vec{t}) = 0$. It follows that $V(G_1,\mathfrak{d}_0) \setminus V_{\text{bad}} \subset V(\mathcal{J})$. By the same argument as in the previous paragraph, since $\mathfrak{d}_0$ is irreducible over $\overline{\mathbb{Z}_q}$, the quotient $\overline{\mathbb{Z}_q}[\vec{A},\vec{Z},\vec{T}]/(G_1,\mathfrak{d}_0) \cong \overline{\mathbb{Z}_q}[\vec{Z},\vec{T}]/(\mathfrak{d}_0)$ has no zero divisors and so $V(G_1,\mathfrak{d}_0)$ is irreducible. Since $(\vec{a},\vec{0},\vec{t}) \in V(G_1,\mathfrak{d}_0)$ for any $\vec{t}$ with $\mathfrak{p}(\vec{t})$ full rank, we have $V_{\text{bad}} \not\supseteq V(G_1,\mathfrak{d}_0)$. From this and the irreducibility of $V(G_1,\mathfrak{d}_0)$, we can then deduce that the Zariski closure of $V(G_1,\mathfrak{d}_0) \setminus V_{\text{bad}} \subset V(\mathcal{J})$ is all of $V(G_1,\mathfrak{d}_0)$, so we have $V(G_1,\mathfrak{d}_0) \subset V(\mathcal{J})$

For the other direction, consider $(\vec{a},\vec{z},\vec{t})$ such that $\vec{a} = \vec{\mathfrak{p}}(\vec{t})$ and there exists $\vec{\omega}$ with $z_i = \sum_j \omega_j a_{i,j}$. We need to show $\mathfrak{d}_0(\vec{z},\vec{t}) = 0$. For this, note that $\det(\mathfrak{p}(\vec{T})\|\sum_j W_j\mathfrak{p}_{i,j}(\vec{T}))$ is the zero polynomial. So $\mathfrak{d}(\sum_j W_j\mathfrak{p}_{i,j}(\vec{T}),\vec{T}) = \mathfrak{c}(\vec{T})\cdot\mathfrak{d}_0(\sum_j W_j\mathfrak{p}_{i,j}(\vec{T}))$ is the zero polynomial. Since $\mathfrak{c}(\vec{T})$ is not the zero polynomial, as otherwise $\mathfrak{d}(\vec{Z},\vec{T})$ would be the zero polynomial, we have that $\mathfrak{d}_0(\sum_j W_j\mathfrak{p}_{i,j}(\vec{T}),\vec{T})$ is the zero polynomial. It follows that $\mathfrak{d}_0(\vec{z},\vec{t}) = \mathfrak{d}_0(\sum_j \omega_j\mathfrak{p}_{i,j}(\vec{t}),\vec{t}) = 0$, finishing the proof of $V(G_1,\mathfrak{d}_0) \supset V(\mathcal{J})$.

□

This lemma allows us to easily prove Theorem 4, which states:

Let $\ell = k+1$ and $\mathcal{D}_{k+1,k}$ be a matrix distribution, which outputs matrices $\mathbf{A} = \mathfrak{p}(\vec{t})$ for uniform $\vec{t}$. Let $\mathfrak{d}$ be the determinant of $(\mathfrak{p}(\vec{T})\|\vec{Z})$ as a polynomial in $\vec{Z},\vec{T}$.

1. If the matrices output by $\mathcal{D}_{k+1,k}$ always have full rank (not just with overwhelming probability), even for t_i from the algebraic closure $\overline{\mathbb{Z}_q}$, then $\mathfrak{d}$ is irreducible over $\overline{\mathbb{Z}_q}$.
2. If all $\mathfrak{p}_{i,j}$ have degree at most 1, $\mathfrak{d}$ is irreducible over $\overline{\mathbb{Z}_q}$ and the total degree of $\mathfrak{d}$ is $k+1$, then the $\mathcal{D}_{k+1,k}$-MDDH assumption holds in *generic* k-linear groups.

Proof. Let notation be as in the lemmas above.
(1): If $\mathfrak{c}$ is non-constant, it would have some roots $(\vec{z},\vec{t})$ in $\overline{\mathbb{Z}_q}$. At these roots $\mathfrak{p}(\vec{t})$ can't have full rank, since

$\det(\mathfrak{p}(\vec{t})\|\vec{z}) = 0$ for all $\vec{z}$. Hence $\mathfrak{d} = \mathfrak{d}_0$, which is irreducible over $\overline{\mathbb{Z}_q}$.
(2): W.l.o.g. we may assume that $\vec{\mathfrak{p}}$ is injective (otherwise we drop some T-variables), so we can express the T_i's as linear polynomials in the $A_{i,j}$'s. Computing a Gröbner basis (for an appropriate elimination ordering) for $\mathcal{J}_0 = \mathcal{J} \cap \mathcal{S}$ from $\mathcal{J}$ just means expressing all T_i's by $A_{i,j}$'s. Since $\mathcal{J}$ is generated by $\mathfrak{d} = \mathfrak{d}_0$ and G_1 by the above Lemma 21, a Gröbner basis for $\mathcal{J}_0$ is just given by G_1 and $\mathfrak{d}$, expressed by the $A_{i,j}$'s. Since this invertible linear variable substitution does not change total degree, the theorem follows. □

C Proof of Theorem 10

The proof is rather technical because we need an explicit construction of a sequence of subspaces with special properties. The key idea is using a consequence of Lemma 9: for any nontrivial subspace $U \subset \mathbb{Z}_q^k$, $\dim(f_0(U)+f_1(U)) > \dim U$, and for any nontrivial subspace $V \subset f_0(\mathbb{Z}_q^k) \cap f_1(\mathbb{Z}_q^k)$, $\dim(f_0^{-1}(V)+f_1^{-1}(V)) > \dim V$. This allows us to build a sequence of subspaces with strictly increasing dimensions having some interesting properties. We will then use these subspaces to build the bases claimed in the theorem.

Consider the following sequences of subspaces, for a suitable value of $m \in \mathbb{Z}$

$$U_1 \subset U_2 \subset \cdots \subset U_m = \mathbb{Z}_q^k\ ; \qquad V_1 \subset V_2 \subset \cdots \subset V_m \subset \mathbb{Z}_q^{k+1}$$

such that $V_i = f_0(U_i) \cap f_1(U_i)$ and $U_{i-1} = f_0^{-1}(V_i) \cap f_1^{-1}(V_i)$. The sequences are well defined because we know that $V_i \subset f_0(U_i)$ and $U_{i-1} \subset f_0^{-1}(V_i)$, and then $U_{i-1} \subset f_0^{-1}(V_i) \subset f_0^{-1}(f_0(U_i)) = U_i$, since f_0 is injective, and similarly $V_{i-1} \subset f_0(U_{i-1}) \subset f_0(f_0^{-1}(V_i)) \subset V_i$. On the other hand, from the injectivity of the maps $\dim U_i = \dim f_0(U_i) = \dim f_1(U_i)$ and $\dim V_i = \dim f_0^{-1}(V_i) = \dim f_1^{-1}(V_i)$. Now, by Lemma 9 we know that $f_0(U_i) \neq f_1(U_i)$, if U_i is nontrivial, and similarly $f_0^{-1}(V_i) \neq f_1^{-1}(V_i)$, if V_i is nontrivial. Therefore, if $\dim V_i > 0$ then

$$\dim U_{i-1} = \dim(f_0^{-1}(V_i) \cap f_1^{-1}(V_i)) < \dim V_i$$

and if $\dim U_i > 0$ then

$$\dim V_i = \dim(f_0(U_i) \cap f_1(U_i)) < \dim U_i$$

On the other hand, since $f_0^{-1}(V_i) \subset U_i$ and $f_1^{-1}(V_i) \subset U_i$ then $f_0^{-1}(V_i) + f_1^{-1}(V_i) \subset U_i$, and analogously $f_0(U_i) + f_1(U_i) \subset V_{i+1}$. Putting all equations together, if U_i is nontrivial,

$$1 \leq \dim U_i - \dim V_i = \dim U_i - \dim(f_0(U_i) \cap f_1(U_i)) = \dim(f_0(U_i) + f_1(U_i)) - \dim U_i \leq \dim V_{i+1} - \dim U_i$$

and similarly, if V_i is nontrivial, $1 \leq \dim V_i - \dim U_{i-1} \leq \dim U_i - \dim V_i$. But

$$\dim U_m - \dim V_m = \dim(f_0(U_m) + f_1(U_m)) - \dim U_m = \dim \mathbb{Z}_q^{k+1} - \dim \mathbb{Z}_q^k = 1$$

and then all the equalities hold. As a consequence, if k is even, taking $k = 2m$ we have shown that $\dim V_i = 2i - 1$ and $\dim U_i = 2i$. Otherwise, we take $k = 2m - 1$ and $\dim V_i = 2i - 2$ and $\dim U_i = 2i - 1$ (hence, V_1 is trivial here).

In addition, the previous equalities of dimensions imply the corresponding equalities of subspaces $U_i = f_0^{-1}(V_i) + f_1^{-1}(V_i)$ and $V_{i+1} = f_0(U_i) + f_1(U_i)$, which in particular mean that a generating set of U_i can be constructed by computing the preimages of a generating set in V_i for both f_0 and f_1 (these preimages always exist for vectors in any $V_i \subset V_m = f_0(U_m) \cap f_1(U_m)$). Similarly, we can build a generating set of V_{i+1} by applying f_0 and f_1 to a generating set of U_i. We will also use the fact that $\mathbb{Z}_q^{m+1} = f_0(U_m) + f_1(U_m)$ to complete a basis of $\mathbb{Z}_q^{k+1}$.

At this point, we have constructed two sequences of subspaces which dimensions grow regularly, and we can build bases of the spaces by cleverly picking vectors from them. We consider separately the cases k even and k odd.

For $k = 2m$, we know that $\dim V_1 = 1$. Let $\vec{y} \in \mathbb{Z}_q^{k+1}$ be a nonzero vector in V_1. Then, $\vec{x}_0 = f_0^{-1}(\vec{y})$ and $\vec{x}_1 = f_1^{-1}(\vec{y})$ form a basis of U_1, since it is a generating set and $\dim U_1 = 2$. Similarly, we build a generating set $\{f_1(\vec{x}_0), f_0(\vec{x}_0), f_1(\vec{x}_1), f_0(\vec{x}_1)\}$ of V_2, but actually $f_0(\vec{x}_0) = f_1(\vec{x}_1) = \vec{y}$. Since $\dim V_2 = 3$ we know that

the three different vectors form a basis. Observe that we can write it as $\{(f_1 \circ f_0^{-1})(\vec{y}), \vec{y}, (f_0 \circ f_1^{-1})(\vec{y})\}$, where f_0^{-1} (and similarly f_1^{-1}) denotes here the inverse map of f_0 restricted to its image $f_0(\mathbb{Z}_q^k)$, so it is well defined on any subspace V_i. Now, computing the preimages for f_0 and f_1 and removing the repeated vectors we can build a basis of U_2. Following the same procedure iteratively, we can build the bases

$$B_1 = \{(f_0^{-1} \circ f_1)^{m-1}(\vec{x}_0), \ldots, (f_0^{-1} \circ f_1)(\vec{x}_0), \vec{x}_0, \vec{x}_1, (f_1^{-1} \circ f_0)(\vec{x}_1), \ldots, (f_1^{-1} \circ f_0)^{m-1}(\vec{x}_1)\}$$

and

$$B_2 = \{(f_1 \circ f_0^{-1})^m(\vec{y}), \ldots, (f_1 \circ f_0^{-1})(\vec{y}), \vec{y}, (f_0 \circ f_1^{-1})(\vec{y}), \ldots, (f_0 \circ f_1^{-1})^m(\vec{y})\}$$

of $\mathbb{Z}_q^k$ and $\mathbb{Z}_q^{k+1}$, respectively, with the property that the images of the vectors in B_1 by f_0 are exactly the last k vectors in B_2, and the images of the vectors in B_1 by f_1 are exactly the first k vectors in B_2. This is the same as saying that f_0 and f_1 are represented in those bases by the matrices $\mathbf{J_0}$ and $\mathbf{J_1}$, respectively.

The proof for the odd case $k = 2m - 1$ proceeds similarly, but starting from a nonzero vector $\vec{x} \in U_1$, computing the two images $\vec{y}_0 = f_0(\vec{x})$ and $\vec{y}_1 = f_1(\vec{x})$, and then applying the same iterative procedure as before to obtain the bases

$$B_1 = \{(f_0^{-1} \circ f_1)^{m-1}(\vec{x}), \ldots, (f_0^{-1} \circ f_1)(\vec{x}), \vec{x}, (f_1^{-1} \circ f_0)(\vec{x}), \ldots, (f_1^{-1} \circ f_0)^{m-1}(\vec{x})\}$$

and

$$B_2 = \{(f_1 \circ f_0^{-1})^m(\vec{y}_1), \ldots, (f_1 \circ f_0^{-1})(\vec{y}_1), \vec{y}_1, \vec{y}_0, (f_0 \circ f_1^{-1})(\vec{y}_0), \ldots, (f_0 \circ f_1^{-1})^m(\vec{y}_0)\}$$

of $\mathbb{Z}_q^k$ and $\mathbb{Z}_q^{k+1}$, respectively, with exactly the same property as before.

D Subgroup Membership Proofs for 2-Lin

In this section we exemplify our approach from Section 6.1 for the 2-Lin case. Let

$$\mathbf{A} = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \\ 1 & 1 \end{pmatrix} = (\vec{u}_1, \vec{u}_2), \qquad \mathbf{A} \leftarrow \mathcal{L}_2,$$

and

$$[\mathbf{u}_3] = \begin{cases} [w_1\vec{u}_1 + w_2\vec{u}_2] & \text{binding key (soundness setting)} \\ [w_1\vec{u}_1 + w_2\vec{u}_2 - (0,0,1)^\top] & \text{hiding key (WI setting)} \end{cases},$$

for $w_1, w_2 \leftarrow \mathbb{Z}_q$. We exemplify our new approach to prove $[\Phi] \in \mathcal{L}_{\mathbf{A},\mathcal{PG}} \subset \mathbb{G}^3$. To simplify the notation we define $\vec{v} := \vec{u}_3 + (0,0,1)^\top$. With this notation, $[\iota'(x)] := [x\vec{v}]$.

Standard Groth-Sahai proof. In the standard approach, used for instance in [35], the prover will show that there are two values $r_1, r_2 \in \mathbb{Z}_q$ such that the following equations hold:

$$[r_1 a_1] = [\Phi_1] \tag{9}$$
$$[r_2 a_2] = [\Phi_2] \tag{10}$$
$$[r_1 + r_2] = [\Phi_3]. \tag{11}$$

Therefore, we are in the setting of multiscalar multiplication with $A_1 = \mathbb{Z}_q$ and $A_2 = \mathbb{G}$. The proof consists of the commitments to r_1, r_2, which are two vectors $[\vec{c}_{r_1}], [\vec{c}_{r_2}] \in \mathbb{G}^3$ such that

$$(\vec{c}_{r_1}, \vec{c}_{r_2}) = (\iota'(r_1), \iota'(r_2)) + \mathbf{A}\begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} = (r_1\vec{v}, r_2\vec{v}) + \mathbf{A}\begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}$$

and the vector

$$[\vec{\pi}_{(r_1,r_2)}] = \left[\left((a_1,0)\mathbf{S}^\top,\ (0,a_2)\mathbf{S}^\top, (1,1)\mathbf{S}^\top\right)\right] = ([s_{11}a_1], [s_{21}a_1], [s_{12}a_2], [s_{22}a_2], [s_{11}+s_{12}], [s_{21}+s_{22}]).$$

Therefore, in total, the proof requires 12 group elements.

To simulate the proof, we proceed as if we were proving that the equations

$$\begin{aligned}[r_1 a_1] &= [\delta\Phi_1] \\ [r_2 a_2] &= [\delta\Phi_2] \\ [r_1 + r_2] &= [\delta\Phi_3],\end{aligned}$$

are satisfied by the all zero witness, with the commitment to $\delta = 0$ being $\mathsf{com}'_{[\mathbf{U}],\vec{z}}(0; (w_1, w_2)^\top)$, which, in the witness indistinguishability setting, is equal to $[\iota'(1)] = [\vec{v}] = [\mathbf{A}\vec{w}]$.

New approach. To construct the proof, the prover needs to sample uniformly at random from the space $\mathcal{H} := \{\mathbf{H} \in \mathbb{Z}_q^{2\times 2} : \mathbf{H} + \mathbf{H}^\top = \mathbf{0}\}$. To sample $\mathbf{H} \leftarrow \mathcal{H}$, pick a random value $h \leftarrow \mathbb{Z}_q$ and define $\mathbf{H} = \begin{pmatrix} 0 & h \\ -h & 0 \end{pmatrix}$. The proof is then defined as:

$$[\mathbf{\Pi}] = [\vec{u}_3(r_1, r_2) + \mathbf{AH}] = \begin{pmatrix} [r_1 v_1] & [r_2 v_1 + h a_1] \\ [r_1 v_2 - a_2 h] & [r_2 v_2] \\ [r_1 v_3 - h] & [r_2 v_3 + h] \end{pmatrix}$$

The proof consists of 6 group elements, as claimed.

For simulation, we sample some $\mathbf{H}' \leftarrow \mathcal{H}$ as before and we define:

$$[\mathbf{\Pi}_{\mathrm{sim}}] = [\vec{\Phi}(w_1, w_2) + \mathbf{AH}'].$$

E Concrete Examples from the k-SCasc Assumption

As we promote the k-SCasc Assumption as a replacement of the k-Lin Assumption, we give two concrete instantiations of a KEM and a PRF based on it.

E.1 Key Encapsulation

We build a $\mathsf{KEM}_{\mathsf{Gen},\mathcal{SC}_k}$ from k-SCasc (Example 4).

- $\mathsf{Gen}(1^\lambda)$ runs $\mathcal{G} \leftarrow \mathsf{Gen}(1^\lambda)$ and picks $a \leftarrow \mathbb{Z}_q$. The public/secret-key is

$$pk = (\mathcal{G}, ([a]) \in \mathbb{G}), \quad sk = a \in \mathbb{Z}_q.$$

- Enc_{pk} picks $\vec{w} \leftarrow \mathbb{Z}_q^k$. The ciphertext/key pair is

$$[\vec{c}] = ([aw_1], [w_1 + aw_2] \ldots, [w_{k-1} + aw_k])^T \in \mathbb{G}^k, \quad [K] = [w_k] \in \mathbb{G}.$$

- $\mathsf{Dec}_{sk}([\vec{c}] \in \mathbb{G}^k)$ recomputes the key as

$$[K] = [\vec{x}^\top \vec{c}] \in \mathbb{G},$$

  where the transformation vector $\vec{x} \in \mathbb{Z}_q^k$ is computed from a as $x_i = \frac{(-1)^{k-i}}{a^{k-i}}$ (such that $\vec{x}^\top \mathbf{A}_0 = (0, \ldots, 0, 1)^T$ where $\mathbf{A}_0$ consists of the top k rows of matrix $\mathbf{A}$ from Example 4).

Security of $\mathsf{KEM}_{\mathsf{Gen},\mathcal{SC}_k}$ follows from Theorem 11. Note that the size of the public/secret key is constant, compared to linear (in k) for the k-Lin-based KEM [23, 45]. The ciphertext size remains the same, however.

E.2 Pseudo-Random Function

We build $\mathsf{PRF}_{\mathsf{Gen},\mathcal{SC}_k} = (\mathsf{Gen}, \mathsf{F})$ from k-SCasc.

- $\mathsf{Gen}(1^\lambda)$ runs $\mathcal{G} \leftarrow \mathsf{Gen}(1^\lambda)$ and picks $a_{i,j} \leftarrow \mathbb{Z}_q$ for $1 \le i \le n$, $1 \le j \le k$ and $\vec{h} \leftarrow \mathbb{Z}_q^k$. The secret-key is $K = ((a_{i,j}), \vec{h})$.
- $\mathsf{F}_K(x)$ computes

$$\mathsf{F}_K(x) = \left[\prod_{i:x_i=1} \mathbf{T}_i \cdot \vec{h} \right] \in \mathbb{G}^k,$$

where

$$\mathbf{T}_i = \begin{pmatrix} \frac{(-1)^{k-1}}{a_{i,1}^k} & \cdots & \frac{-1}{a_{i,1}^2} & \frac{1}{a_{i,1}} \\ \vdots & & \vdots & \vdots \\ \frac{(-1)^{k-1}}{a_{i,k}^k} & \cdots & \frac{-1}{a_{i,k}^2} & \frac{1}{a_{i,k}} \end{pmatrix} \in \mathbb{Z}_q^{k \times k},$$

where the transformation matrices $\mathbf{T}_{i,j}$ of $\mathbf{A}_{i,j} \leftarrow \mathcal{SC}_k$ are the row vectors of $\mathbf{T}_i$. Security of $\mathsf{PRF}_{\mathsf{Gen},\mathcal{SC}_k}$ follows from Theorem 12. Note that the size of the secret-key K is nk, compared to nk^2 for the k-Lin-based PRF [6].

Observe that if we add the restriction $a_{i,j} \neq 0$, we can rewrite $\mathbf{T}_i$ as

$$\mathbf{T}_i = - \begin{pmatrix} b_{i,1}^k & \cdots & b_{i,1}^2 & b_{i,1} \\ \vdots & & \vdots & \vdots \\ b_{i,k}^k & \cdots & b_{i,k}^2 & b_{i,k} \end{pmatrix},$$

where now $b_{i,j} = -\frac{1}{a_{i,j}}$ are random nonzero elements. If we associate $\vec{h} = (h_1, h_2, \ldots, h_k)$ to the polynomial $\mathfrak{h} = h_1 X^k + \ldots + h_{k-1} X^2 + h_k X \in \mathbb{Z}_q[X]$, then $\mathbf{T}_i \vec{h} = -(\mathfrak{h}(b_{i,1}), \ldots, \mathfrak{h}(b_{i,k}))$, and the PRF can be interpreted as a sequence of transformations applied to a random polynomial. More specifically, for every bit $x_i = 1$, the i-th step replaces the coefficients of a polynomial by its evaluations (up to the sign) at some random points $b_{i,1}, \ldots, b_{i,k}$.