Optimal Reductions of Some Decisional Problems to the Rank Problem

Jorge Luis Villar*

Universitat Politècnica de Catalunya, Spain
jvillar@ma4.upc.edu

Abstract. In the last years the use of large matrices and their algebraic properties proved to be useful to instantiate new cryptographic primitives like Lossy Trapdoor Functions and encryption schemes with improved security, like Key Dependent Message resilience. In these constructions the rank of a matrix is assumed to be hard to guess when the matrix is hidden by elementwise exponentiation. This problem, that we call here the Rank Problem, is known to be related to the Decisional Diffie-Hellman problem, but in the known reductions between both problems there appears a loss-factor in the advantage which grows linearly with the rank of the matrix.
In this paper, we give a new and better reduction between the Rank problem and the Decisional Diffie-Hellman problem, such that the reduction loss-factor depends logarithmically in the rank. This new reduction can be applied to a number of cryptographic constructions, improving their efficiency. The main idea in the reduction is to build from a DDH tuple a matrix which rank shifts from r to $2r$, and then apply a hybrid argument to deal with the general case. In particular this technique widens the range of possible values of the ranks that are tightly related to DDH.
On the other hand, the new reduction is optimal as we show the nonexistence of more efficient reductions in a wide class containing all the "natural" ones (i.e., black-box and algebraic). The result is twofold: there is no (natural) way to build a matrix which rank shifts from r to $2r + \alpha$ for $\alpha > 0$, and no hybrid argument can improve the logarithmic loss-factor obtained in the new reduction.
The techniques used in the paper extend naturally to other "algebraic" problems like the Decisional Linear or the Decisional 3-Party Diffie-Hellman problems, also obtaining reductions of logarithmic complexity.

Keywords: Rank Problem, Decisional Diffie-Hellman Problem, Black-Box Reductions, Algebraic Reductions, Decision Linear Problem

1 Introduction

Motivation. In the last years the use of large matrices and their algebraic properties proved to be useful to instantiate new cryptographic primitives like

* Partially supported by the Spanish research project MTM2009-07694, and the European Commission through the ICT programme under contract ICT-2007-216676 ECRYPT II.

Lossy Trapdoor Functions [7, 8, 12, 13] and encryption schemes with improved security, like Key Dependent Message [2]. In these constructions the rank of a matrix is assumed to be hard to guess when the matrix is hidden by elementwise exponentiation. This problem, that we call here the Rank Problem, is known to be related to the Decisional Diffie-Hellman (DDH) problem, but in the known reductions between both problems there appears a loss-factor in the adversary's advantage which grows linearly with the rank of the matrix. The Rank Problem first appeared in some papers under the names Matrix-DDH [2] and Matrix d-Linear [10].

In the cryptographic constructions mentioned above, some secret values (messages or keys) are encoded as group element vectors and then hidden by multiplying them by an invertible matrix. The secret value is recovered by inverting the operations: first multiplying by the inverse matrix and then inverting the encoding as group elements. This last step requires to encode a few bits (typically, a single bit) in each group element, forcing the length of the vector and the rank of the matrix to be comparable to the binary length of the secret. Security of these schemes is related to the indistinguishability of full-rank matrices from low-rank (e.g., rank 1) matrices: If the invertible matrix is replaced by a low rank one, the secret value is information-theoretically hidden. Therefore, the security of these schemes is related to the hardness of the Rank problem for matrices of large rank (e.g., 320 or 1024).

Reductions of the DDH problem to the Rank problem are based in the obvious relationship between them in the case of 2×2 matrices. Namely, from a DDH problem tuple (g, g^x, g^y, g^z) one can build a matrix $g^M = \begin{pmatrix} g & g^x \\ g^y & g^z \end{pmatrix}$, which is the elementwise exponentiation of the $\mathbb{Z}_q$ matrix $M = \begin{pmatrix} 1 & x \\ y & z \end{pmatrix}$. For a 0-instance of DDH (i.e., $z = xy$), $\det M = 0$, while for a 1-instance (i.e., $z \neq xy$), $\det M \neq 0$, and therefore, the rank of M shifts from 1 to 2 depending on the DDH instance. This technique can be applied to larger (even non-square) matrices by just padding the previous 2×2 block with some ones in the diagonal and zeroes elsewhere, just increasing the rank from 1 or 2 to $r + 1$ or $r + 2$, where r is the number of ones added to the diagonal.

Now, a general reduction of DDH to any instance of the rank problem (i.e., telling apart hidden matrices of ranks r_1 and r_2) is obtained by applying a hybrid argument, incurring into a loss-factor in the adversary's advantage which grows linearly in the rank difference $r_2 - r_1$.

This loss-factor has an extra impact on the efficiency of the cryptographic schemes based on matrices: For the same security level the size of the group has to be increased, and therefore the sizes of public keys, ciphertexts, etc. increase accordingly.

Until now it was an open problem to find a tighter reduction of DDH to the Rank problem. To face this kind of problems one can choose between building new tighter reductions or showing impossibility results. However, most of the known impossibility results are quite limited because they only claim the nonex-

istence of reductions of certain type (e.g., black-box, algebraic, etc.). But still these negative results have some value since they capture all possible 'natural' reductions between computational problems, at least in the generic case (e.g., without using specific properties of certain groups and their representation).

Main Results. In this paper, we give a new and better reduction between the Rank and the DDH problems, such that the reduction loss-factor grows logarithmically with the rank of the matrices. This new reduction can be applied to a number of cryptographic constructions improving their efficiency. The main idea in the reduction is to build a matrix from a DDH tuple which rank shifts from r to $2r$, and then apply a hybrid argument to deal with the general case.

On the other hand, the new reduction is optimal: We show the nonexistence of more efficient reductions in a wide class containing all the "natural" ones (i.e., black-box and algebraic). The result is twofold: There is no (natural) way to build a matrix which rank shifts from r to $2r + \alpha$ for $\alpha > 0$, and no hybrid argument can improve the logarithmic loss-factor obtained in the new reduction.

Basically, the new reduction achieves the following result.

(Informal) Theorem 1 *For any ℓ_1, ℓ_2, r_1, r_2 such that $1 \leq r_1 < r_2 \leq \min(\ell_1, \ell_2)$ there is a reduction of the DDH problem to the Rank problem for $\ell_1 \times \ell_2$ matrices of rank either r_1 or r_2, where the advantage of the problem solvers fulfil*

$$\mathbf{AdvRank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) \leq \left\lceil \log_2 \frac{r_2}{r_1} \right\rceil \mathbf{AdvDDH}(\mathcal{G}; t')$$

and their running times t and t' are essentially equal.

In particular, our reduction relates the DDH Problem to the hardness of telling apart $\ell \times \ell$ full rank matrices from rank 1 matrices with a loss-factor of only $\log_2(\ell)$, instead of the factor ℓ obtained in previous reductions. Moreover, the previous reductions are tight only for ranks r_1 and r_2 such that $r_2 = r_1 + 1$, while our results show that there exists a tight reduction for $r_1 < r_2 \leq 2r_1$.

At this point, it arises the natural question of whether a tight reduction exists for a wider range of the ranks r_1 and r_2. However, we show the optimality of the new reduction by the following negative result.

(Informal) Theorem 2 *For any ℓ_1, ℓ_2, r_1, r_2 such that $1 \leq r_1 < r_2 \leq \min(\ell_1, \ell_2)$ and any 'natural' reduction $\mathcal{R}$ of DDH to the Rank problem, the advantages of the Rank problem solver $\mathcal{A}$ and the DDH solver $\mathcal{R}([\mathcal{A}])$ fulfil*

$$\mathbf{AdvRank}_{\mathcal{R}[\mathcal{A}]}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) \geq \left\lceil \log_2 \frac{r_2}{r_1} \right\rceil \mathbf{AdvDDH}_{\mathcal{A}}(\mathcal{G}; t') - \varepsilon$$

where the running times t, t' are similar and ε is a negligible quantity.

Here, 'natural reduction' basically means a black-box reduction which transforms a DDH tuple into a hidden matrix by performing only (probabilistic)

algebraic manipulations, which are essentially linear combinations of the exponents with known integer coefficients, depending on the random coins of the reduction.

All generic reductions from computational problems based on cyclic groups fall into this category. Therefore, this result has to be interpreted as one cannot expect finding a tighter reduction for a large class of groups unless a new (non-black-box or not algebraic) technique is used. Nevertheless, falsifying this negative result would imply an improvement on the efficiency of the cryptosystems based on matrices, or even the discovery of a new reduction technique.

The techniques used in the paper extend naturally to other "algebraic" problems like the Decisional Linear (DLin) or the Decisional 3-Party Diffie-Hellman (D3DH) problems, also obtaining reductions with logarithmic complexity. Actually, these reductions recently appeared in [4] and [5].

(Informal) Theorem 3 *For any ℓ_1, ℓ_2, r_1, r_2 such that $2 \leq r_1 < r_2 \leq \min(\ell_1, \ell_2)$ there is a reduction of the DLin problem to the Rank problem for $\ell_1 \times \ell_2$ matrices of rank either r_1 or r_2, where the advantage of the problem solvers fulfil*

$$\mathbf{AdvRank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) \leq \left\lceil 1.71 \log_2 \frac{r_2}{r_1 - 1} \right\rceil \mathbf{AdvDLin}(\mathcal{G}; t')$$

and their running times t and t' are essentially equal.

(Informal) Theorem 4 *For any ℓ_1, ℓ_2, r_1, r_2 such that $2 \leq r_1 < r_2 \leq \min(\ell_1, \ell_2)$ there is a reduction of the D3DH problem to the Rank problem for $\ell_1 \times \ell_2$ matrices of rank either r_1 or r_2, where the advantage of the problem solvers fulfil*

$$\mathbf{AdvRank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) \leq \left\lceil 1.71 \log_2 \frac{r_2}{r_1 - 1} \right\rceil \mathbf{AdvD3DH}(\mathcal{G}; t')$$

and their running times t and t' are essentially equal.

Negative results similar to Theorem 2 are also given, but in these two cases the reductions are shown to be optimal up to a constant factor of 1.71.

Further Research. Some of the ideas and techniques used in the paper suggest that the problem of the optimality of certain type of reductions for a class of decisional assumptions can be studied under the Algebraic Geometric point of view. In particular, this could help to close the gap in the loss-factor between the reduction and the lower bound when reducing DLin or D3DH to Rank, and could made possible to obtain similar results for a broad class of computational problems. A second open problem is how the techniques and results adapt to the case of composite order groups, specially when the factorization of the order, or the order itself is unknown.

Roadmap. The paper starts with some notation and basic lemmas, in Section 2. Then the Rank Problem and the new reduction of DDH is presented in Section 3. The optimality of the reduction is studied in Section 4. In the last section of the paper, the previous results are extended to other "algebraic" decisional problems like DLin or D3DH.

2 Notation and Basic Lemmas

Let $\mathcal{G}$ be a group of prime order q, and let g be a random generator of $\mathcal{G}$. For convenience we will use additive notation for all groups. In particular, $0_{\mathcal{G}}$ denotes the neutral element in $\mathcal{G}$, whereas $1_{\mathcal{G}}$ denotes the generator g. Analogously, $x1_{\mathcal{G}}$, or simply $x_{\mathcal{G}}$, denotes the result of g^x, for any integer $x \in \mathbb{Z}_q$. The additive notation extends to vectors and matrices of elements in $\mathcal{G}$, in the natural way. That is, given a vector $\boldsymbol{x} = (x_1, \ldots, x_\ell) \in \mathbb{Z}_q^\ell$, we will write $\boldsymbol{x}_{\mathcal{G}} = ((x_1)_{\mathcal{G}}, \ldots, (x_\ell)_{\mathcal{G}})$, and the same for matrices. ${\mathbb{Z}_q}^{\ell_1 \times \ell_2}$ denotes the set of all $\ell_1 \times \ell_2$ matrices, and ${\mathbb{Z}_q}^{\ell_1 \times \ell_2; r}$ is used for the subset of those matrices with rank r. In the special case of invertible matrices we will write $\mathrm{GL}_\ell(\mathbb{Z}_q) = {\mathbb{Z}_q}^{\ell \times \ell; \ell}$. The sets of matrices with entries in $\mathcal{G}$, which we write ${\mathcal{G}}^{\ell_1 \times \ell_2}$, ${\mathcal{G}}^{\ell_1 \times \ell_2; r}$ and $\mathrm{GL}_\ell(\mathcal{G})$, are defined in the natural way by replacing every matrix M by $M_{\mathcal{G}}$. Notice that the sets are independent of the choice of the group generator $1_{\mathcal{G}}$.

An element $x_{\mathcal{G}} = x1_{\mathcal{G}} \in \mathcal{G}$ and an integer $a \in \mathbb{Z}_q$ can be operated together: $ax_{\mathcal{G}} = (ax \mod q)1_{\mathcal{G}} = (ax)_{\mathcal{G}} = xa_{\mathcal{G}}$. These operations extend to vectors and matrices in the natural way. Therefore, for any two matrices $A \in {\mathbb{Z}_q}^{\ell_1 \times \ell_2}$ and $B \in {\mathbb{Z}_q}^{\ell_2 \times \ell_3}$, we have $A_{\mathcal{G}} B = AB_{\mathcal{G}} = (AB)_{\mathcal{G}}$.

For convenience we will use the notation $A \oplus B$ for block matrix concatenation:

$$A \oplus B = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right)$$

In addition, I_ℓ and $0_{\ell_1 \times \ell_2}$ respectively denote the neutral element in $\mathrm{GL}_\ell(\mathbb{Z}_q)$ and the null matrix in ${\mathbb{Z}_q}^{\ell_1 \times \ell_2}$. The shorthand $0_\ell = 0_{\ell \times \ell}$ is also used. Given a matrix $A \in {\mathbb{Z}_q}^{\ell_1 \times \ell_2}$, the transpose of A is denoted by $A^\top$, and the vector subspace spanned by the columns of A is denoted by $\operatorname{Span} A \subseteq \mathbb{Z}_q^{\ell_1}$, which dimension equals $\operatorname{rank} A$.

Uniform sampling of a set S is written as $x \in_{\mathrm{R}} S$. In addition, sampling from a probability distribution D which support is included in S is denoted by $x \leftarrow D$, while $x \leftarrow \mathcal{A}(a)$ denotes that x is the result of running a (probabilistic) algorithm $\mathcal{A}$ on some input a.

As it is usual, a positive function $f : \mathbb{Z}^+ \to \mathbb{R}^+$ is called negligible if $f(\lambda)$ decreases faster than λ^{-c} for any positive constant c. We denote this by $f(\lambda) \in \mathbf{negl}(\lambda)$. Similarly, $f(\lambda) > \mathbf{negl}(\lambda)$ denotes that $f(\lambda)$ is non negligible in λ.

Lemma 1. *The following three natural group actions are transitive:*[1]

1. *the left-action of $GL_{\ell_1}(\mathbb{Z}_q)$ on ${\mathbb{Z}_q}^{\ell_1 \times \ell_2; \ell_2}$, for $\ell_1 \geq \ell_2$, defined by $A \mapsto UA$, where $U \in GL_{\ell_1}(\mathbb{Z}_q)$ and $A \in {\mathbb{Z}_q}^{\ell_1 \times \ell_2; \ell_2}$,*
2. *the right-action of $GL_{\ell_2}(\mathbb{Z}_q)$ on ${\mathbb{Z}_q}^{\ell_1 \times \ell_2; \ell_1}$, for $\ell_1 \leq \ell_2$, defined by $A \mapsto AV$, where $V \in GL_{\ell_2}(\mathbb{Z}_q)$ and $A \in {\mathbb{Z}_q}^{\ell_1 \times \ell_2; \ell_1}$,*

[1] The action of a group G on a set A is transitive if for any $a, b \in A$ there exists $g \in G$ such that $b = g \cdot a$. As a consequence, if $g \in_{\mathrm{R}} G$ then for any $a \in A$, $g \cdot a$ is uniform in A.

3. *the left-right-action of* $GL_{\ell_1}(\mathbb{Z}_q) \times GL_{\ell_2}(\mathbb{Z}_q)$ *on* ${\mathbb{Z}_q}^{\ell_1 \times \ell_2; r}$*, defined by* $A \mapsto UAV$*, where* $U \in GL_{\ell_1}(\mathbb{Z}_q)$*,* $V \in GL_{\ell_2}(\mathbb{Z}_q)$ *and* $A \in {\mathbb{Z}_q}^{\ell_1 \times \ell_2; r}$*.*

Lemma 2 (Rank Decomposition). *Given any matrix* $A \in {\mathbb{Z}_q}^{\ell_1 \times \ell_2; r}$*, there exist matrices* $L \in {\mathbb{Z}_q}^{\ell_1 \times r; r}$ *and* $R \in {\mathbb{Z}_q}^{r \times \ell_2; r}$ *such that* $A = LR$*.*

3 The Rank Problem and The New Reduction of DDH to Rank

We consider an assumption related to matrices, which is weaker than some well-known assumptions like the Decisional Diffie-Hellman, the Decisional Linear [1] and the Decisional 3-Party Diffie-Hellman [3, 6, 9] assumptions. Given an (additive) cyclic group $\mathcal{G}$ of prime order q of binary length λ, the $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ problem informally consists of distinguishing if a given matrix in ${\mathbb{Z}_q}^{\ell_1 \times \ell_2}$ has either rank r_1 or rank r_2, for given integers $r_1 < r_2$. The problem is formally defined through the following two experiments between a challenger and a distinguisher $\mathcal{A}$.

Experiment $\mathbf{ExpRank}^b_{\mathcal{A}}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ is defined as follows, for $b = 0, 1$.

1. If $b = 0$, the challenger chooses $M \in_{\mathrm{R}} {\mathbb{Z}_q}^{\ell_1 \times \ell_2; r_1}$ and sends $M_{\mathcal{G}}$ to $\mathcal{A}$.
   If $b = 1$, the challenger chooses $M \in_{\mathrm{R}} {\mathbb{Z}_q}^{\ell_1 \times \ell_2; r_2}$ and sends $M_{\mathcal{G}}$ to $\mathcal{A}$.
2. The distinguisher $\mathcal{A}$ outputs a bit $b' \in \{0, 1\}$.

Let Ω_b be the event that $\mathcal{A}$ outputs $b' = 1$ in $\mathbf{ExpRank}^b_{\mathcal{A}}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$. The advantage of $\mathcal{A}$ is defined as $\mathbf{AdvRank}_{\mathcal{A}}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2) = |\Pr[\Omega_0] - \Pr[\Omega_1]|$. We can then define

$$\mathbf{AdvRank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) = \max_{\mathcal{A}} \{\mathbf{AdvRank}_{\mathcal{A}}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)\}$$

where the maximum is taken over all $\mathcal{A}$ running within time t.

Definition 1. *The* $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ *assumption in a group* $\mathcal{G}$ *states that* $\mathbf{AdvRank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t)$ *is negligible in* $\lambda = \log |\mathcal{G}|$ *for any value of* t *that is polynomial in* λ*.*

The Rank assumption appeared in recent papers under the names Matrix-DDH [2] and Matrix d-Linear [10]. However, the reduction given in the next proposition substantially improves the reductions previously known. Namely, the loss factor in the new reduction grows no longer linearly but logarithmically in the rank.

Firstly, note that the $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ problem is random self-reducible, since by Lemma 1 given $M_0 \in {\mathbb{Z}_q}^{\ell_1 \times \ell_2; k}$, for random $L \in_{\mathrm{R}} \mathrm{GL}_{\ell_1}(\mathbb{Z}_q)$ and $R \in_{\mathrm{R}} \mathrm{GL}_{\ell_2}(\mathbb{Z}_q)$ the product LM_0R is uniformly distributed in ${\mathbb{Z}_q}^{\ell_1 \times \ell_2; k}$.

Lemma 3. *Any distinguisher for* $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, k - \delta, k)$*,* $\ell_1, \ell_2 \geq 2$*,* $k \geq 2$*,* $1 \leq \delta \leq \left\lfloor \frac{k}{2} \right\rfloor$ *can be converted into a distinguisher for the Decisional Diffie-Hellman* (DDH) *problem, with the same advantage and with essentially the same running time.*

Proof. Given a DDH instance $(1, x, y, z)_{\mathcal{G}}$, the DDH distinguisher builds the $\ell_1 \times \ell_2$ matrix

$$M_{\mathcal{G}} = \underbrace{\begin{pmatrix} 1 & x \\ y & z \end{pmatrix}_{\mathcal{G}} \oplus \cdots \oplus \begin{pmatrix} 1 & x \\ y & z \end{pmatrix}_{\mathcal{G}}}_{\delta \text{ times}} \oplus I_{k-2\delta_{\mathcal{G}}} \oplus 0_{(\ell_1-k)\times(\ell_2-k)_{\mathcal{G}}}$$

and submits the randomized matrix $LM_{\mathcal{G}}R$ to the $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, k-\delta, k)$ distinguisher, where $L \in_{\mathrm{R}} \mathrm{GL}_{\ell_1}(\mathbb{Z}_q)$ and $R \in_{\mathrm{R}} \mathrm{GL}_{\ell_2}(\mathbb{Z}_q)$. Notice that if $z = xy \bmod q$ then the resulting matrix is a random matrix in $\mathcal{G}^{\ell_1 \times \ell_2; k-\delta}$. Otherwise, it is a random matrix in $\mathcal{G}^{\ell_1 \times \ell_2; k}$. □

Theorem 1. *For any ℓ_1, ℓ_2, r_1, r_2 such that $1 \le r_1 < r_2 \le \min(\ell_1, \ell_2)$ we have,*

$$\mathbf{AdvRank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) \le \left\lceil \log_2 \frac{r_2}{r_1} \right\rceil \mathbf{AdvDDH}(\mathcal{G}; t')$$

where $t' = t + O(\ell_1 \ell_2 (\ell_1 + \ell_2))$, taking the cost of a scalar multiplication in $\mathcal{G}$ as one time unit.

Proof. We proceed by applying a hybrid argument. Let us consider the sequence of integers $\{n_i\}$ defined by $n_i = r_1 2^i$, and let k be the smallest index such that $n_k \ge r_2$, that is $k = \lceil \log_2 r_2 - \log_2 r_1 \rceil$. Then define a sequence of random matrices $\{M_{i\mathcal{G}}\}$, where $M_i \in_{\mathrm{R}} {\mathbb{Z}_q}^{\ell_1 \times \ell_2; n_i}$ for $i = 0, \ldots, k-1$, and $M_k \in_{\mathrm{R}} {\mathbb{Z}_q}^{\ell_1 \times \ell_2; r_2}$. For any distinguisher $\mathcal{A}_{\mathbf{Rank}}$ with running time upper bounded by t, let $p_i = \Pr[1 \leftarrow \mathcal{A}_{\mathbf{Rank}}(M_{i\mathcal{G}})]$. By Lemma 3,

$$|p_{i+1} - p_i| = \mathbf{AdvRank}_{\mathcal{A}_{\mathbf{Rank}}}(\mathcal{G}, \ell_1, \ell_2, n_i, n_{i+1}) \le \mathbf{AdvDDH}(\mathcal{G}; t')$$

for $i = 0, \ldots, k-2$, and

$$|p_k - p_{k-1}| = \mathbf{AdvRank}_{\mathcal{A}_{\mathbf{Rank}}}(\mathcal{G}, \ell_1, \ell_2, n_{k-1}, r_2) \le \mathbf{AdvDDH}(\mathcal{G}; t')$$

Therefore,

$$\begin{aligned} \mathbf{AdvRank}_{\mathcal{A}_{\mathbf{Rank}}}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2) = |p_k - p_0| \le |p_1 - p_0| + \ldots + |p_k - p_{k-1}| \le \\ \le k \cdot \mathbf{AdvDDH}(\mathcal{G}; t') \end{aligned}$$

which leads to the desired result. □

4 Optimality of the Reduction

In this section we show that there does not exist any reduction of DDH to the Rank problem that improves the result in Theorem 1, unless it falls out of the class of reductions that we call *black-box algebraic* reductions.

4.1 Black-Box Algebraic Reductions

Formally, a reduction $\mathcal{R}$ of a computational problem $\mathcal{P}_1$ to a problem $\mathcal{P}_2$ efficiently transforms any probabilistic polynomial time algorithm $\mathcal{A}_2$ solving $\mathcal{P}_2$ with a non-negligible advantage ε_2 into another probabilistic polynomial time algorithm $\mathcal{A}_1 = \mathcal{R}[\mathcal{A}_2]$ solving $\mathcal{P}_1$ with a non-negligible advantage ε_1. The reduction $\mathcal{R}$ is called *black-box* if $\mathcal{A}_1$ is just a probabilistic polynomial time algorithm with oracle access to $\mathcal{A}_2$.

In this paper we focus on the optimality of a reduction, measured in terms of the advantages of $\mathcal{A}_1$ and $\mathcal{A}_2$. However, to be meaningful we need to add another requirement to the reduction: The running times of $\mathcal{A}_1$ and $\mathcal{A}_2$ are similar. Otherwise, one can arbitrarily increase the advantage of $\mathcal{A}_1$ by repetition, thus making more than one oracle call to $\mathcal{A}_2$. We must add a qualifier and say that the reduction is then *time-preserving* black-box. However, for simplicity we will omit it and simply refer to black-box reductions.

Following [11], we say that $\mathcal{R}$ is *algebraic* with respect to a group $\mathcal{G}$ if it only performs group operations on the elements of $\mathcal{G}$ (i.e., group operation, inversion and comparison for equality), while there is no limitation in the operations performed on other data types. Although the notion of black-box algebraic reduction is theoretically very limited, it captures all the 'natural' reductions, since all known reductions between problems related to the discrete logarithm in cyclic groups fall into this category. See [11] for a deeper discussion on algebraic reductions and their relation with the generic group model.

In the definition of an algebraic algorithm $\mathcal{R}$ it is assumed that there exists an efficient extractor that, from the inputs of $\mathcal{R}$ (including the random tape) and the code of $\mathcal{R}$, it extracts a representation of every group element in $\mathcal{R}$'s output as a multiexponentiation of the base formed by the group elements in the input of $\mathcal{R}$. However, here we only require that for every value of the random tape of $\mathcal{R}$ there exists such representation, and it is independent of the group elements on the input of $\mathcal{R}$. More precisely, if $g_1, \ldots, g_m \in \mathcal{G}$ are the group elements in the input of $\mathcal{R}$ and $h_1, \ldots, h_n \in \mathcal{G}$ are the group elements in the output, then for any choice of the other inputs and the random tape, there exist coefficients $\alpha_{ij} \in \mathbb{Z}_q$ such that $h_i = \alpha_{i1} g_1 + \ldots + \alpha_{im} g_m$, for $i = 1, \ldots, n$. Notice that this is true as long as $\mathcal{R}$ performs only group operations on the group elements.

We insist in the possible existence of reductions using more intricate operations other than the group operations defined in $\mathcal{G}$. However, there is little hope to be able to control the rank of the manipulated matrices, except for the trivial fact that a random matrix has maximal rank with overwhelming probability.

4.2 Canonical Solvers

In this paper, we consider only reductions $\mathcal{R}$ of some decisional problem (like DDH) to the Rank problem (say $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$). Therefore, in a (time-preserving) black-box reduction, having oracle access to a solver $\mathcal{A}_2$ of Rank exactly means that $\mathcal{R}$ computes some matrix in $\mathcal{G}^{\ell_1 \times \ell_2}$, and uses it as input of

$\mathcal{A}_1$, then obtaining a bit $b' \in \{0,1\}$ as its output. Therefore, $\mathcal{R}$ is nothing more than a way to obtain a matrix from a DDH instance by an algebraic function.

As Rank problem is random self-reducible, one can consider the notion of a *canonical* solver $\overline{\mathcal{A}}$ for $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$. In a first stage, a canonical solver, on the input of a matrix $M_{\mathcal{G}} \in \mathcal{G}^{\ell_1 \times \ell_2}$, computes the randomized matrix $M'_{\mathcal{G}} = LM_{\mathcal{G}}R$ for randomly chosen $L \in \mathrm{GL}_{\ell_1}(\mathbb{Z}_q)$ and $R \in \mathrm{GL}_{\ell_2}(\mathbb{Z}_q)$, and then uses it as input of the second stage. Observe that $M_{\mathcal{G}}$ and $M'_{\mathcal{G}}$ have always the same rank, and they are nearly independent. Indeed $M_{\mathcal{G}}$ and $M'_{\mathcal{G}}$ conditioned to any specific value of the rank r are independent random variables, and $M'_{\mathcal{G}}$ is uniformly distributed in $\mathcal{G}^{\ell_1 \times \ell_2; r}$.

Moreover, for any solver $\mathcal{A}$ of $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ we build a canonical solver $\overline{\mathcal{A}}$ from $\mathcal{A}$ with the same advantage, by just inserting the initial randomization step. As a consequence, to obtain a negative result about the existence of black-box reductions of some problem to $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$, we only need to consider how the reduction works for canonical solvers of $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$.

Finally, it should be noticed that a canonical solver is completely characterized by a probability vector $\boldsymbol{p}_{\mathcal{A}} = (p_{\mathcal{A},i})_{i \in \mathbb{Z}^+}$, where $p_{\mathcal{A},i} = \Pr[1 \leftarrow \mathcal{A}(M_{\mathcal{G}}) : M_{\mathcal{G}} \in_{\mathrm{R}} \mathcal{G}^{\ell_1 \times \ell_2; i}]$. The advantage of a canonical solver is then $\mathbf{AdvRank}_{\mathcal{A}} = |p_{\mathcal{A},r_2} - p_{\mathcal{A},r_1}|$. Dealing with all canonical solvers of $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ means considering all possible probability vectors $\boldsymbol{p}_{\mathcal{A}}$ such that $|p_{\mathcal{A},r_2} - p_{\mathcal{A},r_1}|$ is non-negligible.

4.3 More Linear Algebra

Let us see the implications of restricting the reductions to be algebraic. Since here we reduce the decisional problem DDH to the $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ problem, the reduction $\mathcal{R}$ will receive as input either a 0-instance (i.e., $(1_{\mathcal{G}}, x_{\mathcal{G}}, y_{\mathcal{G}}, xy_{\mathcal{G}})$) or a 1-instance (i.e., $(1_{\mathcal{G}}, x_{\mathcal{G}}, y_{\mathcal{G}}, (xy+s)_{\mathcal{G}})$) of the decisional problem (where $x, y, s \in_{\mathrm{R}} \mathbb{Z}_q$). In spite of the instance received, $\mathcal{R}$ will compute a matrix $M_{\mathcal{G}} \in \mathcal{G}^{\ell_1 \times \ell_2}$ that depends 'algebraically' on the input group elements. Therefore, for any value of the random tape of $\mathcal{R}$ there exist matrices $B_1, B_2, B_3, B_4 \in {\mathbb{Z}_q}^{\ell_1 \times \ell_2}$ such that $M = B_1 + xB_2 + yB_3 + (xy+s)B_4$, where either $s = 0$ or $s \in_{\mathrm{R}} \mathbb{Z}_q$, depending on the type of instance received by $\mathcal{R}$.

Therefore, we need some properties of the sets of matrices that are linear combinations of some fixed matrices with coefficients that are multivariate polynomials. The following lemma informally states that matrices in a linear variety of ${\mathbb{Z}_q}^{\ell \times \ell}$ (of any dimension) are invertible with either zero or overwhelming probability.

Lemma 4. *Let $\mathcal{M}$ be a coset of a $\mathbb{Z}_q$-vector subspace of ${\mathbb{Z}_q}^{\ell \times \ell}$, that is, there exist matrices $A, B_1, \ldots, B_k \in {\mathbb{Z}_q}^{\ell \times \ell}$ for some integer k such that $\mathcal{M} = \{A + x_1B_1 + \ldots + x_kB_k \mid x_1, \ldots, x_k \in \mathbb{Z}_q\}$. If $GL_\ell(\mathbb{Z}_q) \cap \mathcal{M} \neq \emptyset$ then,*

$$\nu_{\mathcal{M}} = \frac{|GL_\ell(\mathbb{Z}_q) \cap \mathcal{M}|}{|\mathcal{M}|} > 1 - \frac{\ell}{q-1}$$

Proof. [2] Let us choose $A \in \mathrm{GL}_\ell(\mathbb{Z}_q) \cap \mathcal{M}$ and let $\{B_1, \ldots, B_k\}$ be a base of the vector space $\mathcal{M} - A$. In any line $\mathcal{L} \subset \mathcal{M}$ containing A there can be at most ℓ matrices $M \in \mathcal{L}$ such that $\operatorname{rank} M < \ell$ (i.e., $\det M = 0$). Indeed, for any line $\mathcal{L}$ there is a nonzero vector $\boldsymbol{x} = (x_1, \ldots, x_k) \in \mathbb{Z}_q^k$ such that $\mathcal{L} = \{A + \mu(x_1 B_1 + \ldots + x_k B_k) \mid \mu \in \mathbb{Z}_q\}$. Therefore the polynomial equation $\det(A + \mu(x_1 B_1 + \ldots + x_k B_k)) = 0$, which is equivalent to $Q_{\boldsymbol{x}}(\mu) = \det(I_\ell + \mu(x_1 B_1 A^{-1} + \ldots + x_k B_k A^{-1})) = 0$, has at most ℓ roots because $Q_{\boldsymbol{x}}(0) = 1$ and $\lambda^{-\ell} Q_{\boldsymbol{x}}(1/\lambda) = \det(\lambda I_\ell + x_1 B_1 A^{-1} + \ldots + x_k B_k A^{-1}) = 0$ if and only if λ is an eigenvalue of $x_1 B_1 A^{-1} + \ldots + x_k B_k A^{-1}$. Finally, since there are exactly $\left|\mathbb{P}\mathbb{Z}_q^{k-1}\right| = \frac{q^k - 1}{q - 1}$ different lines in $\mathcal{M}$ containing A,

$$\nu_F \geq 1 - \frac{\ell(q^k - 1)/(q - 1)}{q^k} > 1 - \frac{\ell}{q - 1}$$

as k is the dimension of the vector space $\mathcal{M} - A$, and then $|\mathcal{M}| = q^k$. □

This lemma can be easily generalized to parametrical subsets of linear varieties by replacing each variable x_j, $j = 1, \ldots, k$, by a multivariate polynomial $p_j(y_1, \ldots, y_n) \in \mathbb{Z}_q[y_1, \ldots, y_n]$ (or simply, $\mathcal{M}$ is now the range of a multivariate polynomial with matrix coefficients). Here we cannot ensure that the mapping between the parameter vector $\boldsymbol{y} = (y_1, \ldots, y_n)$ and the matrices in $\mathcal{M}$ is one-to-one. Therefore we will define $\nu_{\mathcal{M}}$ as the probability of obtaining a full-rank matrix when $\boldsymbol{y} \in_{\mathrm{R}} \mathbb{Z}_q^n$ is sampled with the uniform distribution.

Lemma 5. *Let $\mathcal{M}$ be a subset of ${\mathbb{Z}_q}^{\ell \times \ell}$ defined as $\mathcal{M} = \{p_1(\boldsymbol{y})B_1 + \ldots + p_k(\boldsymbol{y})B_k \mid \boldsymbol{y} \in \mathbb{Z}_q^n\}$, where $p_1(\boldsymbol{y}), \ldots, p_k(\boldsymbol{y}) \in \mathbb{Z}_q[\boldsymbol{y}]$ are multivariate polynomials of total degree at most d, and $B_1, \ldots, B_k \in {\mathbb{Z}_q}^{\ell \times \ell}$ for some integer k. If $GL_\ell(\mathbb{Z}_q) \cap \mathcal{M} \neq \emptyset$ then,*

$$\begin{aligned}\nu_{\mathcal{M}} = \Pr[M \in GL_\ell(\mathbb{Z}_q) \,:\, M = p_1(\boldsymbol{y})B_1 + \ldots + p_k(\boldsymbol{y})B_k,\, \boldsymbol{y} \in_{\mathrm{R}} \mathbb{Z}_q^n] \geq \\ \geq 1 - \frac{\ell d}{q - 1}\frac{q^n - 1}{q^n} > 1 - \frac{\ell d}{q - 1}\end{aligned}$$

Proof. The proof is similar, but now we choose $A = p_1(\boldsymbol{y}_0)B_1 + \ldots + p_k(\boldsymbol{y}_0)B_k \in \mathrm{GL}_\ell(\mathbb{Z}_q) \cap \mathcal{M}$ and define the new polynomials $q_i(\boldsymbol{z}) = p_i(\boldsymbol{y}_0 + \boldsymbol{z}) - p_i(\boldsymbol{y}_0)$ for $i = 1, \ldots k$. Now, $\mathcal{M} \setminus \{A\}$ is partitioned into subsets $\mathcal{L}^* = \{A + q_1(\mu\boldsymbol{z})B_1 + \ldots + q_k(\mu\boldsymbol{z})B_k) \mid \mu \in \mathbb{Z}_q^\times\}$, where $\boldsymbol{z} \in \mathbb{Z}_q^n \setminus \{\boldsymbol{0}\}$, each one containing at most ℓd singular matrices, since the polynomial $Q_{\boldsymbol{z}}(\mu) = \det(I_\ell + q_1(\mu\boldsymbol{z})B_1 A^{-1} + \ldots + q_k(\mu\boldsymbol{z})B_k A^{-1})$ is nonzero (as $Q_{\boldsymbol{z}}(0) = 1$), and it has degree at most $d\ell$. Finally, the claimed inequality follows from the fact that there are $(q^n - 1)/(q - 1)$ different subsets $\mathcal{L}^*$. □

The above lemmas refer only to invertible matrices but a similar result applies to (even rectangular) matrices with respect to a specific value of the rank.

[2] This lemma and the following one can alternatively be proved by using the Schwartz lemma [15] (also referred to as Schwartz-Zippel lemma).

Lemma 6. *Let $\mathcal{M}$ be a subset of $\mathbb{Z}_q^{\ell_1 \times \ell_2}$ defined as $\mathcal{M} = \{p_1(\boldsymbol{y})B_1 + \ldots + p_k(\boldsymbol{y})B_k \mid \boldsymbol{y} \in \mathbb{Z}_q^n\}$, where $p_1(\boldsymbol{y}), \ldots, p_k(\boldsymbol{y}) \in \mathbb{Z}_q[\boldsymbol{y}]$ are multivariate polynomials of total degree at most d, and $B_1, \ldots, B_k \in \mathbb{Z}_q^{\ell_1 \times \ell_2}$ for some integer k. If $r_m = \max_{m \in \mathcal{M}} \operatorname{rank} M$ then,*

$$\nu_{\mathcal{M}} = \Pr[\operatorname{rank} M = r_m \,:\, M = p_1(\boldsymbol{y})B_1 + \ldots + p_k(\boldsymbol{y})B_k,\, \boldsymbol{y} \in_{\mathrm{R}} \mathbb{Z}_q^n] > 1 - \frac{r_m d}{q-1}$$

Proof. We just apply the previous lemma to a projection of the set $\mathcal{M}$. Firstly choose $M_0 \in \mathcal{M}$ such that $\operatorname{rank} M_0 = r_m$ and find matrices $L \in \mathbb{Z}_q^{r_m \times \ell_1; r_m}$ and $R \in \mathbb{Z}_q^{\ell_2 \times r_m; r_m}$ such that $\operatorname{rank} LM_0R = r_m$, that is $LM_0R \in \mathrm{GL}_{r_m}(\mathbb{Z}_q)$. This matrices are really easy to build, since by Lemma 2 there exist $L_0 \in \mathbb{Z}_q^{\ell_1 \times r_m; r_m}$ and $R_0 \in \mathbb{Z}_q^{r_m \times \ell_2; r_m}$ such that $M_0 = L_0R_0$. Therefore, we take any L such that $LL_0 \in \mathrm{GL}_{r_m}(\mathbb{Z}_q)$. For instance, take L as a the all-zero matrix and put r_m ones in its main diagonal, in positions corresponding to r_m linearly independent rows of L_0. We similarly proceed with R_0 and R.

Now, the projected set $\mathcal{M}' = \{LMR \mid M \in \mathcal{M}\}$ fulfils the conditions of Lemma 5 and it contains at least one invertible matrix LM_0R. Thus,

$$\nu_{\mathcal{M}'} = \Pr[M' \in \mathrm{GL}_{r_m}(\mathbb{Z}_q) \,:\, M' = L(p_1(\boldsymbol{y})B_1 + \ldots + p_k(\boldsymbol{y})B_k)R,\, \boldsymbol{y} \in_{\mathrm{R}} \mathbb{Z}_q^n] > \\ > 1 - \frac{\ell r_m}{q-1}$$

Moreover, since $\operatorname{rank}(LMR) \leq \operatorname{rank} M \leq r_m$ for all $M \in \mathcal{M}$, then $\operatorname{rank}(LMR) = r_m$ implies $\operatorname{rank} M = r_m$, and

$$\Pr[\operatorname{rank} M = r_m \,:\, M = p_1(\boldsymbol{y})B_1 + \ldots + p_k(\boldsymbol{y})B_k,\, \boldsymbol{y} \in_{\mathrm{R}} \mathbb{Z}_q^n] \geq \nu_{\mathcal{M}'} > 1 - \frac{\ell r_m}{q-1}$$

□

This lemma basically says that in a set $\mathcal{M}$ defined and sampled as above the matrices have a specific rank (the maximal rank in the set) with overwhelming probability, and ranks below the maximal one occur only with negligible probability.

4.4 The Case of DDH

Now let us consider the specific case of the sets $\mathcal{M}_0$ and $\mathcal{M}_1$ generated by a black-box algebraic reduction $\mathcal{R}$ from a DDH 0-tuple or 1-tuple, respectively, for a fixed random tape of $\mathcal{R}$. More precisely, $\mathcal{M}_{\text{DDH-0}} = \{B_0 + xB_1 + yB_2 + xyB_3 \mid x, y \in \mathbb{Z}_q\}$, while $\mathcal{M}_{\text{DDH-1}} = \{B_0 + xB_1 + yB_2 + (xy+s)B_3 \mid x, y, s \in \mathbb{Z}_q\}$, for some matrices $B_0, B_1, B_2, B_3 \in \mathbb{Z}_q^{\ell_1 \times \ell_2}$ that could depend on the random tape. Let r_{m0} and r_{m1} be the maximal ranks respectively in $\mathcal{M}_{\text{DDH-0}}$ and $\mathcal{M}_{\text{DDH-1}}$. Since the former is a subset of the latter, $r_{m0} \leq r_{m1}$. In addition, it is clear that $\operatorname{rank} B_0 \leq r_{m0}$, but one can also prove that $\operatorname{rank} B_3 \leq r_{m0}$ and therefore $r_{m1} \leq 2r_{m0}$, as claimed in the following lemma.

Lemma 7. *Let r_{m0} and r_{m1} be the maximal ranks respectively in $\mathcal{M}_{DDH\text{-}0}$ and $\mathcal{M}_{DDH\text{-}1}$. Then $r_{m0} \leq r_{m1} \leq 2r_{m0}$.*

Proof. The left inequality is trivial, as mentioned above. To prove the right one we firstly use Lemma 6 to show that $\operatorname{rank} B_3 \leq r_{m0}$. Indeed, the subset $\mathcal{M}^*_{\text{DDH-0}} = \{B_0 + xB_1 + yB_2 + xyB_3 \mid x, y \in \mathbb{Z}_q^\times\}$ differs from $\mathcal{M}_{\text{DDH-0}}$ in that a negligible fraction of it has been removed. Therefore, the probability distributions on both sets (induced by uniformly sampling x and y) are statistically close. Since for all $x, y \in \mathbb{Z}_q^\times$, $\operatorname{rank}(B_0 + xB_1 + yB_2 + xyB_3) = \operatorname{rank}(\frac{1}{xy}B_0 + \frac{1}{y}B_1 + \frac{1}{x}B_2 + B_3)$, and the inversion map $x \mapsto 1/x$ is a bijection in $\mathbb{Z}_q^\times$, the probability distributions of the ranks in $\mathcal{M}^*_{\text{DDH-0}}$ and in $\overline{\mathcal{M}}^*_{\text{DDH-0}} = \{B_3 + xB_2 + yB_1 + xyB_0 \mid x, y \in \mathbb{Z}_q^\times\}$ are identical. Therefore, matrices in $\overline{\mathcal{M}}_{\text{DDH-0}} = \{B_3 + xB_2 + yB_1 + xyB_0 \mid x, y \in \mathbb{Z}_q\}$ have rank r_{m0} with overwhelming probability. Moreover, by Lemma 6, r_{m0} is precisely the maximal rank in $\overline{\mathcal{M}}_{\text{DDH-0}}$ and then, $\operatorname{rank} B_3 \leq r_{m0}$.[3]

Finally, observe that for any $M \in \mathcal{M}_{\text{DDH-1}}$, $M = B_0 + xB_1 + yB_2 + (xy + s)B_3 = (B_0 + xB_1 + yB_2 + xyB_3) + sB_3$ and $\operatorname{rank} M \leq \operatorname{rank}(B_0 + xB_1 + yB_2 + xyB_3) + \operatorname{rank}(sB_3) \leq 2r_{m0}$, because $B_0 + xB_1 + yB_2 + xyB_3 \in \mathcal{M}_{\text{DDH-0}}$. □

The previous discussion deals with a fixed arbitrary random tape of the reduction $\mathcal{R}$. However, the overall performance of $\mathcal{R}$ depends on the aggregation of the contributions of all possible values of the random tape. Technically, given a particular canonical solver $\mathcal{A}$ of $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$, described by its probability vector $\boldsymbol{p}_\mathcal{A}$ as defined in Section 4.2, the advantage of $\mathcal{R}[\mathcal{A}]$ can be computed as

$$\mathbf{AdvDDH}_{\mathcal{R}[\mathcal{A}]}(\mathcal{G}) = \left| \sum_{r=0}^{\min(\ell_1,\ell_2)} (\pi_{0,r} - \pi_{1,r}) p_{\mathcal{A}r} \right| = |(\boldsymbol{\pi}_0 - \boldsymbol{\pi}_1) \cdot \boldsymbol{p}_\mathcal{A}|$$

where

$$\pi_{0,r} = \Pr[\operatorname{rank} M = r \,:\, M \leftarrow \mathcal{R}(1_\mathcal{G}, x_\mathcal{G}, y_\mathcal{G}, xy_\mathcal{G}), x, y \in_\text{R} \mathbb{Z}_q]$$

and

$$\pi_{1,r} = \Pr[\operatorname{rank} M = r \,:\, M \leftarrow \mathcal{R}(1_\mathcal{G}, x_\mathcal{G}, y_\mathcal{G}, (xy + s)_\mathcal{G}), x, y, s \in_\text{R} \mathbb{Z}_q]$$

For convenience, we also introduce the cumulative probabilities $\mathit{\Pi}_{b,r} = \sum_{i=0}^{r} \pi_{b,i}$, $b \in \{0, 1\}$.

Since the reduction $\mathcal{R}$ must work for any successful solver $\mathcal{A}$, for every probability vector $\boldsymbol{p}_\mathcal{A}$ such that $|p_{\mathcal{A}r_1} - p_{\mathcal{A}r_2}| = \mathbf{AdvRank}_\mathcal{A}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ is non-negligible, the advantage $\mathbf{AdvDDH}_{\mathcal{R}[\mathcal{A}]}(\mathcal{G})$ must be also non-negligible. This implies the existence of $\alpha > \mathbf{negl}(\lambda)$ such that[4]

$$|\pi_{0,r} - \pi_{1,r}| \in \mathbf{negl}(\lambda) \qquad \forall r \notin \{r_1, r_2\}$$

[3] A very similar trick also shows that $\operatorname{rank} B_1$ and $\operatorname{rank} B_2$ ar at most r_{m0}. However, it is not clear how to extend this argument to arbitrary multivariate polynomials.

[4] To prove it, consider the fact that there cannot exist any probability vector $\boldsymbol{p}_\mathcal{A}$ orthogonal to $\boldsymbol{\pi}_0 - \boldsymbol{\pi}_1$ such that $|p_{\mathcal{A}r_1} - p_{\mathcal{A}r_2}| > \mathbf{negl}(\lambda)$.

$$|\pi_{0,r_1} - \pi_{1,r_1}| = \alpha$$
$$|\pi_{0,r_2} - \pi_{1,r_2}| = \alpha \pm \mathbf{negl}(\lambda) \tag{1}$$

Moreover,

$$\begin{aligned}\mathbf{AdvDDH}_{\mathcal{R}[\mathcal{A}]}(\mathcal{G}) &\leq |p_{\mathcal{A}r_1} - p_{\mathcal{A}r_2}|\,\alpha + \mathbf{negl}(\lambda) = \\ &= \alpha\mathbf{AdvRank}_{\mathcal{A}}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2) + \mathbf{negl}(\lambda)\end{aligned}$$

All that remains is to find an upper bound of the reduction loss-factor α.

By Lemma 6 we know that for every value of the random tape, $\Pr[\operatorname{rank} M < r_{mb} : M \leftarrow \mathcal{M}_{\text{DDH-b}}] \in \mathbf{negl}\,\lambda$ for $b \in \{0,1\}$, and by definition of r_{mb}, $\Pr[\operatorname{rank} M \leq r_{mb} : M \leftarrow \mathcal{M}_{\text{DDH-b}}] = 1$. Therefore, considering all values of the random tape of $\mathcal{R}$,[5]

$$\Pi_{b,i} = \Pr[r_{mb} \leq i] + \mathbf{negl}(\lambda) \qquad\qquad b \in \{0,1\} \tag{2}$$

where now r_{m0} and r_{m1} are random variables. By Lemma 7, $r_{m0} \leq r_{m1} \leq 2r_{m0}$, which implies[6] $\Pr[r_{m1} \leq i] \leq \Pr[r_{m0} \leq i] \leq \Pr[r_{m1} \leq 2i]$, for arbitrary i, and by (2),

$$\Pi_{1,i} - \mathbf{negl}(\lambda) \leq \Pi_{0,i} \leq \Pi_{1,2i} + \mathbf{negl}(\lambda) \tag{3}$$

Now, using left hand side of (3) for $i = r_1$ we get $\Pi_{1,r_1} \leq \Pi_{0,r_1} + \mathbf{negl}(\lambda)$, and combined with (1), we obtain $\pi_{0,r_1} = \pi_{1,r_1} + \alpha$ and $\pi_{1,r_2} \leq \pi_{0,r_2} + \alpha + \mathbf{negl}(\lambda)$. In addition, for any i such that $r_1 \leq i < r_2$,

$$\Pi_{0,i} = \Pi_{1,i} + \alpha \pm \mathbf{negl}(\lambda) \tag{4}$$

Let us assume now that $r_2 > 2^k r_1$ for some $k \geq 1$. Then, applying the right hand side of (3) and (4),

$$\Pi_{0,2^k r_1} = \Pi_{1,2^k r_1} + \alpha \pm \mathbf{negl}(\lambda) \geq \Pi_{0,2^{k-1} r_1} + \alpha - \mathbf{negl}(\lambda)$$

and by induction,

$$\Pi_{0,2^k r_1} \geq \Pi_{0,r_1} + k\alpha - \mathbf{negl}(\lambda) \geq (k+1)\alpha - \mathbf{negl}(\lambda)$$

where (4) is used again in the last step.

Finally, since the leftmost sum is upper bounded by 1,

$$\alpha \leq \frac{1 + \mathbf{negl}(\lambda)}{k+1}$$

for any $k < \log_2 r_2 - \log_2 r_1$. Therefore,

$$\alpha \leq \frac{1 + \mathbf{negl}(\lambda)}{\lceil \log_2 r_2 - \log_2 r_1 \rceil}$$

The above discussion proves the following theorem.

[5] If $r_{mb} \leq i$ then $\operatorname{rank} M \leq i$ with probability 1. Otherwise, $\operatorname{rank} M \leq i$ only with negligible probability.

[6] Observe that $r_{m1} \leq i \;\Rightarrow\; r_{m0} \leq i \;\Rightarrow\; r_{m1} \leq 2i$.

Theorem 2. *For any ℓ_1, ℓ_2, r_1, r_2 such that $1 \leq r_1 < r_2 \leq \min(\ell_1, \ell_2)$ and any time-preserving black-box algebraic reduction $\mathcal{R}$ of* $\mathbf{DDH}(\mathcal{G})$ *to the* $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2)$ *problem, any canonical Rank solver $\mathcal{A}$ and the corresponding DDH solver $\mathcal{R}([\mathcal{A}])$ fulfil*

$$\mathbf{AdvRank}_{\mathcal{R}[\mathcal{A}]}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) \geq \left\lceil \log_2 \frac{r_2}{r_1} \right\rceil \mathbf{AdvDDH}_{\mathcal{A}}(\mathcal{G}; t') - \mathbf{negl}(\lambda)$$

where the running times t, t' are similar.

□

5 Reductions of Other Decisional Problems

We consider now other well-known computational problems, namely the Decisional Linear (DLin) [1] and the Decisional 3-Party Diffie-Hellman (D3DH) [3, 6, 9] problems.

The techniques described above can be applied to these problems by defining a suitable basic matrix block M (of suitable size) where the problem instance is embedded, and use as many copies of it as possible. More precisely, we call *algebraic* to any decisional problem (such as DDH, DLin or D3DH) in which the problem instance is defined by a tuple of elements in a (cyclic) group which discrete logarithms fulfil or not a specific algebraic equation. The way the problem instance is embedded into the matrix M is by rewriting the algebraic equation as $\det M = 0$.

5.1 The Decisional Linear Problem

The Decisional Linear problem consists on distinguishing between the distributions $(x_{\mathcal{G}}, y_{\mathcal{G}}, z_{\mathcal{G}}, t_{\mathcal{G}}, (x^{-1}z + y^{-1}t)_{\mathcal{G}}) \in \mathcal{G}^5$ and $(x_{\mathcal{G}}, y_{\mathcal{G}}, z_{\mathcal{G}}, t_{\mathcal{G}}, u_{\mathcal{G}}) \in \mathcal{G}^5$, where $x, y, z, t, u \in_{\mathrm{R}} \mathbb{Z}_q$ are chosen independently and uniformly at random. More formally, we consider the following two experiments between a challenger and a distinguisher $\mathcal{A}$.

Experiment $\mathbf{ExpDLin}^b_{\mathcal{A}}(\mathcal{G})$ is defined as follows, for $b = 0, 1$.

1. The challenger chooses random $x, y, z, t, u \in_{\mathrm{R}} \mathbb{Z}_q$. If $b = 0$, the challenger sends the tuple $(1_{\mathcal{G}}, x_{\mathcal{G}}, y_{\mathcal{G}}, z_{\mathcal{G}}, t_{\mathcal{G}}, (x^{-1}z + y^{-1}t)_{\mathcal{G}}) \in \mathcal{G}^6$ to $\mathcal{A}$. Otherwise, it sends the tuple $(1_{\mathcal{G}}, x_{\mathcal{G}}, y_{\mathcal{G}}, z_{\mathcal{G}}, t_{\mathcal{G}}, u_{\mathcal{G}}) \in \mathcal{G}^6$.
2. The distinguisher $\mathcal{A}$ outputs a bit $b' \in \{0, 1\}$.

Let Ω_b be the event that $\mathcal{A}$ outputs $b' = 1$ in $\mathbf{ExpDLin}^b_{\mathcal{A}}(\mathcal{G})$. The advantage of $\mathcal{A}$ is $\mathbf{AdvDLin}_{\mathcal{A}}(\mathcal{G}) = |\Pr[\Omega_0] - \Pr[\Omega_1]|$. We can then define $\mathbf{AdvDLin}(\mathcal{G}; t) = \max_{\mathcal{A}} \{\mathbf{AdvDLin}_{\mathcal{A}}(\mathcal{G})\}$, where the maximum is taken over all $\mathcal{A}$ running within time t.

Definition 2 (DLin). *The* Decisional Linear assumption *in a group $\mathcal{G}$ states that* $\mathbf{AdvDLin}(\mathcal{G}; t)$ *is negligible in $\lambda = \log |\mathcal{G}|$ for any value of t that is polynomial in λ.*

Lemma 8. *Any distinguisher for* $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, k-\delta, k)$, $\ell_1, \ell_2 \geq 3$, $k \geq 3$, $1 \leq \delta \leq \left\lfloor \frac{k}{3} \right\rfloor$ *can be converted into a distinguisher for the Decisional Linear (DLin) problem, with the same advantage and running essentially within the same time.*

Proof. Given a DLin instance $(1, x, y, z, t, u)_{\mathcal{G}}$ the DLin distinguisher builds the $\ell_1 \times \ell_2$ matrix

$$M_{\mathcal{G}} = \underbrace{\begin{pmatrix} x & 0 & 1 \\ 0 & y & t \\ z & 1 & u \end{pmatrix}_{\mathcal{G}} \oplus \cdots \oplus \begin{pmatrix} x & 0 & 1 \\ 0 & y & t \\ z & 1 & u \end{pmatrix}_{\mathcal{G}}}_{\delta \text{ times}} \oplus I_{k-3\delta_{\mathcal{G}}} \oplus 0_{(m-k)\times(n-k)_{\mathcal{G}}}$$

and submits the randomized matrix $LM_{\mathcal{G}}R$ to the $\mathbf{Rank}(\mathcal{G}, \ell_1, \ell_2, k-\delta, k)$ distinguisher, where $L \in_{\mathrm{R}} \mathrm{GL}_{\ell_1}(\mathbb{Z}_q)$ and $R \in_{\mathrm{R}} \mathrm{GL}_{\ell_2}(\mathbb{Z}_q)$. Notice that if $u = x^{-1}z + y^{-1}t \mod q$ then the resulting matrix is a random matrix in $\mathcal{G}^{\ell_1 \times \ell_2; k-\delta}$. Otherwise, it is a random matrix in $\mathcal{G}^{\ell_1 \times \ell_2; k}$. □

Theorem 3. *For any* ℓ_1, ℓ_2, r_1, r_2 *such that* $2 \leq r_1 < r_2 \leq \min(\ell_1, \ell_2)$,

$$\mathbf{AdvRank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) \leq \left\lceil \frac{\log(3r_2) - \log(3r_1 - 2)}{\log 3 - \log 2} \right\rceil \mathbf{AdvDLin}(\mathcal{G}; t') \leq$$
$$\leq \left\lceil 1.71 \log_2 \frac{r_2}{r_1 - 1} \right\rceil \mathbf{AdvDLin}(\mathcal{G}; t')$$

Proof. We can apply a hybrid argument similar to the one used in Theorem 1. Let us consider the sequence of integers $\{n_i\}$ defined by the recurrence $n_0 = r_1$ and $n_{i+1} = \left\lfloor \frac{3n_i}{2} \right\rfloor$, and let k be the smallest index such that $n_k \geq r_2$. Then define a sequence of random matrices $\{M_{i\mathcal{G}}\}$, where $M_i \in_{\mathrm{R}} {\mathbb{Z}_q}^{\ell_1 \times \ell_2; n_i}$ for $i = 0, \ldots, k-1$, and $M_k \in_{\mathrm{R}} {\mathbb{Z}_q}^{\ell_1 \times \ell_2; r_2}$. For any distinguisher $\mathcal{A}_{\mathbf{Rank}}$ with running time upper bounded by t, let $p_i = \Pr[1 \leftarrow \mathcal{A}_{\mathbf{Rank}}(M_{i\mathcal{G}})]$. By Lemma 8,

$$|p_{i+1} - p_i| = \mathbf{AdvRank}_{\mathcal{A}_{\mathbf{Rank}}}(\mathcal{G}, \ell_1, \ell_2, n_i, n_{i+1}) \leq \mathbf{AdvDLin}(\mathcal{G}; t')$$

for $i = 0, \ldots, k-2$, and

$$|p_k - p_{k-1}| = \mathbf{AdvRank}_{\mathcal{A}_{\mathbf{Rank}}}(\mathcal{G}, \ell_1, \ell_2, n_{k-1}, r_2) \leq \mathbf{AdvDLin}(\mathcal{G}; t')$$

Therefore,

$$\mathbf{AdvRank}_{\mathcal{A}_{\mathbf{Rank}}}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2) = |p_k - p_0| \leq |p_1 - p_0| + \ldots + |p_k - p_{k-1}| \leq$$
$$\leq k \cdot \mathbf{AdvDLin}(\mathcal{G}; t')$$

On the other hand, as $\left\lfloor \frac{3x}{2} \right\rfloor \geq \frac{3x-1}{2}$ then $n_k \geq \left(\frac{3}{2}\right)^k \left(r_1 - \frac{2}{3}\right)$ which implies that $k \leq \frac{\log(3r_2) - \log(3r_1 - 2)}{\log 3 - \log 2}$. □

The optimality of the reduction presented above can be analyzed with the same tools described in Section 4, but adapting some parts of Subsection 4.4.

First of all, we can describe the 0-instances and the 1-instances for the DLin problem in a slightly different way. Namely, $\mathcal{M}_{\text{DLin-0}} = \{B_1 + xB_2 + yB_3 + x\alpha B_4 + y\beta B_5 + (\alpha+\beta)B_6 \mid x, y, \alpha, \beta \in \mathbb{Z}_q\}$, while $\mathcal{M}_{\text{DLin-1}} = \{B_1 + xB_2 + yB_3 + x\alpha B_4 + y\beta B_5 + (\alpha+\beta+s)B_6 \mid x, y, \alpha, \beta, s \in \mathbb{Z}_q\}$, for some matrices $B_1, B_2, B_3, B_4, B_5, B_6 \in \mathbb{Z}_q^{\ell_1 \times \ell_2}$ that could depend on the random tape of the reduction. By a similar trick one can manage to reprove Lemma 7 also for DLin and the rest of the analysis works equally well. The trick in this case is excluding the case $\alpha + \beta = 0$ (which affects to a negligible fraction of the matrices) and then using a more elaborate bijection which transforms $B_1 + xB_2 + yB_3 + x\alpha B_4 + y\beta B_5 + (\alpha+\beta)B_6$ into $\gamma B_1 + x\gamma B_2 + y\gamma B_3 + x\alpha\gamma B_4 + y(1-\alpha\gamma)B_5 + B_6$, where $\gamma = 1/(\alpha+\beta)$.

However, the logarithmic expression (which is identical to the one in Theorem 2) for the maximal loss-factor in the reduction is different from the loss-factor in the above reduction, leaving a gap that could mean that a better 'natural' reduction is still possible. Nevertheless, the authors think that a more detailed analysis of the maximal ranks r_{m0} and r_{m1} could be possible, which would improve the negative result obtained here.

5.2 The D3DH Problem

The Decisional 3-Party Diffie-Hellman (D3DH) problem [3, 6, 9] consists in telling apart the two distributions $(x_{\mathcal{G}}, y_{\mathcal{G}}, z_{\mathcal{G}}, (xyz)_{\mathcal{G}}) \in \mathcal{G}^4$ and $(x_{\mathcal{G}}, y_{\mathcal{G}}, z_{\mathcal{G}}, t_{\mathcal{G}}) \in \mathcal{G}^4$, where $x, y, z, t \in_{\text{R}} \mathbb{Z}_q$ are chosen independently at random. The problem is formally defined through the following two experiments between a challenger and a distinguisher $\mathcal{A}$.

Experiment $\mathbf{ExpD3DH}^b_{\mathcal{A}}(\mathcal{G})$ is defined as follows, for $b = 0, 1$.

1. The challenger chooses random $x, y, z, t \in_{\text{R}} \mathbb{Z}_q$. If $b = 0$, the challenger sends the tuple $(1_{\mathcal{G}}, x_{\mathcal{G}}, y_{\mathcal{G}}, z_{\mathcal{G}}, (xyz)_{\mathcal{G}}) \in \mathcal{G}^5$ to $\mathcal{A}$. Otherwise, it sends the tuple $(1_{\mathcal{G}}, x_{\mathcal{G}}, y_{\mathcal{G}}, z_{\mathcal{G}}, t_{\mathcal{G}}) \in \mathcal{G}^5$.
2. The distinguisher $\mathcal{A}$ outputs a bit $b' \in \{0, 1\}$.

Let Ω_b be the event that $\mathcal{A}$ outputs $b' = 1$ in $\mathbf{ExpD3DH}^b_{\mathcal{A}}(\mathcal{G})$. The advantage of $\mathcal{A}$ is $\mathbf{AdvD3DH}_{\mathcal{A}}(\mathcal{G}) = |\Pr[\Omega_0] - \Pr[\Omega_1]|$ and we define $\mathbf{AdvD3DH}(\mathcal{G}, t) = \max_{\mathcal{A}} \{\mathbf{AdvD3DH}_{\mathcal{A}}(\mathcal{G})\}$, where the maximum is taken over all $\mathcal{A}$ running within time t.

Definition 3. *The* Decisional 3-Party Diffie-Hellman assumption *in a group $\mathcal{G}$ states that* $\mathbf{AdvD3DH}(\mathcal{G}, t)$ *is negligible in $\lambda = \log|\mathcal{G}|$ for any value of t that is polynomial in λ.*

Similar to the Decisional Linear problem, it turns out that the D3DH problem is easier than the Rank problem.

Theorem 4. *For any ℓ_1, ℓ_2, r_1, r_2 such that $2 \le r_1 < r - 2 \le \min(\ell_1, \ell_2)$,*

$$\mathbf{AdvRank}(\mathcal{G}, \ell_1, \ell_2, r_1, r_2; t) \le \left\lceil \frac{\log(3r_2) - \log(3r_1 - 2)}{\log 3 - \log 2} \right\rceil \mathbf{AdvD3DH}(\mathcal{G}; t') \le$$

$$\leq \left\lceil 1.71 \log_2 \frac{r_2}{r_1 - 1} \right\rceil \mathbf{AdvD3DH}(\mathcal{G}; t')$$

Proof. The proof only differs from the proof of Proposition 3 in the 3×3 blocks built from a problem instance, in the proof of Lemma 3. Indeed, given the D3DH instance $(1, x, y, z, t)_{\mathcal{G}}$ the matrix

$$\begin{pmatrix} x & -1 & 0 \\ 0 & y & 1 \\ t & 0 & z \end{pmatrix}$$

has rank 2 or 3 depending on whether $t = xyz \mod q$. □

The analysis of the optimality of this reduction is comparable to the case of the Decisional Linear problem. Here the sets of matrices are $\mathcal{M}_{\text{D3DH-0}} = \{B_1 + xB_2 + yB_3 + zB_4 + xyzB_5 \mid x, y, z \in \mathbb{Z}_q\}$ and $\mathcal{M}_{\text{D3DH-1}} = \{B_1 + xB_2 + yB_3 + zB_4 + (xyz + s)B_5 \mid x, y, z, s \in \mathbb{Z}_q\}$, for some matrices $B_1, B_2, B_3, B_4, B_5 \in \mathbb{Z}_q^{\ell_1 \times \ell_2}$ that could depend on the random tape of the reduction. The same gap between the constructive and negative results is obtained.

5.3 Further Generalizations

The ideas presented before, both the constructive and the negative results for reductions of some decisional problems to the Rank problem seems to be easily applicable to a wide class of decisional problems. On the one hand, the construction of a reduction to the Rank problem only needs a way to encode the difference the 0-instance and the 1-instance of the problem as the determinant of a square matrix M built up from the group elements in the instances. Typically a 0-instance corresponds to $\det M = 0$. Following this approach, it is straightforward to obtain efficient reductions for instance for the family of Decisional r-Linear Problems, with arbitrary r.

On the other hand, the negative results about the existence of efficient reductions also rely on algebraic considerations, mainly related to the sets $\mathcal{M}$ which can be seen as special affine algebraic varieties. It is an open problem to obtain a description of a wide class of algebraic decisional problems for which a general negative result can be derived.

In this paper, only prime order groups are considered. However, it would be interesting to investigate whether the techniques presented here can be applied to composite order groups, where the matrices involved in the analysis are defined over rings, and this can introduce some extra difficulties to deal with notions like the rank and the random self-reducibility.

6 Acknowledgements

The authors are grateful to Dennis Hofheinz, David Galindo, Javier Herranz, Eike Kiltz, Gottfried Herold and Alexander May for insightful discussions and comments.

References

1. Dan Boneh, Xavier Boyen, and Hovav Shacham. Short group signatures. In Matthew K. Franklin, editor, *CRYPTO*, volume 3152 of *Lecture Notes in Computer Science*, pages 41–55. Springer, 2004.
2. Dan Boneh, Shai Halevi, Michael Hamburg, and Rafail Ostrovsky. Circular-secure encryption from decision diffie-hellman. In David Wagner, editor, *CRYPTO*, volume 5157 of *Lecture Notes in Computer Science*, pages 108–125. Springer, 2008.
3. Dan Boneh, Amit Sahai, and Brent Waters. Fully collusion resistant traitor tracing with short ciphertexts and private keys. In Serge Vaudenay, editor, *EUROCRYPT*, volume 4004 of *Lecture Notes in Computer Science*, pages 573–592. Springer, 2006.
4. David Galindo, Javier Herranz, and Jorge Luis Villar. Identity-based encryption with master key-dependent message security and applications. *IACR Cryptology ePrint Archive*, 2012:142, 2012.
5. David Galindo, Javier Herranz, and Jorge Luis Villar. Identity-based encryption with master key-dependent message security and leakage resilience. In Sara Foresti and Moti Yung, editors, *Computer Security - ESORICS 2012*, volume 7459 of *Lecture Notes in Computer Science*, pages 627–642. Springer, 2012.
6. Matthew Green and Susan Hohenberger. Practical adaptive oblivious transfer from simple assumptions. In Yuval Ishai, editor, *TCC*, volume 6597 of *Lecture Notes in Computer Science*, pages 347–363. Springer, 2011.
7. Dennis Hofheinz. All-but-many lossy trapdoor functions. Cryptology ePrint Archive, Report 2011/230, 2011. http://eprint.iacr.org/.
8. Dennis Hofheinz. All-but-many lossy trapdoor functions. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT*, volume 7237 of *Lecture Notes in Computer Science*, pages 209–227. Springer, 2012.
9. Fabien Laguillaumie, Pascal Paillier, and Damien Vergnaud. Universally convertible directed signatures. In Roy [14], pages 682–701.
10. Moni Naor and Gil Segev. Public-key cryptosystems resilient to key leakage. In Shai Halevi, editor, *CRYPTO*, volume 5677 of *Lecture Notes in Computer Science*, pages 18–35. Springer, 2009.
11. Pascal Paillier and Damien Vergnaud. Discrete-log-based signatures may not be equivalent to discrete log. In Roy [14], pages 1–20.
12. Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. *IACR Cryptology ePrint Archive*, 2007:279, 2007.
13. Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. In Cynthia Dwork, editor, *STOC*, pages 187–196. ACM, 2008.
14. Bimal K. Roy, editor. *Advances in Cryptology - ASIACRYPT 2005, 11th International Conference on the Theory and Application of Cryptology and Information Security, Chennai, India, December 4-8, 2005, Proceedings*, volume 3788 of *Lecture Notes in Computer Science*. Springer, 2005.
15. Jacob T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, 1980.