Limits on the Efficiency of (Ring) LWE based Non-Interactive Key Exchange

Siyao Guo[1⋆], Pritish Kamath[2⋆⋆], Alon Rosen[3⋆⋆⋆], and Katerina Sotiraki[4†]

[1] NYU Shanghai, Shanghai, 200122, China siyao.guo@nyu.edu
[2] TTIC, Chicago, IL 60637, USA pritish@ttic.edu
[3] IDC Herzliya, Herzliya, 4610101, Israel alon.rosen@idc.ac.il
[4] MIT, Cambridge, MA 02139, USA katesot@mit.edu

Abstract. LWE based key-exchange protocols lie at the heart of post-quantum public-key cryptography. However, all existing protocols either lack the *non-interactive* nature of Diffie-Hellman key-exchange or *polynomial* LWE-modulus, resulting in unwanted efficiency overhead.
We study the possibility of designing non-interactive LWE-based protocols with *polynomial* LWE-modulus. To this end,

- We identify and formalize simple non-interactive and polynomial LWE-modulus variants of existing protocols, where Alice and Bob *simultaneously* exchange one or more (ring) LWE samples with polynomial LWE-modulus and then run individual key reconciliation functions to obtain the shared key.
- We point out central barriers and show that such non-interactive key-exchange protocols are impossible if:
  1) the reconciliation functions first compute the inner product of the received LWE sample with their private LWE secret. This impossibility is information theoretic.
  2) One of the reconciliation functions does not depend on the error of the transmitted LWE sample. This impossibility assumes hardness of LWE.
- We give further evidence that progress in either direction, of giving an LWE-based NIKE protocol or proving impossibility of one will lead to progress on some other well-studied questions in cryptography.

Overall, our results show possibilities and challenges in designing simple (ring) LWE-based non-interactive key exchange protocols.

⋆ Supported by Shanghai Eastern Young Scholar Program.
⋆⋆ Work done while at MIT, supported by NSF awards CCF-1733808, IIS-1741137 and MIT-IBM Watson AI Lab and Research Collaboration Agreement No. W1771646.
⋆⋆⋆ Supported by ISF grant No. 1399/17 and via Project PROMETHEUS (Grant 780701).
† Research supported in part by NSF/BSF grant #1350619, an MIT-IBM grant, a DARPA Young Faculty Award, MIT Lincoln Laboratories and Analog Devices.

1 Introduction

In 1976, Diffie and Hellman [DH76] proposed an extremely elegant key-exchange protocol, in which two parties, Alice and Bob, exchange respective group elements g^a, g^b *simultaneously*, where g is a generator of a publicly chosen group $\mathcal{G}$ and $a, b \in [|\mathcal{G}|]$ are uniformly chosen secret elements. Alice and Bob then locally perform a single group exponentiation in order to derive the shared key, g^{ab}. This simple idea lies at the foundation of public key cryptography, and has been widely used in practice throughout the years.
Two decades later, Shor [Sho94] showed that efficient quantum algorithms could, in principle, break the Diffie-Hellman key-exchange protocol, as well as other widely used assumptions (e.g. Factoring). Thus, with the development of quantum computers on the horizon, the importance of designing post-quantum secure key-exchange protocols, that can replace current standards, has been recognized. As part of this effort, the National Institute of Standards and Technology (NIST) decided to look into post-quantum cryptography standardization and is hosting a post-quantum cryptography call of proposals [NIS]. One of the major primitives that they seek is a key-encapsulation mechanism.

1.1 (Ring) LWE based Key Exchange Protocols

A significant portion of the algorithms qualified to the second round of the NIST call for proposals [SAB+17], [NAB+17], [LLJ+17], [PAA+17], [GMZB+17] is based on the (ring) learning with errors (LWE) assumption [Reg05,LPR10]. A remarkable feature of this assumption (and consequently of the proposals) is that its *average-case* hardness is based on the *worst-case* hardness of lattice problems, which themselves are conjectured to be secure against efficient quantum algorithms.
Those proposals use two routes to achieve key-exchange, one is through public-key encryption and the other is through reconciliation. However, all of them *lack the non-interactive nature* of the key-exchange protocol of Diffie-Hellman, as explained below.

Key-exchange through public-key encryption. In the first case, Alice samples a secret & public-key pair and sends her public-key to Bob. Then, Bob picks a desired shared key and sends it to Alice, encrypted under her public-key. Finally, Alice decrypts Bob's message to recover the shared key. While conceptually simple, this approach lacks some of the advantages of the Diffie-Hellman protocol. Firstly, Bob has complete control over the shared key. Secondly, the protocol is inherently interactive – the parties need at least two rounds of interaction.

Key-exchange through reconciliation. The reconciliation approach was introduced by Ding et al. [DXL12] and Peikert [Pei14] and was implemented and improved in later works [ADPS16,BCNS14]. The most basic version of such reconciliation-based protocols has a simple description[5]

[5] For simplicity, we only describe the LWE-based variant; the ring version is obtained by replacing $\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{e}_1, \boldsymbol{e}_2$ with ring elements from some chosen polynomial ring and using the corresponding polynomial multiplication.

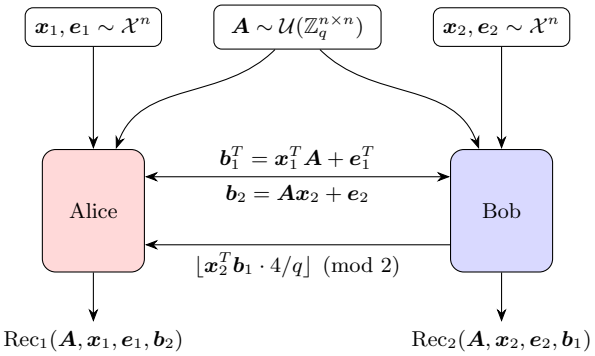

Fig. 1. Alice and Bob *simultaneously* exchange LWE samples using the same public matrix $\boldsymbol{A}$. *After receiving* $\boldsymbol{b}_2$, Bob sends the second most significant bit of $\boldsymbol{x}_2^T \boldsymbol{b}_1$ to Alice. Both players then apply their respective *key reconciliation functions* on the variables they have to produce a shared key.

(See Figure 1) : Let $\boldsymbol{A}$ be a random public $n \times n$ matrix over $\mathbb{Z}_q$ where q is *polynomial* in n and let $\mathcal{X}$ be a *noise distribution*, then the parties act as follows: Alice randomly picks $\boldsymbol{x}_1, \boldsymbol{e}_1$ from $\mathcal{X}^n$ and sends $\boldsymbol{b}_1 = \boldsymbol{x}_1^T \boldsymbol{A} + \boldsymbol{e}_1$ to Bob, while Bob *simultaneously* picks random $\boldsymbol{x}_2, \boldsymbol{e}_2$ from $\mathcal{X}^n$ and sends $\boldsymbol{b}_2 = \boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{e}_2$ to Alice. *After receiving* $\boldsymbol{b}_1$, Bob sends to Alice the second most significant bit of $\boldsymbol{x}_2^T \boldsymbol{b}_1$, i.e., $\lfloor 4/q \cdot \boldsymbol{x}_2^T \boldsymbol{b}_1 \rfloor \pmod 2$. To agree on a common key, Alice and Bob first compute the inner product of their secret and incoming message and obtain $\boldsymbol{x}_1^T \boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{x}_1^T \boldsymbol{e}_2$ and $\boldsymbol{x}_1^T \boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{e}_1^T \boldsymbol{x}_2$ respectively. The small magnitude of Alice and Bob's secret and noise already allows them to achieve approximate agreement: the most significant bit of $\boldsymbol{x}_1^T \boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{x}_1^T \boldsymbol{e}_2$ and $\boldsymbol{x}_1^T \boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{e}_1^T \boldsymbol{x}_2$ is often the same. To achieve exact agreement, they run a simple *key reconciliation* procedure, where Bob sends the second most significant bit as an additional hint.

1.2 (Ring) LWE based Non-Interactive Key Exchange?

As discussed above, Diffie-Hellman key exchange allows parties to send their messages simultaneously or communicate in a non-interactive way (e.g. by publishing them on Alice's and Bob's public websites). In the contrast, current proposed LWE-based key exchange protocols require additional interactions. Even though the additional interaction is only a single bit (as is the case in Figure 1), one extra round of a practical key exchange protocol may result in significant delays when used at a large scale (such as that of the internet). This motivates the main question that we study in this paper:

Can we have practical (ring) LWE-based non-interactive key exchange protocols? Or are such protocols inherently interactive?

A remark on LWE-modulus. Throughout the paper, we focus on polynomial LWE-modulus. We observe that if superpolynomial LWE-modulus is to be considered, LWE-based key exchange in Figure 1 *can be made non-interactive*. That's because the most significant bits of $\boldsymbol{x}_1^T \boldsymbol{b}_2$ and $\boldsymbol{x}_2^T \boldsymbol{b}_1$ agree with probability $1 - \Theta(nB^2/q)$, for a noise distribution $\mathcal{X}$ whose support is included in $[-B, B]$. If the modulus to noise rate is large (i.e. superpolynomial in the security parameter), then the probability of disagreement of their most significant bits is negligible, and hence the above non-interactive protocol is sufficient. However, in the case of a polynomially bounded q, the disagreement probability is non-negligible. Given the extremely demanding efficiency constraints on practical implementations[6], it would be highly desirable to have variants of such LWE-based key-exchange protocol in which the disagreement probability is negligible even in the case that q is as small as a polynomial in the security parameter. Additionally, requiring a large modulus to noise rate affects the hardness of the corresponding LWE assumption, since the worst-to-average case reductions translate this rate to the gap in the promise lattice problems [Pei09]. Namely, LWE with large modulus-to-noise rate is a stronger assumption (i.e. more susceptible to polynomial-time attacks) than LWE with a smaller modulus-to-noise rate.

1.3 Our Results

In this paper, we explore the possibility of attaining (ring) LWE-based non-interactive key exchange (NIKE) (with modulus polynomial in the security parameter).

Our focus. We focus on the setting where Alice and Bob only send one or a few (ring) LWE samples to each other; similarly to the protocol in Figure 1, but without the last message sent from Bob to Alice.
The main motivation for studying this setting is that perhaps it is the simplest setting which captures natural non-interactive variants of current LWE based key exchange protocols. Therefore, impossibility results will give a theoretical justification for current LWE based key exchange protocols. On the other hand, possibility results will yield Diffe-Hellman like non-interactive protocols.

Moreover, NIKE in this setting is simply characterized by two efficiently computable *key reconciliation functions* $\mathrm{Rec}_1, \mathrm{Rec}_2$, such that

- The outputs of Alice and Bob agree with each other with overwhelming probability, that is, $\mathrm{Rec}_1(\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \boldsymbol{b}_2) = \mathrm{Rec}_2(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1)$ holds with overwhelming probability (recall that $\boldsymbol{b}_1 := \boldsymbol{A}^T \boldsymbol{x}_1 + \boldsymbol{e}_1$ and $\boldsymbol{b}_2 := \boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{e}_2$).

[6] a typical size of q is $\approx 2^{13}$ and there are proposals that even use $q = 257$ [LLJ+17].

– The output of the protocol is pseudo-random even when conditioned on the transcript, that is, it is hard to predict $\mathrm{Rec}_1(\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \boldsymbol{b}_2)$ given $\boldsymbol{A}, \boldsymbol{b}_1, \boldsymbol{b}_2$.

Natural choices of reconciliation functions. Observe that in Figure 1, Alice and Bob achieve approximate agreement by computing $\boldsymbol{x}_1^T \boldsymbol{b}_2$ and $\boldsymbol{x}_2^T \boldsymbol{b}_1$, respectively. These values are noisy versions of $\boldsymbol{x}_1^T \boldsymbol{A} \boldsymbol{x}_2$ and their most significant bit agrees with probability $1 - \Theta(nB^2/q)$ when the support of $\mathcal{X}$ is in $[-B, B]$. Based on this observation, one may consider the following three families of reconciliation functions (in increasing order of generality).

1. Rec_1 and Rec_2 are arbitrary efficient functions (not necessarily the most significant bit) on $\boldsymbol{x}_1^T \boldsymbol{b}_2$ and $\boldsymbol{x}_2^T \boldsymbol{b}_1$ respectively.
2. Rec_1 and Rec_2 are arbitrary efficient functions on $\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{b}_2$ and $\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{b}_1$ respectively.
3. Rec_1 and Rec_2 are arbitrary efficient functions on $\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \boldsymbol{b}_2$ and $\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1$ respectively.

Note that the third family captures all possible reconciliation functions. Our main results rule out the first and second families of reconciliation functions even when multiple LWE samples are exchanged, and point out central efficiency barriers for the third family.

First result (Section 3). One natural idea to remove the interaction would be to somehow "amplify" the agreement probability by sending more LWE samples and generating more independent samples from the joint distribution $(\boldsymbol{X}, \boldsymbol{Y})$ where $\boldsymbol{X} := \boldsymbol{x}_1^T \boldsymbol{b}_2$ and $\boldsymbol{Y} := \boldsymbol{x}_2^T \boldsymbol{b}_1$, then apply Rec_1 and Rec_2 on independent samples from $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively.
In Theorem 1, we show that for any m, balanced $\mathrm{Rec}_1, \mathrm{Rec}_2$ (see Definition 1) and non-trivial noise distribution, $\mathrm{Rec}_1(\boldsymbol{X}^m) = \mathrm{Rec}_2(\boldsymbol{Y}^m)$ holds with probability at most $1 - \Omega(1/q^2)$. This implies that such reconciliation functions cannot exist (this impossibility is information theoretic and holds even for computationally inefficient reconciliation functions). Our results naturally extend to the case of ring LWE.

Second result (Section 4). Even though the above result captures known constructions, it does not rule out a slightly more general case where the reconciliation functions depend on $\boldsymbol{A}$. Indeed, given $\boldsymbol{X}' := (\boldsymbol{A}, \boldsymbol{X})$ and $\boldsymbol{Y}' := (\boldsymbol{A}, \boldsymbol{Y})$, Alice and Bob can agree on an *insecure* random bit with probability 1 by evaluating a balanced function of $\boldsymbol{A}$ (while ignoring $\boldsymbol{X}$ and $\boldsymbol{Y}$). Of course, such protocols are not suitable for key agreement, since the common random bit is not pseudo-random conditioned on $\boldsymbol{A}$.
In Theorem 3, we show that the reconciliation fuctions Rec_1 and Rec_2 have to depend on the LWE noises $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ respectively. For instance, the above theorem excludes a more general case than family 2 where the reconciliation functions are of the form $\mathrm{Rec}_1(\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \boldsymbol{b}_2) = h_1(\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{b}_2)$ and $\mathrm{Rec}_2(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1) = h_2(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1)$. In particular, it rules out the case where the joint distribution is $(\boldsymbol{X}', \boldsymbol{Y}')$. However, in contrast to Theorem 1 which holds unconditionally, Theorem 3 assumes the hardness of the LWE problem. Our results extend to the case of ring LWE and to a polynomial number of samples.

Third result (Section 5). The above two results rule out the most natural choices of key reconciliation functions based on variants of inner product, unconditionally or under the LWE assumption. In Section 5.1, we show that the existence of efficient Rec_1 and Rec_2, which depend on all of their inputs, cannot be ruled out (at least as long as the existence of iO is a possibility). In particular, in Theorem 4, we show that there exists an instantiation of the NIKE protocol in our framework that is based on indistinguishability obfuscation (iO) and puncturable PRFs [BZ17].
However, we identify a crucial restriction on the complexity of reconciliation functions. In Theorem 5, we show that the reconciliation functions themselves actually have to contain cryptographic hardness, in the sense that they *directly* yield weak pseudorandom functions. Therefore, the reconciliation functions have to be at least as complex as weak pseudorandom functions and hence suffer from the complexity limitations and attacks on weak pseudorandom functions. Moreover, this connection shows that any NIKE protocol based on hardness of LWE with polynomial modulus, gives rise to new constructions of weak pseudorandom functions based on the hardness of LWE with polynomial modulus. Such constructions have been an open problem almost since the introduction of the LWE assumption, and thus we view Theorem 5 as an indication that finding appropriate reconciliation functions requires new techniques.

1.4 Discussion and Open Problems

When parties exchange only LWE samples, we rule out the most natural choices of key reconciliation functions. Additionally, we point out that non-interactive key reconciliation functions, unlike interactive ones, have to be as complex as weak pseudorandom functions. Overall, our results show possibilities and challenges in designing simple (ring) LWE-based non-interactive key exchange protocols.
An interesting open direction is to understand what happens when the messages contain extra information, apart from the LWE samples. To this end, one would have to come up with a natural and simple form of messages (based on LWE) and explore the possibility of basing non-interactive key exchange on it. For instance, a natural idea is to consider LWE samples together with some *leakage* about the secrets. We remark that Theorem 5 continues to hold even if the leakage function is pseudorandom.

2 Preliminaries

We now provide some useful notation and definitions. We denote a sample drawn from $\mathcal{D}$ by $x \sim \mathcal{D}$ and a sample of the uniform distribution over S by $x \sim S$.

Definition 1. *A function $f : S \to \{0,1\}$ is called balanced respect to distribution $\mathcal{D}$ if $\mathbb{E}_{x\sim\mathcal{D}}[f(x)] = 1/2$.*

Definition 2. *A distribution $\mathcal{X}$ over $\mathbb{Z}_q$ is B-bounded if its support is included in $[-B, B]$.*

We formally define the class of all non-interactive key exchange protocols that could exist. We use $\mathsf{negl}(\lambda)$ to denote any function $g : \mathbb{R} \to \mathbb{R}$ that satisfies $g(\lambda) \leq O(n^{-c})$ for all constants $c \in \mathbb{N}$.

Definition 3. *For a security parameter $\lambda > 0$, a non-interactive key-exchange protocol consists of two $\mathsf{poly}(\lambda)$-time algorithms b_1 and b_2 and two $\mathsf{poly}(\lambda)$-time computable boolean functions Rec_1 and Rec_2 that satisfy the conditions below (where $(\boldsymbol{r}, \boldsymbol{r}_1, \boldsymbol{r}_2)$ is a random source where $\boldsymbol{r}$ is a source of shared randomness and $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ are private sources of randomness of the two parties)*

1. $\Pr_{\boldsymbol{r},\boldsymbol{r}_1,\boldsymbol{r}_2}[\mathrm{Rec}_1(\boldsymbol{r}, \boldsymbol{r}_1, b_2(\boldsymbol{r}, \boldsymbol{r}_2)) = \mathrm{Rec}_2(\boldsymbol{r}, \boldsymbol{r}_2, b_1(\boldsymbol{r}, \boldsymbol{r}_1))] \geq 1 - \mathsf{negl}(\lambda)\,,$
2. *For any probabilistic $\mathsf{poly}(\lambda)$-time algorithm $\mathcal{A}$,*

$$\Pr_{\boldsymbol{r},\boldsymbol{r}_1,\boldsymbol{r}_2}[\mathcal{A}(\boldsymbol{r}, b_1(\boldsymbol{r}, \boldsymbol{r}_1), b_2(\boldsymbol{r}, \boldsymbol{r}_2)) = \mathrm{Rec}_1(\boldsymbol{r}, \boldsymbol{r}_1, b_2(\boldsymbol{r}, \boldsymbol{r}_2))] \leq \frac{1}{2} + \mathsf{negl}(\lambda)\,.$$

Finally, we describe the Learning-with-Errors (LWE) assumption.

Definition 4. *[Reg05] The LWE assumption for integers n, m, q and noise distribution $\mathcal{X}$ over $\mathbb{Z}_q$ states that,*

$$(\boldsymbol{A}, \boldsymbol{b} := \boldsymbol{x}^T\boldsymbol{A} + \boldsymbol{e}) \approx_c (\boldsymbol{A}, \boldsymbol{u}),$$

where $\boldsymbol{A} \sim \mathbb{Z}_q^{n\times m}$, $\boldsymbol{u} \sim \mathbb{Z}_q^m$, $\boldsymbol{x} \sim \mathcal{X}^n$ and $\boldsymbol{e} \sim \mathcal{X}^m$.

3 (Information Theoretic) Impossibility of Amplification with Multiple Samples

Before stating the main Theorem of this section, we provide some definitions and notation.

Definition 5. *A distribution $\mathcal{X}$ over any group G (e.g. $G = \mathbb{Z}_q$) is* symmetric *if $\Pr_{X\sim\mathcal{X}}[X = z] = \Pr_{X\sim\mathcal{X}}[X = -z]$ for any $z \in G$.*

Given a distribution $\mathcal{X}$ over $\mathbb{Z}_q$, let $(\mathcal{X}^n)^*$ be the distribution of $\boldsymbol{w} = (w^{(1)}, w^{(2)}, \ldots, w^{(n)})$ drawn from $\mathcal{X}^n$ conditioned on the event that $\boldsymbol{w}$ is not a zero-divisor, that is $\gcd(w^{(1)}, w^{(2)}, \ldots, w^{(n)}, q) = 1$.

Theorem 1. *Let $n, q \geq 1$ be integers and $\mathcal{X}$ be a symmetric distribution over $\mathbb{Z}_q$ such that for any $a \in \mathbb{Z}_q \setminus \{0\}$, it holds that $\Pr_{X\sim\mathcal{X}}[aX = 0] \leq 9/10$ and $\Pr_{X\sim\mathcal{X}}[aX = q/2] \leq 9/10$. Let $\mu_{\mathcal{X}}(X, Y)$ be the joint distribution of*

$$X = \boldsymbol{x}_1^T\boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{x}_1^T\boldsymbol{e}_2 \text{ and } Y = \boldsymbol{x}_1^T\boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{e}_1^T\boldsymbol{x}_2,$$

where $\boldsymbol{A} \sim \mathcal{U}(\mathbb{Z}_q^{n\times n})$, $\boldsymbol{e}_1, \boldsymbol{e}_2 \sim \mathcal{X}^n$ and $\boldsymbol{x}_1, \boldsymbol{x}_2 \sim (\mathcal{X}^n)^*$. Then, for any $m \geq 1$, and any balanced functions $\mathrm{Rec}_1, \mathrm{Rec}_2 : \mathbb{Z}_q^m \to \{0, 1\}$ respect to the marginal distributions of $\mu_{\mathcal{X}}^{\otimes m}$, it holds that

$$\Pr_{(\boldsymbol{X},\boldsymbol{Y})\sim\mu_{\mathcal{X}}^{\otimes m}}[\mathrm{Rec}_1(\boldsymbol{X}) = \mathrm{Rec}_2(\boldsymbol{Y})] \leq 1 - \Omega(1/q^2).$$

Our theorem also holds for the ring case with the same parameters (See Theorem 6 in Appendix A). This theorem shows that no matter how many independent samples are drawn and no matter what procedures are applied on those samples, Alice and Bob can agree with each other on a random bit with probability at most $1 - \Omega(1/q^2)$. Note that Alice and Bob have to marginally produce a uniform bit as captured in the condition that Rec_1 and Rec_2 are balanced.

Our theorem applies to the most commonly used noise distributions. For instance, the discrete Gaussian distribution $\mathcal{D}_\sigma$ with standard deviation $\sigma > 10$ satisfies the conditions of Theorem 1. First, the discrete Gaussian is a symmetric distribution. Second, if $x \sim \mathcal{D}_\sigma$, then from monotonicity of $\mathcal{D}_\sigma$, for any $a \in \mathbb{Z}_q \setminus \{0\}$, $\Pr[ax = q/2] \leq \Pr[ax = 0]$. Therefore, it is enough to show that for any $a \in \mathbb{Z}_q \setminus \{0\}$, $\Pr[ax = 0] \leq 9/10$ which is straightforward to verify[7].

Additionally, the condition of Theorem 1 that for any $a \in \mathbb{Z}_q \setminus \{0\}$, $\Pr[aX = 0] \leq 9/10$ and $\Pr[aX = q/2] \leq 9/10$ is quite mild. For instance, if $q > 2$ is prime, then this condition simplifies to the assumption that the support of $\mathcal{X}$ is not equal to $\{0\}$. Also, for general q if the support of $\mathcal{X}$ is 1/10-far from a proper subgroup or a coset of a proper subgroup of $\mathbb{Z}_q$, then this assumption is satisfied.

Notice that $\mu_{\mathcal{X}}(X, Y)$ as defined in Theorem 1 does not correspond to the joint distribution described in the introduction, since $\boldsymbol{x}_1, \boldsymbol{x}_2$ are sampled from $(\mathcal{X}^n)^*$. This is without loss of generality because if $\boldsymbol{w} \sim \mathcal{X}^n$, then the probability that $\gcd(w^{(1)}, w^{(2)}, \ldots, w^{(n)}, q) \neq 1$ is smaller than the probability that $w^{(1)}, w^{(2)}, \ldots, w^{(n)}$ all belong to a proper subgroup of $\mathbb{Z}_q$, which is less than $(9/10)^n$. So, the distribution of $(\boldsymbol{X}, \boldsymbol{Y})$ is at most $O(m/(9/10)^n)$ far from the distribution of m samples drawn as described in the introduction. Even though this is a very small change in the protocol, it will simplify our proof a lot, since in this case the value $\boldsymbol{x}_1^T \boldsymbol{A} \boldsymbol{x}_2$ is a uniform element in $\mathbb{Z}_q$ [8].

Our Theorem 1 shows that in this regime, it is *information theoretically* impossible to agree on a common bit with probability $1 - o(1/q^2)$. In fact, the problem of generating common randomness by observing independent samples from two correlated distributions (or a joint distribution) is known as "Non-interactive Agreement Distillation" in the area of information theory (See Section 3.1) and the notion of maximal correlation exactly captures this problem (upto a polynomial factor in the error). Even though we could prove our theorem in a self-contained manner, we feel this connection provides more insight. Therefore, in the next section we present some basic facts about maximal correlation and then present a proof through this notion. In Appendix A, we also present a self-contained proof of Theorem 1 using Fourier analysis and extend this to the ring LWE case (Theorem 6).

[7] Note that by symmetry and monotonicity of $\mathcal{D}_\sigma$, $\Pr[ax = 0] \leq \Pr[a(|x| - 1) = 0] + \Pr[x = 0]$. Combining with the fact that $\Pr[ax = 0] + \Pr[a(|x| - 1) = 0] \leq 1$ for $a \neq 0$, and $\Pr[x = 0] \leq 1/(1 + 2e^{-1/\sigma^2})$, we conclude that $\Pr[ax = 0] \leq (1 + \Pr[x = 0])/2 \leq 9/10$ for $\sigma > 10$.

[8] If $\boldsymbol{w} = (w^{(1)}, w^{(2)}, \ldots, w^{(n)})$ such that $\gcd(w^{(1)}, w^{(2)}, \ldots, w^{(n)}, q) = 1$ and $\mathbf{u}$ is uniform in $\mathbb{Z}_q^n$, then $\boldsymbol{w}^T \mathbf{u}$ is also uniform in $\mathbb{Z}_q$.

3.1 Maximal Correlation and Non-interactive Agreement Distillation

The *Non-interactive Agreement Distillation* problem, parameterized by a joint distribution $\mu(x, y)$ is defined as follows: Two players, Alice and Bob, observe sequences $(X_1, \ldots, X_m)$ and $(Y_1, \ldots, Y_m)$ respectively where $\{(X_i, Y_i)\}_{i=1}^m$ are drawn i.i.d. from $\mu(x, y)$. Both players look at their share of randomness, apply a function and output a bit. Their goal is to maximize the probability that their output bits agree, while ensuring that they are marginally uniform.
Hirschfeld [Hir35] and Gebelein [Geb41] introduced the notion of *maximal correlation*, which was later studied by Rényi [Rén59]. It turns out that maximal correlation (almost tightly) captures the maximum agreement probability that the players can get.

Definition 6 (Maximal Correlation). *For a joint distribution μ over $G_A \times G_B$, its maximal correlation $\rho(\mu)$ is defined as follows,*

$$\sup_{f,g} \left\{ \mathop{\mathbb{E}}_{(x,y)\sim\mu} [f(x) \cdot g(y)] \;\middle|\; \begin{array}{l} f : G_A \to \mathbb{R}, \quad \mathbb{E}_{\mu_{G_A}}[f] = \mathbb{E}_{\mu_{G_B}}[g] = 0 \\ g : G_B \to \mathbb{R}, \, \mathrm{Var}_{\mu_{G_A}}[f] = \mathrm{Var}_{\mu_{G_B}}[g] = 1 \end{array} \right\},$$

where μ_{G_A} and μ_{G_B} are the marginal distributions of μ.

In order to analytically capture maximal correlation, let us define, for any joint distribution μ over $G_A \times G_B$, the $|G_A| \times |G_B|$ matrix M_μ given by

$$M_\mu(x, y) = \frac{\mu(x, y)}{\sqrt{\mu_A(x)\mu_B(y)}} \, .$$

where μ_A and μ_B are the marginal distributions of μ.

Fact 2 *The maximal correlation $\rho(\mu)$ is equal to the second largest singular value of M_μ, denoted as $\sigma_2(M_\mu)$.*[9]

In the seminal work of [Wit75], it was shown that maximal correlation actually captures (up to a square root factor), the best agreement probability that the players can get even with an infinite number of samples!

Lemma 1. *Suppose $\rho(\mu) = 1 - \varepsilon$, then for any $m \geq 1$, $f : G_A^m \to \{0, 1\}$ and $g : G_B^m \to \{0, 1\}$ with $\mathbb{E}_{\mu_{\boldsymbol{X}}^{\otimes m}}[f] = \mathbb{E}_{\mu_{\boldsymbol{Y}}^{\otimes m}}[g] = 1/2$, it holds that*

$$\Pr_{(\boldsymbol{X},\boldsymbol{Y})\sim\mu^{\otimes m}} [f(\boldsymbol{X}) = g(\boldsymbol{Y})] \leq 1 - \varepsilon/2. \tag{1}$$

Moreover, there exist m, f, g such that $\mathbb{E}_{\mu_{\boldsymbol{X}}^{\otimes m}}[f] = \mathbb{E}_{\mu_{\boldsymbol{Y}}^{\otimes m}}[g] = 1/2$ and

$$\Pr_{(\boldsymbol{X},\boldsymbol{Y})\sim\mu^{\otimes m}} [f(\boldsymbol{X}) = g(\boldsymbol{Y})] \;\geq\; 1 - \frac{\arccos(\rho(\mu))}{\pi} \;\geq\; 1 - \sqrt{2\varepsilon}. \tag{2}$$

[9] The top singular value being 1.

3.2 Bounding Maximal Correlation

Given Lemma 1, it suffices to upper bound the maximal correlation of $\mu_{\mathcal{X}}(X, Y)$. We exploit the special form of our distribution, namely that X is distributed uniformly in $\mathbb{Z}_q$ and $X - Y$ is distributed as some "noise distribution" ξ. For such distributions, the maximal correlation is much easier to analyze. In this section, we prove the following lemma.

Lemma 2. *Let $n, q \geq 1$ be integers. For a distribution $\mathcal{X}$ over $\mathbb{Z}_q$ and the joint distribution $\mu_{\mathcal{X}}$ that satisfies the conditions of Theorem 1, it holds that*

$$\rho(\mu_{\mathcal{X}}) \leq 1 - \Omega(1/q^2).$$

Theorem 1 follows immediately by combining Lemma 1 and Lemma 2. To prove Lemma 2, we consider a more general class of joint distributions called Cayley Distributions and characterize their maximal correlation.

Definition 7 (Cayley Distributions). *A joint distribution μ over $\mathbb{Z}_q^k \times \mathbb{Z}_q^k$ is said to be a* Cayley distribution *if there exists a "noise distribution" $\xi : \mathbb{Z}_q^k \to \mathbb{R}_{\geq 0}$, such that,*
(i) $\xi(\boldsymbol{z}) = \xi(-\boldsymbol{z})$ for all $\boldsymbol{z} \in \mathbb{Z}_q^k$ and
(ii) $\mu(\boldsymbol{x}, \boldsymbol{y}) = \frac{\xi(\boldsymbol{x}-\boldsymbol{y})}{q^k}$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}_q^k$.[10]

A Cayley distribution can be viewed as sampling $\boldsymbol{x}$ uniformly at random in $\mathbb{Z}_q^k$, sampling $\boldsymbol{z} \sim \xi$ and setting $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}$. Note that a Cayley distribution μ is symmetric and has uniform marginals on $\mathbb{Z}_q^k$, so its maximal correlation is given by the second largest eigenvalue of M_μ (by Theorem 2 and the fact that for symmetric matrices, singular values are same as eigenvalues). Interestingly, the eigenvectors of M_μ can be completely characterized in a way that does not depend on the noise distribution ξ. This makes it easy to get a handle on the eigenvalues, which leads to the following lemma.

Lemma 3 (Maximal Correlation of Cayley Distributions [Lov75]). *For $\boldsymbol{a} \in \mathbb{Z}_q^k$, define the* character *$\chi_{\boldsymbol{a}} : \mathbb{Z}_q^k \to \mathbb{C}$ as $\chi_{\boldsymbol{a}}(\boldsymbol{x}) = e^{-2\pi i \cdot \langle \boldsymbol{a}, \boldsymbol{x} \rangle / q}$. Let μ be any Cayley distribution over $\mathbb{Z}_q^k \times \mathbb{Z}_q^k$, with associated noise function ξ. Then*

$$\rho(\mu) = \max_{\boldsymbol{a} \in Z_q^k \setminus \{\boldsymbol{0}^k\}} \mathop{\mathbb{E}}_{\boldsymbol{e} \sim \xi} [\chi_{\boldsymbol{a}}(\boldsymbol{e})].$$

We point out that Definition 7 and Lemma 3 generalize to all finite abelian groups G. However for concreteness, we only focus on our special case of $G = \mathbb{Z}_q^k$. While this lemma is standard, we include a proof for completeness.

Proof. We interpret $\chi_{\boldsymbol{a}}$ as a vector in $\mathbb{C}^{q^k}$ indexed by elements in $\mathbb{Z}_q^k$. It is straightfoward to verify that $\chi_{\boldsymbol{a}} \in \mathbb{C}^{q^k}$ is an eigenvector of M_μ with corresponding eigenvalue $\mathbb{E}_{\boldsymbol{e} \sim \xi}[\chi_{\boldsymbol{a}}(\boldsymbol{e})]$. Note that since μ is a Caley

[10] Observe that since ξ is a probability distribution over $\mathbb{Z}_q^k$, it follows that μ is also a probability distribution.

distribution, $M_\mu(\boldsymbol{x},\boldsymbol{y}) = q^k \cdot \mu(\boldsymbol{x},\boldsymbol{y})$. Fix any $\boldsymbol{a} \in \mathbb{Z}_q^k$. For any $\boldsymbol{x} \in \mathbb{Z}_q^k$, it holds that

$$\begin{aligned}
(M_\mu \chi_{\boldsymbol{a}})(\boldsymbol{x}) &= \sum_{\boldsymbol{y}\in\mathbb{Z}_q^k} M_\mu(\boldsymbol{x},\boldsymbol{y}) \cdot \chi_{\boldsymbol{a}}(\boldsymbol{y}) = \sum_{\boldsymbol{y}\in\mathbb{Z}_q^k} (q^k \cdot \mu(\boldsymbol{x},\boldsymbol{y})) \cdot \chi_{\boldsymbol{a}}(\boldsymbol{y}) \\
&= \sum_{\boldsymbol{y}\in\mathbb{Z}_q^k} \xi(\boldsymbol{y}-\boldsymbol{x}) \cdot \chi_{\boldsymbol{a}}(\boldsymbol{y}) = \sum_{\boldsymbol{e}\in\mathbb{Z}_q^k} \xi(\boldsymbol{e}) \cdot \chi_{\boldsymbol{a}}(\boldsymbol{x}+\boldsymbol{e}) \\
&= \left(\sum_{\boldsymbol{e}\in\mathbb{Z}_q^k} \xi(\boldsymbol{e}) \cdot \chi_{\boldsymbol{a}}(\boldsymbol{e}) \right) \cdot \chi_{\boldsymbol{a}}(\boldsymbol{x}) \\
&= \mathop{\mathbb{E}}_{\boldsymbol{e}\sim\xi}[\chi_{\boldsymbol{a}}(\boldsymbol{e})] \cdot \chi_{\boldsymbol{a}}(\boldsymbol{x}) \,.
\end{aligned}$$

Note that the largest eigenvalue is $\mathbb{E}_{\boldsymbol{e}\sim\xi}[\chi_{\boldsymbol{a}}(\boldsymbol{e})] = 1$ given by $\boldsymbol{a} = \mathbf{0}^k$ because for any $\boldsymbol{e} \in \mathbb{Z}_q^k$, $\chi_{\mathbf{0}^k}(\boldsymbol{e}) = 1$ and $|\chi_{\boldsymbol{a}}(\boldsymbol{e})| \le 1$ if $\boldsymbol{a} \neq \mathbf{0}^k$. Hence, $\rho(\mu)$, which is the second largest eigenvalue of M_μ, is $\max_{\boldsymbol{a}\in\mathbb{Z}_q^k\setminus\{\mathbf{0}^k\}} \mathbb{E}_{\boldsymbol{e}\sim\xi}[\chi_{\boldsymbol{a}}(\boldsymbol{e})]$.

Proof (Proof of Lemma 2). Note that $\mu_{\mathcal{X}}$ is a Cayley distribution over $\mathbb{Z}_q \times \mathbb{Z}_q$ with associated noise distribution $\xi(z) = \Pr[\boldsymbol{x}_1^T\boldsymbol{e}_2 - \boldsymbol{e}_1^T\boldsymbol{x}_2 = z]$, where $\boldsymbol{e}_1, \boldsymbol{e}_2$ are drawn from $\mathcal{X}^n$ and $\boldsymbol{x}_1, \boldsymbol{x}_2$ are drawn from $(\mathcal{X}^n)^*$. First, $\xi(z) = \xi(-z)$ for any $z \in \mathbb{Z}_q$, since $\boldsymbol{x}_1^T\boldsymbol{e}_2$ and $\boldsymbol{e}_1^T\boldsymbol{x}_2$ are drawn from the same distribution, and so $\boldsymbol{x}_1^T\boldsymbol{e}_2 - \boldsymbol{e}_1^T\boldsymbol{x}_2$ is distributed identically to $\boldsymbol{e}_1^T\boldsymbol{x}_2 - \boldsymbol{x}_1^T\boldsymbol{e}_2$. Second, because $\boldsymbol{x}_1^T\boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{x}_1^T\boldsymbol{e}_2$ is distributed uniformly over $\mathbb{Z}_q$ and is independent from $\boldsymbol{x}_1^T\boldsymbol{e}_2 - \boldsymbol{e}_1^T\boldsymbol{x}_2$, we have that $\mu_{\mathcal{X}}(X,Y) = \Pr[\boldsymbol{x}_1^T\boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{x}_1^T\boldsymbol{e}_2 = X \textbf{ and } \boldsymbol{x}_1^T\boldsymbol{e}_2 - \boldsymbol{e}_1^T\boldsymbol{x}_2 = X - Y] = \frac{\xi(X-Y)}{q}$.
By Lemma 3, $\rho(\mu_{\mathcal{X}}) = \max_{a\in\mathbb{Z}_q\setminus\{0\}} \mathbb{E}_{e\sim\xi}[\chi_a(e)]$. Fix an arbitrary $a \in \mathbb{Z}_q \setminus \{0\}$, we need to show that $|\mathbb{E}_{e\sim\xi}[\chi_a(e)]| \le 1 - \Omega(1/q^2)$. This is implied by Claim 1 and Claim 2 below.

Claim 1 $|\mathbb{E}_{e\sim\xi}[\chi_a(e)]| \le \max_{\boldsymbol{c}\in\mathbb{Z}_q^n\setminus\{\mathbf{0}^n\}} |\mathbb{E}_{\boldsymbol{e}\sim\mathcal{X}^n}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]|$.

Proof. Note that

$$\begin{aligned}
|\mathop{\mathbb{E}}_{e\sim\xi}[\chi_a(e)]| &= \Big|\mathop{\mathbb{E}}_{\boldsymbol{x}_1,\boldsymbol{x}_2\sim(\mathcal{X}^n)^*}\Big[\mathop{\mathbb{E}}_{\boldsymbol{e}_1,\boldsymbol{e}_2\sim\mathcal{X}^n}[\chi_a(\boldsymbol{x}_1^T\boldsymbol{e}_2 - \boldsymbol{e}_1^T\boldsymbol{x}_2)]\Big]\Big| \\
&\le \mathop{\mathbb{E}}_{\boldsymbol{x}_1,\boldsymbol{x}_2\sim(\mathcal{X}^n)^*}\Big[\Big|\mathop{\mathbb{E}}_{\boldsymbol{e}_2\sim\mathcal{X}^n}[\chi_{a\boldsymbol{x}_1}(\boldsymbol{e}_2)] \cdot \mathop{\mathbb{E}}_{\boldsymbol{e}_1\sim\mathcal{X}^n}[\chi_{a\boldsymbol{x}_2}(-\boldsymbol{e}_1)]\Big|\Big] \\
&\le \mathop{\mathbb{E}}_{\boldsymbol{x}_1\sim(\mathcal{X}^n)^*}\Big[\Big|\mathop{\mathbb{E}}_{\boldsymbol{e}_2\sim\mathcal{X}^n}[\chi_{a\boldsymbol{x}_1}(\boldsymbol{e}_2)]\Big|\Big]
\end{aligned}$$

where the second line follows from triangle inequality and the independence of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, the third line is because $\mathbb{E}_{\boldsymbol{e}_2\sim\mathcal{X}^n}[\chi_{a\boldsymbol{x}_1}(\boldsymbol{e}_2)]$ and $\mathbb{E}_{\boldsymbol{e}_1\sim\mathcal{X}^n}[\chi_{a\boldsymbol{x}_2}(-\boldsymbol{e}_1)]$ are reals of absolute value at most 1. Observe that for any fixed $\boldsymbol{x}_1$ from $(\mathcal{X}^n)^*$, $a\boldsymbol{x}_1 \neq \mathbf{0}^n$ so that $\big|\mathbb{E}_{\boldsymbol{e}_2\sim\mathcal{X}^n}[\chi_{a\boldsymbol{x}_1}(\boldsymbol{e}_2)]\big|$ is at most $\max_{\boldsymbol{c}\in\mathbb{Z}_q^n\setminus\{\mathbf{0}^n\}} |\mathbb{E}_{\boldsymbol{e}\sim\mathcal{X}^n}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]|$ and the desired conclusion follows.

Claim 2 *For any* $\boldsymbol{c} \in \mathbb{Z}_q^n \setminus \{\mathbf{0}^n\}$, $|\mathbb{E}_{\boldsymbol{e}\sim\mathcal{X}^n}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]| \le 1 - \Omega(1/q^2)$.

Proof. Because each coordinate of $\boldsymbol{e}$ is drawn independently from $\mathcal{X}$,

$$\mathop{\mathbb{E}}_{\boldsymbol{e}\sim\mathcal{X}^n}[\chi_{\boldsymbol{c}}(\boldsymbol{e})] = \prod_{i=1}^{n} \mathop{\mathbb{E}}_{z\sim\mathcal{X}}[\chi_{c_i}(z)].$$

Since $\mathcal{X}$ is symmetric, for any $i \in [n]$, $\mathbb{E}_{z\sim\mathcal{X}}[\chi_{c_i}(z)]$ is real with absolute value at most 1. Therefore, it suffices to show that $|\mathbb{E}_{z\sim\mathcal{X}}[\chi_{c_i}(z)]| \leq 1 - \Omega(1/q^2)$ for an arbitrary $i \in [n]$. Fix an $i \in [n]$ such that $c_i \neq 0$ and observe that

$$\mathop{\mathbb{E}}_{z\sim\mathcal{X}}[\chi_{c_i}(z)] \;\leq\; 1 - \Pr_{z\sim\mathcal{X}}[c_i z \neq 0] \cdot \Omega\left(\frac{1}{q^2}\right),$$

because if $c_i z \neq 0$, then the real part of $\chi_{c_i}(z)$ is at most $\cos(\frac{2\pi}{q}) \leq 1 - (1/q^2)$[11]. Similarly,

$$\mathop{\mathbb{E}}_{z\sim\mathcal{X}}[\chi_{c_i}(z)] \;\geq\; -1 + \Pr_{z\sim\mathcal{X}}[c_i z \neq q/2] \cdot \Omega\left(\frac{1}{q^2}\right)$$

holds because if $c_i z \neq q/2$, then the real part of $\chi_{c_i}(z)$ is at least $\cos(\pi + \frac{2\pi}{q}) \geq -1+(1/q^2)$[12]. By our assumption on $\mathcal{X}$, we have that $\Pr_{z\sim\mathcal{X}}[c_i z \neq q/2] \geq 0.1$ and $\Pr_{z\sim\mathcal{X}}[c_i z \neq 0] \geq 0.1$. Hence, $|\mathbb{E}_{z\sim\mathcal{X}}[\chi_{c_i}(z)]| \leq 1 - \Omega(1/q^2)$ which concludes the proof.

For the interested reader, we provide a more self-contained proof in Appendix A which is equivalent to an unrolling of the above proof, but is much more succinct because we do not use the more general statement of Lemma 1 about maximal correlation. In Appendix A, we also give an extension of the proof to the case of Ring-LWE.

4 (Computational) Impossibility of Noise-Ignorant Key Reconciliation Functions

Let us set up some basic notation. For distributions $\mathcal{X}, \mathcal{Y}$ over G, we use $\mathrm{RD}_2(\mathcal{X}||\mathcal{Y}) = \mathbb{E}_{a\sim\mathcal{X}}\left[\Pr_{x\sim\mathcal{X}}[x = a]/\Pr_{y\sim\mathcal{Y}}[y = a]\right]$ to denote the powers of their Rényi divergence [vEH14]. We use $1 + \mathcal{X}$ to denote the distribution which samples x from $\mathcal{X}$ then outputs $1 + x$. And $\mathcal{X} + \mathcal{X}'$ is the distribution obtained as $x + x'$ for $x \sim \mathcal{X}$ and $x' \sim \mathcal{X}'$.

Theorem 3. *Let $n \geq 1, q = \mathsf{poly}(n), m = \mathsf{poly}(n)$ be integers and $\mathcal{X}$ be a noise distribution over $\mathbb{Z}_q$ such that $\mathrm{RD}_2(1 + \mathcal{X}||\mathcal{X}) = 1 + \gamma$. Let $\mu_{\mathcal{X}}(\boldsymbol{X}, \boldsymbol{Y})$ be the joint distribution of*

$$\boldsymbol{X} = (\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \boldsymbol{b}_2) \text{ and } \boldsymbol{Y} = (\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{b}_1),$$

where $\boldsymbol{A} \sim \mathcal{U}(\mathbb{Z}_q^{n\times n})$, $\boldsymbol{e}_1, \boldsymbol{e}_2 \sim \mathcal{X}^n$ and $\boldsymbol{x}_1, \boldsymbol{x}_2 \sim \mathcal{X}^n$, $\boldsymbol{b}_1 = \boldsymbol{x}_1^T\boldsymbol{A} + \boldsymbol{e}_1^T$ and $\boldsymbol{b}_2 = \boldsymbol{A}\boldsymbol{x}_2 + \boldsymbol{e}_2$.
Suppose that f and g are efficiently computable boolean functions that reach key agreement with error at most ε. The domains of Rec_1 and Rec_2 are the support of the marginal distributions $\mu_{\boldsymbol{X}}^{\otimes m}$ and $\mu_{\boldsymbol{Y}}^{\otimes m}$ respectively. Then, m independent samples of $(\boldsymbol{A}, \boldsymbol{b}_2)$ can be efficiently distinguished from m independent samples $(\boldsymbol{A}, \boldsymbol{u})$ where $\boldsymbol{u} \sim \mathcal{U}(\mathbb{Z}_q^n)$ with advantage at least $\Omega(1/q^4 m\gamma) - O(\sqrt{\varepsilon})$.

[11] Because for $x \in [-\pi/2, \pi/2]$, $\cos(x) \leq 1 - x^2/(4\pi^2)$.
[12] Because for $x \in [-\pi/2, \pi/2]$, $\cos(\pi + x) \geq -1 + x^2/(4\pi^2)$.

Our theorem also holds for the ring case. This theorem implies that as long as $\mathrm{RD}_2(1+\mathcal{X}||\mathcal{X})$ is polynomial in n and one party's key reconciliation function does not depend on its noise, then (ring) LWE samples (associated with error $\mathcal{X}$) are not pseudorandom. The condition of $\mathrm{RD}_2(1+\mathcal{X}||\mathcal{X})$ captures a large class of noise distributions including the discrete Gaussian distribution [13].
Let $\mathcal{X}'$ over $\mathbb{Z}_q$ be the distribution that outputs 1 with probability $\alpha = \sqrt{1/m\gamma}$ and outputs 0 otherwise. Let $\mathcal{Z} = \mathcal{U}(\mathbb{Z}_q)^{n\times n} \times \mathcal{X}^n \times \mathcal{X}^n$. Theorem 3 follows from the next two lemmas.

Lemma 4. *Let $\{\mathbf{U}_i\}_{i=1}^m \sim \mathcal{Z}^{\otimes m}$, $\{\boldsymbol{u}_i\}_{i=1}^m, \sim \mathcal{U}(\mathbb{Z}_q^n)^{\otimes m}$, $\{\boldsymbol{u}'_i\}_{i=1}^m \sim \mathcal{U}(\mathbb{Z}_q^n)^{\otimes m}$ and $\{\boldsymbol{w}_i\}_{i=1}^m \sim (\mathcal{X}'^n)^{\otimes m}$. Then,*

$$\begin{aligned}&\Pr[f(\{\mathbf{U}_i, \boldsymbol{u}_i\}_{i=1}^m) \neq f(\{\mathbf{U}_i, \boldsymbol{u}'_i\}_{i=1}^m)]\\ &\leq \Pr[f(\{\mathbf{U}_i, \boldsymbol{u}_i\}_{i=1}^m) \neq f(\{\mathbf{U}_i, \boldsymbol{u}_i + \boldsymbol{w}_i\}_{i=1}^m)] \cdot O\left(q^2\sqrt{m\gamma}\right).\end{aligned}$$

Lemma 5. *Let $\boldsymbol{b}_i = \boldsymbol{A}_i\boldsymbol{x}_i + \boldsymbol{e}_i$ and $\boldsymbol{b}'_i = \boldsymbol{A}_i\boldsymbol{x}'_i + \boldsymbol{e}'_i$, where $\{\boldsymbol{A}_i\}_{i=1}^m \sim \mathcal{U}(\mathbb{Z}_q^{n\times n})^{\otimes m}$ and $\{\boldsymbol{x}_i\}_{i=1}^m, \{\boldsymbol{e}_i\}_{i=1}^m, \{\boldsymbol{x}'_i\}_{i=1}^m, \{\boldsymbol{e}'_i\}_{i=1}^m \sim (\mathcal{X}^n)^{\otimes m}$ and let $\{\boldsymbol{y}_i\}_{i=1}^m \sim (\mathcal{X}^n)^{\otimes 2}$, $\{\boldsymbol{w}_i\}_{i=1}^m \sim (\mathcal{X}'^n)^{\otimes m}$. It holds that*

$$\Pr[f(\{\boldsymbol{A}_i, \boldsymbol{y}_i, \boldsymbol{b}_i\}_{i=1}^m) \neq f(\{\boldsymbol{A}_i, \boldsymbol{y}_i, \boldsymbol{b}'_i\}_{i=1}^m)] \geq 1/2 - 2\varepsilon, \tag{3}$$

and

$$\Pr[f(\{\boldsymbol{A}_i, \boldsymbol{y}_i, \boldsymbol{b}_i + \boldsymbol{w}_i\}_{i=1}^m) \neq f(\{\boldsymbol{A}_i, \boldsymbol{y}_i, \boldsymbol{b}_i\}_{i=1}^m)] \leq O(\sqrt{\varepsilon}). \tag{4}$$

We first prove Theorem 3 using Lemmas 4 and 5. In the rest of this section, we prove Lemmas 4 and 5. Lemma 4 is based on Fourier analysis and works for any boolean function f. Lemma 5 relies on the assumption that f, g are efficient key reconciliation functions and g does not depend on its noise.

4.1 Proof of Theorem 3

Let f and g be key reconciliation functions satisfying the conditions of Theorem 3. We wish to distinguish between m i.i.d. samples $\{(\boldsymbol{A}_i, \boldsymbol{b}_i)\}_{i=1}^m$ from m i.i.d. samples $\{(\boldsymbol{A}_i, \boldsymbol{u}_i)\}_{i=1}^m$.
First, note that if

$$|\Pr[f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{u}_i\}_{i=1}^m) = 0] - \Pr[f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{b}_i\}_{i=1}^m) = 0]| \geq \alpha/q^2,$$

where $\{\boldsymbol{x}_i\}_{i=1}^m, \{\boldsymbol{e}_i\}_{i=1}^m \sim (\mathcal{X}^n)^{\otimes m}$, there exists a polynomial time distinguisher, since $\boldsymbol{x}_i$ and $\boldsymbol{e}_i$ are efficiently sampleable.
Otherwise, from Equation (3) of Lemma 5, we have that

$$\Pr[f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{u}_i\}_{i=1}^m) \neq f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{u}'_i\}_{i=1}^m)] \geq 2\varepsilon + 2\alpha/q^2,$$

[13] In particular, Bogdanov et al. [BGM+16] showed that $\mathrm{RD}_2(1+\mathcal{D}_\sigma||\mathcal{D}_\sigma) \leq \exp(2\pi(1/\sigma)^2)$ is at most a constant for any discrete Gaussian distribution $\mathcal{D}_\sigma$ with standard deviation $\sigma \geq 1$.

where $\boldsymbol{u}_i' \sim \mathcal{U}(\mathbb{Z}_q^n)$. Combining this with Lemma 4, we get that

$$\Pr[f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{u}_i\}_{i=1}^m) \neq f\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{u}_i + \boldsymbol{w}_i\}_{i=1}^m)] \geq \Omega\left(\frac{\alpha^2}{q^4} + \frac{\alpha\varepsilon}{q^2}\right),$$

where $\boldsymbol{w}_i \sim \mathcal{X}'^n$. But, from Equation (4) of Lemma 5, we have that

$$\Pr[f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{b}_i\}_{i=1}^m) \neq f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{b}_i + \boldsymbol{w}_i\}_{i=1}^m)] \leq O(\sqrt{\varepsilon})$$

Thus, we distinguish between m i.i.d. samples $\{(\boldsymbol{A}_i, \boldsymbol{u}_i)\}_{i=1}^m$ and $\{(\boldsymbol{A}_i, \boldsymbol{b}_i)\}_{i=1}^m$ by computing $\Pr[f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{y}_i\}_{i=1}^m) \neq f(\{\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{e}_i, \boldsymbol{y}_i + \boldsymbol{w}_i\}_{i=1}^m)]$, where $\{\boldsymbol{y}_i\}_{i=1}^m$ are the challenge samples. This gives us an advantage of $\Omega(\alpha^2/q^4) - O(\sqrt{\varepsilon})$.

4.2 Proof of Lemma 4

Let $\mathsf{Re}(z)$ denote the real part of any $z \in \mathbb{C}$. We fix $\{\mathbf{U}_i\}_{i=1}^m$ and for any $\boldsymbol{u} = \{\boldsymbol{u}_i\}_{i=1}^m \in (\mathbb{Z}_q^n)^{\otimes m}$, let $F(\boldsymbol{u}) = (-1)^{f(\{(\mathbf{U}_i, \boldsymbol{u}_i)\}_{i=1}^m)}$, then

$$\Pr\left[f(\{(\mathbf{U}_i, \boldsymbol{u}_i + \boldsymbol{w}_i)\}_{i=1}^m) \neq f(\{(\mathbf{U}_i, \boldsymbol{u}_i)\}_{i=1}^m))\right] = \frac{1 - \mathbb{E}[F(\boldsymbol{u})F(\boldsymbol{u} + \boldsymbol{w})]}{2},$$

where $\boldsymbol{u} \sim \mathcal{U}(\mathbb{Z}_q^n)^{\otimes m}, \boldsymbol{w} \sim (\mathcal{X}'^n)^{\otimes m}$ and $\boldsymbol{w} = \{\boldsymbol{w}_i\}_{i=1}^m$.
For any $\boldsymbol{c} \in (\mathbb{Z}_q^n)^m$, let $\widehat{F}(\boldsymbol{c}) = \mathbb{E}_{\boldsymbol{u} \sim \mathcal{U}(\mathbb{Z}_q^n)^{\otimes m}}[F(\boldsymbol{u})\chi_{\boldsymbol{c}}(-\boldsymbol{u})]$. Note that for any $\boldsymbol{u} \in (\mathbb{Z}_q^n)^m$, $F(\boldsymbol{u}) = \sum_{\boldsymbol{c} \in (\mathbb{Z}_q^n)^m} \widehat{F}(\boldsymbol{c})\chi_{\boldsymbol{c}}(\boldsymbol{u})$. Finally, because F is real, $\mathbb{E}[F(\boldsymbol{u})F(\boldsymbol{u} + \boldsymbol{w})] = \mathbb{E}[\overline{F(\boldsymbol{u})}F(\boldsymbol{u} + \boldsymbol{w})]$.

$$\begin{aligned}
&\mathbb{E}[\overline{F(\boldsymbol{u})}F(\boldsymbol{u} + \boldsymbol{w})] \\
&= \left|\widehat{F}(\mathbf{0}^{nm})\right|^2 + \sum_{\boldsymbol{c} \in (\mathbb{Z}_q^n)^m \setminus \{\mathbf{0}^{nm}\}} \left|\widehat{F}(\boldsymbol{c})\right|^2 \mathbb{E}[\chi_{\boldsymbol{c}}(\boldsymbol{w})] \\
&= \left|\widehat{F}(\mathbf{0}^{nm})\right|^2 + \sum_{\boldsymbol{c} \in (\mathbb{Z}_q^n)^m \setminus \{\mathbf{0}^{nm}\}} \left|\widehat{F}(\boldsymbol{c})\right|^2 \mathbb{E}[\mathsf{Re}(\chi_{\boldsymbol{c}}(\boldsymbol{w}))] \\
&\leq \left|\widehat{F}(\mathbf{0}^{nm})\right|^2 + \left(\max_{\boldsymbol{c} \in (\mathbb{Z}_q^n)^m \setminus \{\mathbf{0}^{nm}\}} \mathbb{E}[\mathsf{Re}(\chi_{\boldsymbol{c}}(\boldsymbol{w}))]\right)\left(\sum_{\boldsymbol{c} \in (\mathbb{Z}_q^n)^m \setminus \{\mathbf{0}^{nm}\}} \left|\widehat{F}(\boldsymbol{c})\right|^2\right) \\
&\leq \left|\widehat{F}(\mathbf{0}^{nm})\right|^2 + \left(\max_{\boldsymbol{c} \in (\mathbb{Z}_q^n)^m \setminus \{\mathbf{0}^{nm}\}} \mathbb{E}[\mathsf{Re}(\chi_{\boldsymbol{c}}(\boldsymbol{w}))]\right)\left(1 - \left|\widehat{F}(\mathbf{0}^{nm})\right|^2\right)
\end{aligned}$$

where the first line is by expanding F using its Fourier representation and linearity of expectation, the second line is because $\mathbb{E}[\overline{F(\boldsymbol{u})}F(\boldsymbol{u} + \boldsymbol{w})]$ is real, and the last line uses Parseval's identity, which states that $\sum_{\boldsymbol{c}} \left|\widehat{F}(\boldsymbol{c})\right|^2 = \mathbb{E}\left[|F(\boldsymbol{u})|^2\right] = 1$.
Similarly to the analysis of Claim 2, $\max_{\boldsymbol{c} \in (\mathbb{Z}_q^n)^m \setminus \{\mathbf{0}^{nm}\}} \mathbb{E}[\mathsf{Re}(\chi_{\boldsymbol{c}}(\boldsymbol{w}))] \leq 1 - \Omega(\alpha/q^2)$, because for any $\boldsymbol{c} \neq \mathbf{0}^{nm}$, $\Pr[\boldsymbol{c}^T\boldsymbol{w} \neq 0] \geq \alpha$ and $\mathsf{Re}(\chi_{\boldsymbol{c}}(\boldsymbol{w})) \leq 1 - \Omega(1/q^2)$ whenever $\boldsymbol{c}^T\boldsymbol{w} \neq 0$. Therefore,

$$\begin{aligned}
&\Pr_{\boldsymbol{u} \sim \mathcal{U}(\mathbb{Z}_q^n)^{\otimes m}, \boldsymbol{w} \sim (\mathcal{X}'^n)^m}\left[f(\{(\mathbf{U}_i, \boldsymbol{u}_i + \boldsymbol{w}_i)\}_{i=1}^m) \neq f(\{(\mathbf{U}_i, \boldsymbol{u}_i)\}_{i=1}^m))\right] \\
&\geq \Omega(\alpha/q^2)\frac{1 - \left|\widehat{F}(\mathbf{0}^{nm})\right|^2}{2}.
\end{aligned}$$

Since $\Pr_{\boldsymbol{u},\boldsymbol{u}'\sim\mathcal{U}(\mathbb{Z}_q^n)^{\otimes m}}\left[f(\{(\mathbf{U}_i,\boldsymbol{u}_i)\}_{i=1}^m) \neq f(\{(\mathbf{U}_i,\boldsymbol{u}_i')\}_{i=1}^m))\right] = \frac{1-\left|\widehat{F}(\mathbf{0}^{nm})\right|^2}{2}$, the lemma follows by averaging over $\{\mathbf{U}_i\}_{i=1}^m$.

4.3 Proof of Lemma 5

Let $\{\boldsymbol{y}_i\}_{i=1}^m = \{(\boldsymbol{x}_i'', \boldsymbol{e}_i'')\}_{i=1}^m$, $\boldsymbol{b}_i'' = \boldsymbol{A}_i\boldsymbol{x}_i'' + \boldsymbol{e}_i''$ and suppose Equation (3) is not true, then together with the correctness condition, it holds that

$$\Pr[g(\{(\boldsymbol{A}_i, \boldsymbol{x}_i', \boldsymbol{b}_i'')\}_{i=1}^m) = f(\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{b}_i')\}_{i=1}^m)] > 1/2 + \varepsilon,$$

which breaks the soundness condition because an adversary can sample fresh $\{\boldsymbol{x}_i'\}_{i=1}^m \sim (\mathcal{X}^n)^{\otimes m}$ and compute $g(\{(\boldsymbol{A}_i, \boldsymbol{x}_i', \boldsymbol{b}_i'')\}_{i=1}^m)$.
To prove Equation (4), we first show the following two claims

Claim.

$$\begin{aligned}&\Pr[f(\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{b}_i + \boldsymbol{w}_i)\}_{i=1}^m) \neq g(\{(\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{b}_i'')\}_{i=1}^m))]\\ &\leq \sqrt{\varepsilon \cdot \mathrm{RD}_2^m(\mathcal{X} + \mathcal{X}' || \mathcal{X})}.\end{aligned}$$

Proof. We rely on two elementary properties of Rényi divergence: for any two distributions $\boldsymbol{X}$ and $\boldsymbol{Y}$ and any event E, $(\Pr[\boldsymbol{X} \in E])^2 \leq \Pr[\boldsymbol{Y} \in E] \cdot \mathrm{RD}_2(\boldsymbol{X}||\boldsymbol{Y})$, and for any m, $\mathrm{RD}_2(\boldsymbol{X}^m||\boldsymbol{Y}^m) = (\mathrm{RD}_2(\boldsymbol{X}||\boldsymbol{Y}))^m$. For any fixed choice of $\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{x}_i)\}_{i=1}^m$, let E be the event that f disagrees with g. Then, by the properties of Rényi divergence,

$$\begin{aligned}&\left(\Pr[f(\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{b}_i + \boldsymbol{w}_i)\}_{i=1}^m) \neq g(\{(\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{b}_i'')\}_{i=1}^m)]\right)^2\\ &\leq \Pr[f(\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{b}_i)\}_{i=1}^m) \neq g(\{(\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{b}_i'')\}_{i=1}^m)] \cdot \mathrm{RD}_2((\mathcal{X} + \mathcal{X}')^{\otimes m} || (\mathcal{X})^{\otimes m})\\ &= \Pr[f(\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{b}_i)\}_{i=1}^m) \neq g(\{(\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{b}_i'')\}_{i=1}^m)] \cdot \left(\mathrm{RD}_2(\mathcal{X} + \mathcal{X}' || \mathcal{X})\right)^m.\end{aligned}$$

The desired conclusion follows by averaging over $\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{x}_i)\}_{i=1}^m$ and the fact that for any random variable z, $(\mathbb{E}[z])^2 \leq \mathbb{E}[z^2]$.

Claim. $\mathrm{RD}_2(\mathcal{X} + \mathcal{X}' || \mathcal{X}) = 1 + \alpha^2\gamma$

Proof. By the definition of RD_2 and $\mathcal{X}'$,

$$\begin{aligned}&\mathrm{RD}_2(\mathcal{X} + \mathcal{X}' || \mathcal{X})\\ &= \sum_{a \in G} \frac{((1-\alpha)\Pr_{X\sim\mathcal{X}}[X = a] + \alpha \Pr_{X\sim\mathcal{X}}[X + 1 = a])^2}{\Pr_{X\sim\mathcal{X}}[X = a]}\\ &= (1-\alpha)^2 + 2(1-\alpha)\alpha + \alpha^2 \mathrm{RD}_2(\mathcal{X} + 1 || \mathcal{X})\\ &= 1 + \alpha^2(\mathrm{RD}_2(\mathcal{X} + 1 || \mathcal{X}) - 1).\end{aligned}$$

From the correctness condition, which is

$$\Pr[f(\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{b}_i)\}_{i=1}^m) \neq g(\{(\boldsymbol{A}_i, \boldsymbol{x}_i, \boldsymbol{b}_i'')\}_{i=1}^m)] \leq \varepsilon$$

and the above two claims and union bound,

$$\begin{aligned}&\Pr[f(\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{b}_i + \boldsymbol{w}_i)\}_{i=1}^m) \neq f(\{(\boldsymbol{A}_i, \boldsymbol{x}_i'', \boldsymbol{e}_i'', \boldsymbol{b}_i)\}_{i=1}^m)]\\ &\leq \varepsilon + \sqrt{\varepsilon(1 + \alpha^2\gamma)^m}.\end{aligned}$$

The Equation (4) follows from our choice of $\alpha = \sqrt{1/m\gamma}$.

5 Connections to other cryptographic primitives

Thus far, our results focused on specific classes of reconciliation functions showing that they are not powerful enough to give NIKE in our framework. Extending our previous results either on the positive or negative direction hits barriers. The positive direction, which is to propose a NIKE protocol that avoids our impossibility results implies cryptographic constructions still unknown from polynomial modulus LWE. In particular, a positive result would imply direct constructions of special structured weak pseudorandom functions from polynomial modulus LWE. The negative direction, which is to prove a completely general impossibility result, is ruled out if iO exists.

5.1 From iO To NIKE

Even though our results show that there are many limitations in building practical NIKE from polynomial modulus LWE, assuming indistinguishability obfuscation (iO) constructing NIKE is, at least theoretically, possible. Therefore, unless there are breakthrough advancements that rule out the possibility of construction iO, showing a general impossibility of NIKE is out of range. In this section, we sketch the iO-based construction of NIKE of Boneh and Zhandry [BZ17] and explain why it can be implemented in our framework.

Theorem 4 ([BZ17]). *Assuming a secure pseudorandom generator, a secure punctured pseudorandom function family and a secure indistinguishability obfuscator, there exists a secure* NIKE.

Additionally to the matrix $\boldsymbol{A}$, in this protocol the parties share the following *obfuscated* program:

Inputs: $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \mathcal{X}, s_1, s_2 \in \mathcal{S}$
Constants: PRF
If $\boldsymbol{b}_1 = \mathrm{PRG}(s_1)$, output $\mathrm{PRF}(\boldsymbol{b}_1, \boldsymbol{b}_2)$.
If $\boldsymbol{b}_2 = \mathrm{PRG}(s_2)$, output $\mathrm{PRF}(\boldsymbol{b}_1, \boldsymbol{b}_2)$.
Otherwise, output $\perp$.

During the protocol, the parties exchange LWE samples $\boldsymbol{b}_1, \boldsymbol{b}_2$, evaluate the obfuscated program with $s_1 = (\boldsymbol{x}_1, \boldsymbol{e}_1)$ and $s_2 = (\boldsymbol{x}_2, \boldsymbol{e}_2)$ and set as their shared key the output of the obfuscated program. The LWE samples are computed from a function of the form $G_{\mathbf{M}}(\boldsymbol{x}, \boldsymbol{e}) = \mathbf{M}\boldsymbol{x} + \boldsymbol{e}$, where $\mathbf{M} \in \mathbb{Z}_q^{n \times n}$ and $\boldsymbol{x}, \boldsymbol{e}$ are sampled from a noise distribution. Directly using the LWE assumptions, which states that the output of G is indistinguishable from uniform and the fact that G is expanding, we conclude that G is a PRG. Combining this observation with the known constructions of punctured PRFs from any one-way function, we conclude that there exists a NIKE protocol assuming iO and polynomial modulus LWE.

5.2 From NIKE To weak-PRFs

In this section, we show that reconciliation functions have to be *weak-pseudorandom functions*. A weak-pseudorandom function (weak-PRF) is an efficient function family that is indistinguishable from a random function when we have access only on random evaluations of the function. We focus on the case of boolean weak-pseudorandom functions. Formally:

Definition 8. *Let $\lambda > 0$ be a security parameter. An efficient function family ensemble $\mathcal{F} = \{\mathcal{F}_\lambda : \{0,1\}^{k(\lambda)} \to \{0,1\}\}$ is called weak-pseudorandom function family if for every probabilistic polynomial-time algorithm $\mathcal{A}$:*

$$\Pr_{f,x}[\mathcal{A}^{\mathcal{O}_f}(x) = f(x)] \leq 1/2 + \mathsf{negl}(\lambda),$$

where f is sampled uniformly at random from $\mathcal{F}_\lambda$ and $x \sim U(\{0,1\}^{k(\lambda)})$. Every query to the oracle $\mathcal{O}$ is answered with a tuple of the form $(u, f(u))$, where $u \sim \mathcal{U}(\{0,1\}^{k(\lambda)})$. We call $\left|\Pr_{f,x}[\mathcal{A}^{\mathcal{O}_f}(x) = f(x)] - 1/2\right|$ the success probability of $\mathcal{A}$.

The main theorem of this section shows that the reconciliation functions have to be sampled from a weak-PRF family.

Theorem 5. *Let $\lambda > 0$ be a security parameter and let $f(\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \boldsymbol{b}_2)$ and $g(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1)$ be efficient functions such that:*

- $\Pr[f(\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \boldsymbol{b}_2) = g(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1)] \geq 1 - \mathsf{negl}(\lambda)$
- *For every efficient probabilistic polynomial-time algorithm $\mathcal{D}$ with input $(\boldsymbol{A}, \boldsymbol{b}_1, \boldsymbol{b}_2)$:*

$$\Pr[\mathcal{D}(\boldsymbol{A}, \boldsymbol{b}_1, \boldsymbol{b}_2) = f(\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \boldsymbol{b}_2)] \leq 1/2 + \mathsf{negl}(\lambda),$$

then assuming the LWE *assumption, the function families $\mathcal{F} = \{F_{\boldsymbol{A},\boldsymbol{x}_1,\boldsymbol{e}_1} : \mathbb{Z}_q^n \to \{0,1\}\}$, where $F_{\boldsymbol{A},\boldsymbol{x}_1,\boldsymbol{e}_1}(\cdot) = f(\boldsymbol{A}, \boldsymbol{x}_1, \boldsymbol{e}_1, \cdot)$ and $\mathcal{G} = \{G_{\boldsymbol{A},\boldsymbol{x}_2,\boldsymbol{e}_2} : \mathbb{Z}_q^n \to \{0,1\}\}$, where $G_{\boldsymbol{A},\boldsymbol{x}_2,\boldsymbol{e}_2}(\cdot) = g(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \cdot)$ are weak-PRF families.*

Even though we formally prove that the reconciliation functions should be pseudorandom with access to random evaluations of the functions, they have to satisfy a stronger pseudorandomness property: they should remain pseudorandom even with access to evaluations of *adversarially chosen* LWE samples. Also, our result directly generalizes to the case of *multiple* LWE samples. In fact, the above theorem can be extended to show that in a NIKE protocol where the exchanged messages are indistinguishable from uniform, reconciliation functions have to be sampled from a weak-PRF function family.

Although (weak-)PRFs are equivalent to one-way function [GGM86], the known generic constructions are highly inefficient and unstructured. Constructions of (weak-)PRFs from LWE are only known for superpolynomial modulus [BPR12,BP14] and finding a direct construction based on polynomial modulus is a very interesting open problem in the study of pseudorandom functions [BR17]. We emphasize that even though pseudorandomness is a necessary condition for a reconciliation function and identifies a barrier in building NIKE from LWE, it is definitely not sufficient. Reconciliation functions are very structured as the computation of the common key should be allowed in at least two ways, one for Alice and one for Bob.

Proof. We show that $\mathcal{F}$ is a weak-PRF family and the same analysis holds for $\mathcal{G}$. Assume that there exists a distinguisher $\mathcal{A}$ for $\mathcal{F}$ with success probability α; we use $\mathcal{A}$ to break the soundness of the NIKE protocol. From the correctness condition of NIKE,

$$\Pr[F_{\boldsymbol{A},\boldsymbol{x}_1,\boldsymbol{e}_1}(\boldsymbol{b}_2) = g(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1)] \geq 1 - \mathsf{negl}(\lambda).$$

Hence, with high probability we get evaluations of $F_{\boldsymbol{A},\boldsymbol{x}_1,\boldsymbol{e}_1}$ by sampling LWE secret and noise $\boldsymbol{x}_2, \boldsymbol{e}_2$ and computing $g(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1)$. Additionally, the LWE assumption implies that these evaluations of F are computationally indistinguishable from uniform evaluations, as required by the definition of weak-PRFs.
An adversary $\mathcal{D}$ that breaks the soundness condition of NIKE runs as follows:

- Run the distinguisher $\mathcal{A}$, where instead of uniform evaluations compute evaluations using LWE samples and g as above.
- Use as the challenge query $\boldsymbol{b}_2$.
- Return the output of $\mathcal{A}$.

Let us denote by $\mathcal{E}$ the event that $F_{\boldsymbol{A},\boldsymbol{x}_1,\boldsymbol{e}_1}(\boldsymbol{b}_2) = g(\boldsymbol{A}, \boldsymbol{x}_2, \boldsymbol{e}_2, \boldsymbol{b}_1)$, the success probability of $\mathcal{D}$ is equal to

$$\begin{aligned}
&\Pr[\mathcal{D}(\boldsymbol{A}, \boldsymbol{b}_1, \boldsymbol{b}_2) = F_{\boldsymbol{A},\boldsymbol{x}_1,\boldsymbol{e}_1}(\boldsymbol{b}_2)] \\
&\geq \Pr[\mathcal{D}(\boldsymbol{A}, \boldsymbol{b}_1, \boldsymbol{b}_2) = F_{\boldsymbol{A},\boldsymbol{x}_1,\boldsymbol{e}_1}(\boldsymbol{b}_2) | \mathcal{E}] \Pr[\mathcal{E}] \\
&= \Pr[\mathcal{A}(\boldsymbol{b}_2) = F_{\boldsymbol{A},\boldsymbol{x}_1,\boldsymbol{e}_1}(\boldsymbol{b}_2)] \Pr[\mathcal{E}] \\
&\geq 1/2 + \alpha - \mathsf{negl}(\lambda).
\end{aligned}$$

Hence if $\mathcal{A}$ breaks $\mathcal{F}$, then D breaks the soundness condition of NIKE.

Acknowledgements

The authors thank Martin Albrecht, Jacob Alperin-Sheriff, Leo Ducas and anonymous reviewers for useful comments and advice.

References

ADPS16. Erdem Alkim, Léo Ducas, Thomas Pöppelmann, and Peter Schwabe. Post-quantum key exchange - A new hope. In *25th USENIX Security Symposium, USENIX Security 16, Austin, TX, USA, August 10-12, 2016.*, pages 327–343, 2016. 2

BCNS14. Joppe W. Bos, Craig Costello, Michael Naehrig, and Douglas Stebila. Post-quantum key exchange for the TLS protocol from the ring learning with errors problem. *IACR Cryptology ePrint Archive*, 2014:599, 2014. 2

BGM+16. Andrej Bogdanov, Siyao Guo, Daniel Masny, Silas Richelson, and Alon Rosen. On the hardness of learning with rounding over small modulus. In *Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part I*, pages 209–224, 2016. 13

BP14. Abhishek Banerjee and Chris Peikert. New and improved key-homomorphic pseudorandom functions. In *Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part I*, pages 353–370, 2014. 17

BPR12. Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In *Advances in Cryptology - EUROCRYPT 2012 - 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cambridge, UK, April 15-19, 2012. Proceedings*, pages 719–737, 2012. 17

BR17. Andrej Bogdanov and Alon Rosen. Pseudorandom functions: Three decades later. In *Tutorials on the Foundations of Cryptography.*, pages 79–158. 2017. 17

BZ17. Dan Boneh and Mark Zhandry. Multiparty key exchange, efficient traitor tracing, and more from indistinguishability obfuscation. *Algorithmica*, 79(4):1233–1285, 2017. 6, 16

DH76. Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Trans. Information Theory*, 22(6):644–654, 1976. 2

DXL12. Jintai Ding, Xiang Xie, and Xiaodong Lin. A simple provably secure key exchange scheme based on the learning with errors problem. Cryptology ePrint Archive, Report 2012/688, 2012. http://eprint.iacr.org/2012/688. 2

Geb41. Hans Gebelein. Das statistische problem der korrelation als variations-und eigenwertproblem und sein zusammenhang mit der ausgleichsrechnung. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 21(6):364–379, 1941. 9

GGM86. Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *J. ACM*, 33(4):792–807, 1986. 17

GMZB+17. Oscar Garcia-Morchon, Zhenfei Zhang, Sauvik Bhattacharya, Ronald Rietman, Ludo Tolhuizen, Jose-Luis Torre-Arce, Hayo Baan, Markku-Juhani O. Saarinen, Scott Fluhrer, Thijs Laarhoven, and Rachel Player. Round5. Technical report, National Institute of Standards and Technology, 2017. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions. 2

Hir35. Hermann O Hirschfeld. A connection between correlation and contingency. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 31, pages 520–524. Cambridge Univ Press, 1935. 9

LLJ+17. Xianhui Lu, Yamin Liu, Dingding Jia, Haiyang Xue, Jingnan He, Zhenfei Zhang, Zhe Liu, Hao Yang, Bao Li, and Kunpeng Wang. Lac. Technical report, National Institute of Standards and Technology, 2017. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions. 2, 4

Lov75. László Lovász. Spectra of graphs with transitive groups. *Periodica Math. Hung.*, 6:191–195, 1975. 10

LPR10. Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In *Advances in Cryptology - EUROCRYPT 2010, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Monaco / French Riviera, May 30 - June 3, 2010. Proceedings*, pages 1–23, 2010. 2

NAB+17. Michael Naehrig, Erdem Alkim, Joppe Bos, Leo Ducas, Karen Easterbrook, Brian LaMacchia, Patrick Longa, Ilya Mironov, Valeria Nikolaenko, Christopher Peikert, Ananth Raghunathan, and Douglas Stebila. Frodokem. Technical report, National Institute of Standards and Technology, 2017. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions. 2

NIS. NIST. https://csrc.nist.gov/csrc/media/projects/post-quantum-cryptography/ documents/call-for-proposals-final-dec-2016.pdf. 2

PAA+17. Thomas Poppelmann, Erdem Alkim, Roberto Avanzi, Joppe Bos, Leo Ducas, Antonio de la Piedra, Peter Schwabe, Douglas Stebila, Martin R. Albrecht, Emmanuela Orsini, Valery Osheter, Kenneth G. Paterson, Guy Peer, and Nigel P. Smart. Newhope. Technical report, National Institute of Standards and Technology, 2017. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions. 2

Pei09. Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 333–342, 2009. 4

Pei14. Chris Peikert. Lattice cryptography for the internet. In *Post-Quantum Cryptography - 6th International Workshop, PQCrypto 2014, Waterloo, ON, Canada, October 1-3, 2014. Proceedings*, pages 197–219, 2014. 2

Reg05. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 84–93, 2005. 2, 7

Rén59. Alfréd Rényi. On measures of dependence. *Acta mathematica hungarica*, 10(3-4):441–451, 1959. 9

SAB+17. Peter Schwabe, Roberto Avanzi, Joppe Bos, Leo Ducas, Eike Kiltz, Tancrede Lepoint, Vadim Lyubashevsky, John M. Schanck, Gregor Seiler, and Damien Stehle. Crystals-kyber. Technical report, National Institute of Standards and Technology, 2017. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions. 2

Sho94. Peter W. Shor. Polynominal time algorithms for discrete logarithms and factoring on a quantum computer. In *Algorithmic Number Theory, First International Symposium, ANTS-I, Ithaca, NY, USA, May 6-9, 1994, Proceedings*, page 289, 1994. 2

vEH14. Tim van Erven and Peter Harremoës. Rényi divergence and kullback-leibler divergence. *IEEE Trans. Information Theory*, 60(7):3797–3820, 2014. 12

Wit75. Hans S Witsenhausen. On sequences of pairs of dependent random variables. *SIAM Journal on Applied Mathematics*, 28(1):100–113, 1975. 9

A A self-contained proof of Theorem 1

In Section 3, we use two lemmas (Lemmas 1 and 3) in order to bound the key agreement probability by $\max_{c\in\mathbb{Z}_q\setminus\{0\}} |\,\mathbb{E}_{e\sim\xi}[\chi_c(e)]$. In this section, we give a self-contained proof for Lemma 2 without explicitly using the notion of maximal correlation. However, this proof is essentially an unrolling of the proof using maximal correlation.

Claim 3 *For any $m \geq 1$, and balanced $f, g : \mathbb{Z}_q^m \to \{0,1\}$, it holds that*

$$\Pr[f(\boldsymbol{X}) = g(\boldsymbol{Y})] \leq \frac{1 + \max_{c\in\mathbb{Z}_q\setminus\{0\}} |\,\mathbb{E}_{e\sim\xi}[\chi_c(e)]|}{2}$$

where for any $z \in \mathbb{Z}_q$, $\xi(z) = \Pr[\boldsymbol{x}_1^T\boldsymbol{e}_2 - \boldsymbol{e}_1^T\boldsymbol{x}_2 = z]$.

Combining the above claim with the fact that $\max_{c\in\mathbb{Z}_q\setminus\{0\}} |\,\mathbb{E}_{e\sim\xi}[\chi_c(e)]| \leq 1 - \Omega(1/q^2)$ (see Claim 1 and 2), Theorem 1 follows.

Proof. Let $F(\boldsymbol{x}) = (-1)^{f(\boldsymbol{x})}$ and $G(\boldsymbol{x}) = (-1)^{g(\boldsymbol{x})}$. For any $\boldsymbol{c} \in \mathbb{Z}_q^m$, let $\hat{F}(\boldsymbol{c}) = \mathbb{E}_{\boldsymbol{x}\sim\mathcal{U}(\mathbb{Z}_q^m)}[F(\boldsymbol{x})\chi_{\boldsymbol{c}}(-\boldsymbol{x})]$ and $\hat{G}(\boldsymbol{c}) = \mathbb{E}_{\boldsymbol{x}\sim\mathcal{U}(\mathbb{Z}_q^m)}[G(\boldsymbol{x})\chi_{\boldsymbol{c}}(-\boldsymbol{x})]$. Note that for any $\boldsymbol{x} \in \mathbb{Z}_q^m$, $F(\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathbb{Z}_q^m} \hat{F}(\boldsymbol{c})\chi_{\boldsymbol{c}}(\boldsymbol{x})$ and $G(\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathbb{Z}_q^m} \hat{G}(\boldsymbol{c})\chi_{\boldsymbol{c}}(\boldsymbol{x})$. Observe that $\boldsymbol{X}$ is distributed uniformly and $\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{e}$.

$$\begin{aligned}
&|\,\mathbb{E}[\overline{F(\boldsymbol{X})}G(\boldsymbol{X}+\boldsymbol{e})]| \\
&= \left|\sum_{\boldsymbol{c}\in\mathbb{Z}_q^m\setminus\{0^m\}} \hat{F}(\boldsymbol{c})\hat{G}(\boldsymbol{c})\,\mathbb{E}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]\right| \\
&\leq \sqrt{\sum_{\boldsymbol{c}\in\mathbb{Z}_q^m\setminus\{0^m\}} |\hat{F}(\boldsymbol{c})|^2 \sum_{\boldsymbol{c}\neq 0^m} |\hat{G}(\boldsymbol{c})|^2}\, \max_{\boldsymbol{c}\neq 0^m} |\,\mathbb{E}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]| \\
&\leq \max_{\boldsymbol{c}\in\mathbb{Z}_q^m\setminus\{0^m\}} |\,\mathbb{E}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]| \\
&\leq \max_{c\in\mathbb{Z}_q\setminus\{0\}} |\,\mathbb{E}[\chi_c(\boldsymbol{e}_i)]|,
\end{aligned}$$

where $\boldsymbol{e} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$. The first equality follows by linearity of expectation and the fact that $\boldsymbol{X}$ is uniform over $\mathbb{Z}_q^m$. For the next inequality, we use triangle inequality and that $|\,\mathbb{E}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]|$ is real, since ξ is symmetric. The next two inequalities follow by Cauchy-Schwarz and Parseval's identity, which states that $\sum_{\boldsymbol{c}} \left|\hat{F}(\boldsymbol{c})\right|^2 = \mathbb{E}\left[|F(\boldsymbol{X})|^2\right] = 1$. The desired conclusion follows from the fact that $\Pr[f(\boldsymbol{X}) = g(\boldsymbol{Y})] = (1 + \mathbb{E}[\overline{F(\boldsymbol{X})}G(\boldsymbol{Y})])/2$.

Ring-LWE case. We get a similar result for the Ring-LWE case. Let R_q be the ring $\mathbb{Z}_q[x]/g(x)$ where g is a polynomial of degree n over Z_q. We identify an element in R_q by its coefficient vector in $\mathbb{Z}_q^n$. We say that $\boldsymbol{w}$ is drawn from $(\mathcal{X}^n)^*$ if its coefficients are drawn from $\mathcal{X}^n$ conditioned on $\boldsymbol{w}$ being a unit of R_q.

Theorem 6. *Let $n, q \geq 1$ be integers and R_q be as above. Assume that the distribution $\mathcal{X}$ over $\mathbb{Z}_q$ is symmetric and for any $a \in \mathbb{Z}_q \setminus \{0\}$, $\Pr[az = 0] \leq 9/10$ and $\Pr[az = q/2] \leq 9/10$ and $(\mathcal{X}^n)^*$ as above.*
Let $\mu_{\mathrm{RLWE},\mathcal{X}}(\boldsymbol{X}, \boldsymbol{Y})$ be the joint distribution of

$$\boldsymbol{X} = \boldsymbol{x}_1 \cdot \boldsymbol{a} \cdot \boldsymbol{x}_2 + \boldsymbol{x}_1 \cdot \boldsymbol{e}_2 \text{ and } \boldsymbol{Y} = \boldsymbol{x}_1 \cdot \boldsymbol{a} \cdot \boldsymbol{e}_1 + \boldsymbol{x}_2 \cdot \boldsymbol{e}_1,$$

where $\cdot$ is polynomial multiplication, $\boldsymbol{a} \sim \mathcal{U}(R_q)$, $\boldsymbol{e}_1, \boldsymbol{e}_2 \sim \mathcal{X}^n$ and $\boldsymbol{x}_1, \boldsymbol{x}_2 \sim (\mathcal{X}^n)^*$. Then for any $m \geq 1$, and any balanced functions $f, g : R^m \to \{0,1\}$ respect to the marginal distributions of $\mu_{\mathrm{RLWE},\mathcal{X}}^{\otimes m}$, it holds that

$$\Pr_{(\boldsymbol{X}^m, \boldsymbol{Y}^m) \sim \mu_{\mathrm{RLWE},\mathcal{X}}^{\otimes m}} [f(\boldsymbol{X}^m) = g(\boldsymbol{Y}^m)] \leq 1 - \Omega(1/q^2).$$

Proof. We proceed as in the LWE case by proving claims similar to Claim 1, 2 and 3. For $\boldsymbol{c} \in R_q$, we define $\chi_{\boldsymbol{c}} \colon R_q \to \mathbb{C}$ as $\chi_{\boldsymbol{c}}(\boldsymbol{x}) = e^{-2\pi i \cdot \langle \boldsymbol{c}, \boldsymbol{x} \rangle / q}$, where $\langle \boldsymbol{c}, \boldsymbol{x} \rangle$ is the inner product of the coefficient vectors of $\boldsymbol{c}, \boldsymbol{x}$ over $\mathbb{Z}_q$. Then, the following claims hold.

Claim 4 *For any $m \geq 1$ and balanced $f, g : R_q^m \to \{0,1\}$, it holds that*

$$\Pr[f(\boldsymbol{X}^m) = g(\boldsymbol{Y}^m)] \leq \frac{1 + \max_{\boldsymbol{c} \in R_q \setminus \{\boldsymbol{0}^n\}} |\mathbb{E}_{\boldsymbol{e} \sim \xi}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]|}{2}$$

where for any $\boldsymbol{z} \in R_q$, $\xi(\boldsymbol{z}) = \Pr[\boldsymbol{x}_1 \cdot \boldsymbol{e}_2 - \boldsymbol{e}_1 \cdot \boldsymbol{x}_2 = \boldsymbol{z}]$.

Claim 5 $|\mathbb{E}_{\boldsymbol{e} \sim \xi}[\chi_{\boldsymbol{a}}(\boldsymbol{e})]| \leq \max_{\boldsymbol{c} \in R_q \setminus \{\boldsymbol{0}^n\}} |\mathbb{E}_{\boldsymbol{e} \sim \mathcal{X}^n}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]|$.

Claim 6 *For any $\boldsymbol{c} \in R_q \setminus \{\boldsymbol{0}^n\}$, $|\mathbb{E}_{\boldsymbol{e} \sim \mathcal{X}^n}[\chi_{\boldsymbol{c}}(\boldsymbol{e})]| \leq 1 - \Omega(1/q^2)$.*

The proofs are almost identical to the corresponding proofs of Claim 3, 2 and 1 and so we omit them.