Learning with Errors is easy with quantum samples

Alex B. Grilo*[1] and Iordanis Kerenidis†[1,2]

[1]IRIF, CNRS, Université Paris Diderot, Paris, France
[2]Centre for Quantum Technologies, National University of Singapore, Singapore

February 28, 2017

Abstract

Learning with Errors is one of the fundamental problems in computational learning theory and has in the last years become the cornerstone of post-quantum cryptography. In this work, we study the quantum sample complexity of Learning with Errors and show that there exists an efficient quantum learning algorithm (with polynomial sample and time complexity) for the Learning with Errors problem where the error distribution is the one used in cryptography. While our quantum learning algorithm does not break the LWE-based encryption schemes proposed in the cryptography literature, it does have some interesting implications for cryptography: first, when building an LWE-based scheme, one needs to be careful about the access to the public-key generation algorithm that is given to the adversary; second, our algorithm shows a possible way for attacking LWE-based encryption by using classical samples to approximate the quantum sample state, since then using our quantum learning algorithm would solve LWE.

1 Introduction

The ubiquity and importance of Machine Learning nowadays is undeniable. The large amount of data arising in the real world, for example through scientific observations, large-scale experiments, internet traffic, social media, etc, makes it necessary to be able to predict some general properties or behaviors of the data from a limited number of samples of the data. In this context, Computational Learning Theory provides rigorous models for learning and studies the necessary and sufficient resources, for example, the number of samples or the running time of the learning algorithm. In his seminal work, Valiant [Val84] introduced the model of PAC learning, and since then this model has been extensively studied and has given rise to numerous extensions.

In another revolutionary direction, Quantum Computing takes advantage of the quantum nature of small-scale systems as a computational resource. In this field, the main question is to understand what problems can be solved more efficiently in a quantum computer than in classical computers. In the intersection of the two fields, we have Quantum Learning Theory, where we ask if quantum learning algorithms can be more efficient than classical ones.

One of course needs to be careful about defining quantum learning and more precisely, what kind of access to the data a quantum learning algorithm has. On one hand, we can just provide classical samples to the quantum learning algorithm that can then use the quantum power in processing these classical data. In the more general scenario, we allow the quantum learning algorithm to receive quantum samples of

*`abgrilo@irif.fr`
†`jkeren@irif.fr`

the data, for a natural notion of a quantum sample as a superposition that corresponds to the classical sample distribution.

More precisely, in classical learning, the learning algorithm is provided with samples of $(x, f(x))$, where x is drawn from some unknown distribution D and f is the function we wish to learn. The goal of the learner in this case is to output a function g such that with high probability (with respect to the samples received), f and g are close, i.e., $Pr\left[f(x) \neq g(x)\right]$ is small when x is drawn from the same distribution D.

The extension of this model to the quantum setting is that the samples now are given in the form of a quantum state $\sum_x \sqrt{D(x)}\,|x\rangle\,|f(x)\rangle$. Note that one thing the quantum learner can do with this state is simply measure it in the computational basis and get a classical sample from the distribution D. Hence, a quantum sample is at least as powerful as a classical sample. The main question is whether the quantum learner can make better use of these quantum samples and provide an advantage in the number of samples and/or running time compared to a classical learner.

In this work we focus on one of the fundamental problems in learning theory, the Learning with Errors (LWE). In LWE, one is given samples of the form

$$(x, \langle x, a\rangle + \varepsilon (mod\, q))$$

where $a \in \mathbb{F}_q^n$ is fixed, $x \in \mathbb{F}_q^n$ is drawn uniformly at random and $\varepsilon \in \mathbb{F}_q$ is an 'error' term drawn from some distribution χ. The goal is to output a, while minimizing the number of samples used and the computation time.

First, LWE is the natural generalisation of the well-studied Learning parity with noise problem (LPN), which is the case of $q = 2$. Moreover, a lot of attention was drawn to this problem when Regev [Reg05] reduced some (expected to be) hard problems involving lattices to LWE. With this reduction, LWE has become the cornerstone of current post-quantum cryptographic schemes. Several cryptographic primitives proposals such as Fully Homomorphic Encryption [BV14], Oblivious Transfer [PVW08], Identity based encryption [GPV08, CHKP12, ABB10], and others schemes are based in the hardness of LWE (for a more complete list see Ref. [MR08] and Ref. [Pei16]).

Classically, Blum et al. [BKW03] proposed the first sub-exponential algorithm for this problem, where the sample complexity is $2^{O(n/\log n)}$. Then, Arora and Ge [AG11] improved the time complexity for LWE with a learning algorithm that runs in $2^{\tilde{O}(n^{2\varepsilon})}$ time. For LPN, Lyubashevsky [Lyu05] has proposed an algorithm with sample complexity $n^{1+\varepsilon}$ at the cost of increasing computation time to $O(2^{n/\log\log n})$.

1.1 Our contributions

In this work we study quantum algorithms for solving LWE with quantum samples. Let us be more explicit on the definition of a quantum sample for the LWE problem. We assume that the quantum learning algorithm receives samples in the form

$$\frac{1}{\sqrt{q^n}} \sum_{x \in \mathbb{F}_q^n} |x_1\rangle \ldots |x_n\rangle \left| \sum_{i=1}^{n} a_i x_i + b_x (mod\, q) \right\rangle, \tag{1}$$

where b_x are iid random variables from some distribution χ over $\mathbb{F}_q$.

As expected, the performance of the learning algorithm, both in the classical and quantum case, is sensitive to the noise model adopted, i.e. to the distribution χ. When LWE is used in cryptographic schemes, the distribution χ has support on a small interval around 0, either uniform or a discrete gaussian. We prove that for such distributions, there exists an efficient quantum learner for LWE.

Main Result[informal] *For error distributions χ used in cryptographic schemes, and for any $\epsilon > 0$, there exists a quantum learning algorithm that solves LWE with probability $1 - \eta$ using $O(n \log \frac{1}{\eta})$*

samples and running time $poly(n, \log \frac{1}{\eta})$.

Another interesting feature of our quantum learner is that it is conceptually a very simple algorithm based on one of the basic quantum operations, the Quantum Fourier Transform. Such algorithms have even started to be implemented, of course for very small input sizes and for the binary case [RdSR+15]. Nevertheless, as far as quantum algorithms are concerned, our learner is quite feasible from an implementation point of view.

The approach to solve the problem is a generalisation of Bernstein-Vazirani algorithm [BV97]: we start with a quantum sample, apply a Quantum Fourier Transform over $\mathbb{F}_q$ on each of the qudit registers, and then, we measure in the computational basis. Our analysis shows that, when the last qudit is not 0, which happens with high probability, the value of the remaining registers gives a with constant probability. We can then repeat this process so that our algorithm outputs a with high probability.

We can also use the same technique to prove a generalisation of the result proposed by Cross et al. [CSS15] for the LPN problem. The main difference with their work is that we start with a quantum sample, i.e. a state where the noise is independent for each element in the superposition.

1.2 Related work

We now review some results on quantum algorithms for learning problems. For a more extended introduction, see the survey by Arunachalam and de Wolf [AdW17].

The first approach on trying to solve learning problems with quantum samples was proposed by Bshouty and Jackson [BJ95], where they prove that DNFs can be learned efficiently, even when the samples are noisy. No such efficient learners are known classically.

Despite not presenting it as a learning problem, Bernstein and Vazirani [BV97] show how to learn parity using a single quantum sample, while classically we need a logarithmic number of samples.

Some years later, Servedio and Gortler [SG04] showed that classical and quantum sample/query complexity of learning problems are polynomially related, but they showed that for time complexity there exist exponential separations between classical and quantum learning (assuming some standard computational hardness assumptions).

Then, Ambainis et al. [AIK+04], Atici and Servedio [AS05], and Hunziker et al. [HMP+10] provided general upper bounds on the query complexity for learning problems that depend on the size of the concept class being learned.

On specific problems, Atici and Servedio [AS07] and Belovs [Bel15] provided quantum algorithms for learning juntas and Cross et al. [CSS15] proposed and implemented quantum learning algorithms for LPN in a different noise model.

Recently, Arunachalam and de Wolf [AdW16] proved optimal bounds for the quantum sample complexity of the Quantum PAC model.

1.3 Relation to LWE-based cryptography

As we have mentioned, LWE is used in cryptography for many different tasks. Let us briefly describe how one can build an encryption scheme based on LWE [Reg05]. The key generation algorithm produces a secret key $a \in \mathbb{F}_q$, while the public key consists of a sequence of classical LWE samples $(x^1, \langle a, x^1 \rangle + \varepsilon_1 (mod\, q)), ..., (x^m, \langle a, x^m \rangle + \varepsilon_m (mod\, q))$, where the error comes from a distribution with support in a small interval around 0. For the encryption of a bit b, the party picks a subset S of $[m]$ uniformly at random and outputs

$$\left(\sum_{i \in S} x^i (mod\, q), b \left\lceil \frac{q}{2} \right\rceil + \sum_{i \in S} \langle a, x^i \rangle + \varepsilon_i (mod\, q) \right).$$

For the decryption, knowing a allows one to find b. The security analysis of the encryption scheme postulates that if an adversary can break the encryption efficiently then he is also able to solve the LWE problem efficiently.

The quantum algorithm we present here does not break the above LWE-based encryption scheme. Nevertheless, it does have some interesting implications for cryptography.

First, our algorithm shows a possible way for attacking LWE-based encryption: use classical samples to approximate the quantum sample state, and then use our algorithm to solve LWE. One potential way for this would be to start with m classical samples and create the following superposition

$$\sum_{S\subseteq[m]} |S\rangle \left|\sum_{i\in S} x^i (mod\, q)\right\rangle \left|\sum_{i\in S} \langle a, x^i\rangle + \varepsilon_i (mod\, q)\right\rangle.$$

This operation is in fact efficient. Then, in order to approximate the quantum sample state, one would need to 'forget' the first register that contains the index information about which subset of the m classical samples we took. In the most general case, such an operation of forgetting the index of the states in a quantum superposition, known as index-erasure (see Aharonov and Ta-Shma [ATS03] and Ambainis et al. [AIK$^+$04]), is exponentially hard, and a number of problems, such as Graph Non-isomorphism, would have an efficient quantum algorithm, if we could do it efficiently. Nevertheless, one may try to use the extra structure of the LWE problem to find sub-exponential algorithms for this case.

A second concern that our algorithm raises is that when building an LWE-based scheme, one needs to be careful on the access to the public-key generation algorithm that is given to the adversary. It is well-known that for example, even in the classical case, if the adversary can ask classical queries to the LWE oracle, then he can easily break the scheme: by asking the same query many times one can basically average out the noise and find the secret a. However, if we just assume that the public key is given as a box that an agent has passive access to it, in the sense that he can request a random sample and receive one, then the encryption scheme is secure classically as long as LWE is difficult. However, imagine that the random sample from LWE is provided by a device that creates a superposition $\frac{1}{\sqrt{q^n}} \sum_{x\in\mathbb{F}_q^n} |x_1\rangle ... |x_n\rangle \, |\sum_{i=1}^n a_i x_i + b_x (mod\, q)\rangle$ and then measures it. Then a quantum adversary that has access to this quantum state can break the scheme. Again, our claim is, by no means, that our algorithm breaks the proposed LWE-based encryption schemes, but more that LWE-based schemes which are secure classically (assuming the hardness of LWE) may stop being secure against quantum adversaries if the access to the public key generation algorithm becomes also quantum.

A similar situation has also appeared in the symmetric key cryptography with the so called superposition attacks [Zha12, BZ13, DFNS14, KLLP16]. There, we assume an attacker that has quantum access to the encryption oracle, in other words he can create a superposition of all possible pairs of (message, ciphertext). Such a quantum adversary can in fact break many schemes that are assumed to be secure classically. While in the case of symmetric cryptography, the quantum attacker must have quantum access to the encryption oracle in order to break the system, our results show that in the case of LWE-based public-key encryption, the quantum attacker must have quantum access to the public key generation algorithm.

2 Preliminaries

2.1 Definitions

For $n \in \mathbb{N}$, we define $[n] := \{1, ..., n\}$. For a complex number $x = a + ib$, $a, b \in \mathbb{R}$, we define its norm $|x|$ by $\sqrt{a^2 + b^2}$, its real part $\Re(x) = a$ and its imaginary part $\Im(w) = b$.

We remind now the notation for quantum information and computation. For readers not familiar with these concepts we refer Ref. [NC11]. Let $\{e_i\}$ be the standard basis for the q-dimensional Hilbert space $\mathbb{C}^q$. We denote here $|i\rangle = e_i$ and a q-dimensional qudit is a unit vector in this space, i.e. $|\psi\rangle = \sum_{i\in\mathbb{F}_q} \alpha_i |i\rangle$, for $\alpha_i \in \mathbb{C}$ and $\sum_{i\in\mathbb{F}_q} |\alpha_i|^2 = 1$. We call the state a qubit when $q = 2$. A k-qudit quantum state is a unit vector in the complex Hilbert space $\mathbb{C}^{q^k}$ and we shorthand the basis states for this space $|i_1\rangle \otimes ... \otimes |i_k\rangle$ with $|i_1\rangle ... |i_k\rangle$.

For a state $|\psi\rangle = \sum_{i_1,...,i_k\in\mathbb{F}_q} \alpha_{i_1,...,i_k} |i_1\rangle ... |i_k\rangle$ and a projector P, i.e. P is a linear transformation such that $P = P^2$, the probability a state $|\psi\rangle$ projects onto P is $\|P|\psi\rangle\|^2$.

2.2 Tail bounds

We state now two tail bound that we use in our calculations.

Chernoff bound for sum of Bernoulli. For $i \in [k]$, let X_i be iid Bernoulli random variables

$$X_i = \begin{cases} 1 & \text{w.p. } 1-\delta \\ 0 & \text{w.p. } \delta \end{cases}$$

Let $X = \sum_{i\in[k]} X_i$ and $\mu = \mathbb{E}[X] = (1-\delta)k$. It follows that

$$Pr[X < (1-\varepsilon)\mu] < e^{-\mu\varepsilon^2/2}.$$

Hoeffding's bound. Consider a set of k independent random variables X_i, such that $a_i \le X_i \le b_i$. Let $c_i = b_i - a_i$, $X = \sum_{i\in[k]} X_i$ and $\mu = \mathbb{E}[X]$. Then, it follows that for any $t > 0$

$$Pr[X \le \mu - t] \le e^{\frac{t^2}{\sum c_i^2}}.$$

3 The quantum learning model

3.1 Definition of the model

In this work, we use the model of learning under the uniform distribution where the learner receives samples according to the uniform distribution and outputs the exact function with high probability. In the quantum setting, the learning algorithm is given quantum samples, namely a uniform superposition of the inputs and function values,

$$\sum_{x\in X} \frac{1}{\sqrt{|X|}} |x\rangle |f(x)\rangle.$$

We also have to define how the noise is added into the quantum superposition. We consider the type of noise defined in Bshouty and Jackson [BJ95], where independent noise is added for each element in the superposition. This model is a natural generalisation for quantum samples with noise since it can be seen as a superposition of the classical samples. Also, as noticed before, this is the kind of state we would get after solving the index erasure problem.

3.2 Quantum algorithm for learning a linear function without error

For completeness, we briefly describe the quantum learning algorithm for learning a linear function over $\mathbb{F}_q$ without any noise. This is a simple generalisation of Bernstein-Vazirani algorithm [BV97]. We can use a single quantum sample

$$|\psi\rangle = \frac{1}{\sqrt{q^n}} \sum_{x\in\mathbb{F}_q^n} |x_1\rangle ... |x_n\rangle \left| \sum_{i=1}^{n} a_i x_i (mod\, q) \right\rangle \tag{2}$$

to retrieve the secret a with constant probability.

For this, we use the Quantum Fourier Transform over $\mathbb{F}_q$, which is the following unitary operation

$$QFT\,|j\rangle = \frac{1}{\sqrt{q^n}} \sum_{k=0}^{q^n-1} \omega^{jk}\,|k\rangle$$

and apply it independently on each qudit register of the quantum sample state of eq. (2). We get the state

$$\frac{1}{q^{n+\frac{1}{2}}} \sum_{x_1,\dots,x_n\in\mathbb{F}_q} \sum_{j_1,\dots,j_{n+1}\in\mathbb{F}_q} \omega^{\sum_{i=1}^n x_i(j_i+j_{n+1}a_i)}\,|j_1\rangle\dots|j_n\rangle\,|j_{n+1}\rangle\,.$$

It is not hard to see that the probability that for all $i\in[n]$, we have $j_i = -j_{n+1}a_i(mod\,q)$ is

$$\begin{aligned}
&\left\| \frac{1}{q^{n+\frac{1}{2}}} \sum_{j_{n+1}\in\mathbb{F}_q} \sum_{x\in\mathbb{F}_q^n} \omega^0\,|-j_{n+1}a_1(mod\,q)\rangle\dots|-j_{n+1}a_n(mod\,q)\rangle\,|j_{n+1}\rangle \right\|^2 \\
&= \frac{1}{q^{2n+1}} \sum_{j_{n+1}\in\mathbb{F}_q} \left(\sum_{x\in\mathbb{F}_q^n} 1 \right)^2 \\
&= 1.
\end{aligned}$$

Therefore, whenever $j_{n+1} \neq 0$, which, as we prove later, happens with probability $\frac{q-1}{q}$, we can retrieve a by outputting for all $i\in[n], a_i = \frac{-j_i}{j_{n+1}}$ (all operations mod q).

4 Learning parity with noise

We start by showing how to solve the Learning Parity with Noise (LPN) problem, which is the LWE problem for $q=2$.

Here, as described in last section, the parity bit is flipped independently for each element in the superposition with probability η. This is the same noise model proposed by Bshouty and Jackson [BJ95]. Note that Cross et al. [CSS15] studied LPN with different noise models: in the first, all parities in the superposition were flipped at the same time with probability η; in the second, each qubit passed through a depolarising channel.

The algorithm is the same as in the previous section, where now the QFT is over $\mathbb{F}_2$ (also called the Hadamard Transform). All operations below are $(mod\,2)$.

Theorem 4.1. *Let $|\psi\rangle = \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x_1\rangle\dots|x_n\rangle\,|\sum_{i=1}^n a_ix_i + b_x\rangle$ be a quantum sample where b_x are iid random variables with value 0 with probability $1-\eta$ and 1 with probability η.*

Applying a Hadamard transform on all qubits and measuring them in the computational basis, provides an outcome $|j_1\rangle\dots|j_{n+1}\rangle$ with the following properties: (i) $Pr\,[j_{n+1}=0] = \frac{1}{2}$; (ii) with probability exponentially close to 1 over the sample, $Pr\,[j_1\dots j_n = a|j_{n+1}=1] \geq (1-\delta)^2(1-2\eta)^2$, for any constant δ; and (iii) for $c\neq a$, $\mathbb{E}\,[Pr\,[j_1,\dots,j_n = c|j_{n+1}=1]] \leq \frac{1}{2^n}$.

Proof. If we apply Hadamards on each qubit of the sample state, we have

$$\begin{aligned}
&H^{\otimes n+1}\frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x_1\rangle\dots|x_n\rangle \left| \sum_{i=1}^n a_ix_i + b_x \right\rangle \\
&= \frac{1}{2^{n+\frac{1}{2}}} \sum_{x\in\{0,1\}^n} \sum_{j\in\{0,1\}^{n+1}} (-1)^{b_xj_{n+1}+\sum_{i=1}^n x_i(j_i+j_{n+1}a_i)}\,|j_1\rangle\dots|j_n\rangle\,|j_{n+1}\rangle
\end{aligned}$$

If we measure all the qubits in the computational basis, the probability that the last qubit is $|0\rangle$ is

$$\left\| \frac{1}{2^{n+\frac{1}{2}}} \sum_{j\in\{0,1\}^n} \sum_{x\in\{0,1\}^n} (-1)^{\sum_{i=1}^n x_i j_i} |j_1\rangle \dots |j_n\rangle |0\rangle \right\|^2 = \left\| \frac{2^n}{2^{n+\frac{1}{2}}} |0\rangle \dots |0\rangle |0\rangle \right\|^2 = \frac{1}{2}.$$

This proves item (i).

Otherwise, assuming that the last qubit is $|1\rangle$, we now calculate the probability that the first qubits are state $|a\rangle$ in the normalized post-measured state

$$\left\| \frac{2^{\frac{1}{2}}}{2^{n+\frac{1}{2}}} \sum_{x\in\{0,1\}^n} (-1)^{b_x + \sum_{i=1}^n x_i(a_i+a_i)} |a_1\rangle |a_2\rangle \dots |a_n\rangle \right\|^2 = \frac{1}{2^{2n}} \left(\sum_{x\in\{0,1\}^n} (-1)^{b_x} \right)^2$$

From the distribution of each b_x, we have that $(-1)^{b_x}$ is 1 w.p. $1-\eta$ and -1 w.p. η, independently. Therefore $\mathbb{E}\left[(-1)^{b_x}\right] = 1-2\eta$ and using Hoeffding's bound we have that

$$Pr\left[\sum_{x\in\{0,1\}^n} (-1)^{b_x} \le (1-\delta)(1-2\eta)2^n \right] < e^{\delta^2(1-2\eta)^2 2^{2n}/4}$$

Therefore, with probability exponentially close to 1, the probability that $j = a$ (i.e. $\forall i \in [n], j_i = a_i$) is at least

$$\frac{1}{2^{2n}}((1-\delta)(1-2\eta)2^n)^2 = (1-\delta)^2(1-2\eta)^2$$

This proves item (ii).

Let us fix now $c \in \{0,1\}^n$ such that $c \neq a$ and let $d_i = a_i + c_i$. The expected probability that $j = c$ is

$$\begin{aligned}
&\frac{1}{2^{2n}} \mathbb{E}\left[\left(\sum_{x\in\{0,1\}^n} (-1)^{b_x + \sum_{i\in[n]} d_i x_i} \right)^2 \right] \\
&= \frac{1}{2^{2n}} \left(var\left[\sum_{x\in\{0,1\}^n} (-1)^{b_x + \sum_{i\in[n]} d_i x_i} \right] + \mathbb{E}\left[\sum_{x\in\{0,1\}^n} (-1)^{b_x + \sum_{i\in[n]} d_i x_i} \right]^2 \right) \\
&= \frac{1}{2^{2n}} \left(\sum_{x\in\{0,1\}^n} var\left[(-1)^{b_x + \sum_{i\in[n]} d_i x_i} \right] + \left(\sum_{x\in\{0,1\}^n} \mathbb{E}\left[(-1)^{b_x + \sum_{i\in[n]} d_i x_i} \right] \right)^2 \right),
\end{aligned}$$

where the last equality holds since the random variables b_x are independent.

For i^* such that $a_{i^*} \neq c_{i^*}$, i.e. $d_{i^*} = 1$, let $\tilde{x} = x_1 x_2 \dots x_{i^*-1} x_{i^*+1} \dots x_n$. We have that

$$\begin{aligned}
&\sum_{x\in\{0,1\}^n} \mathbb{E}\left[(-1)^{b_x + \sum_{i\in[n]} d_i x_i} \right] \\
&= \sum_{x\in\{0,1\}^n} (-1)^{\sum_{i\in[n]} d_i x_i} \mathbb{E}\left[(-1)^{b_x} \right] \\
&= (1-2\eta) \sum_{x\in\{0,1\}^n} (-1)^{\sum_{i\in[n]} d_i x_i} \\
&= (1-2\eta) \sum_{\tilde{x}\in\{0,1\}^{n-1}} (-1)^{\sum_{i\in[n]/\{i^*\}} d_i x_i} \sum_{x_{i^*}\in\{0,1\}} (-1)^{x_{i^*}} \\
&= (1-2\eta) \sum_{\tilde{x}\in\{0,1\}^{n-1}} (-1)^{\sum_{i\in[n]/\{i^*\}} d_i x_i} \cdot 0 \\
&= 0.
\end{aligned}$$

For the variance, we have by its definition that

$$
\begin{aligned}
& var\left[(-1)^{b_x+\sum_{i\in[n]} d_i x_i}\right] \\
&= \mathbb{E}\left[\left((-1)^{b_x+\sum_{i\in[n]} d_i x_i}\right)^2\right] - \left((-1)^{\sum_{i\in[n]} d_i x_i}\right)^2 \mathbb{E}\left[(-1)^{b_x}\right]^2 \\
&= 1 - \mathbb{E}\left[(-1)^{b_x}\right]^2 \\
&= 4\eta - 4\eta^2
\end{aligned}
$$

Therefore

$$
\frac{1}{2^{2n}}\mathbb{E}\left[\left(\sum_{x\in\{0,1\}^n}(-1)^{b_x+\sum_{i\in[n]} d_i x_i}\right)^2\right] = \frac{1}{2^{2n}}\sum_{x\in\{0,1\}^n} var\left[(-1)^{b_x+\sum_{i\in[n]} d_i x_i}\right] = \frac{4\eta-4\eta^2}{2^n} < \frac{1}{2^n}
$$

□

We can amplify the probability of outputting a by repeating the process and outputting the most common string. We can prove using Chernoff bounds that the value a will be output with high probability. We do this analysis for the case of LWE.

5 An efficient quantum learning algorithm for LWE

In this section we show how to solve LWE with quantum samples. We use the noise distributions proposed in Brakerski and Vaikuntanathan [BV14]. There, the field order q is sub-exponential in the dimension n (generally in $[2^{n^\gamma}, 2\cdot 2^{n^\gamma})$ for some constant $\gamma \in (0,1)$ while the noise distribution χ produces samples with magnitude at most polynomial in n (for instance linear). For now, we look at the case where we fix the field order to be q and the noise magnitude to be at most $k \ll q$.

Algorithm 1 for LWE

1. Receive a quantum sample $|\psi\rangle = \frac{1}{\sqrt{q^n}}\sum_{x\in\mathbb{F}_q^n}|x_1\rangle \ldots |x_n\rangle \left|\sum_{i=1}^n a_i x_i + b_x (mod\ q)\right\rangle$
2. Apply QFT^{n+1} on $|\psi\rangle$.
3. Measure in the computational basis, resulting in state $|j_1\rangle \ldots |j_{n+1}\rangle$.
4. **If** $j_{n+1} \neq 0$ Output $\left(\frac{-j_1}{j_{n+1}}(mod\, q), \ldots, \frac{-j_n}{j_{n+1}}(mod\, q)\right)$
   Else Output $\perp$

Running time The algorithm performs $O(n)$ QFTs over $\mathbb{F}_q$ and each one takes time $O(\log q)$, hence the overall running time of the algorithm is $O(n\log q)$.

For the correctness of the algorithm we prove the following theorem.

Theorem 5.1. *Let*

$$
|\psi\rangle = \frac{1}{\sqrt{q^n}}\sum_{x\in\mathbb{F}_q^n}|x_1\rangle \ldots |x_n\rangle \left|\sum_{i=1}^n a_i x_i + b_x(mod\ q)\right\rangle,
$$

where the b_x are independent random variables drawn from a noise distribution that is symmetric around 0 *and the noise magnitude is at most $k = polylog(q)$.*

Let o be the output of Algorithm 1. We have the following properties: (i) $Pr[o = \perp] = \frac{1}{q}$; (ii) $Pr[o = a | o \neq \perp] \geq \frac{q}{24(q-1)k}$; and (iii) for $c \neq a$, $\mathbb{E}[Pr[o = c | o \neq \perp]] \leq \frac{1}{q^n}$.

Proof. If we apply QFT on the state $|\psi\rangle$, we have

$$\begin{aligned}&QFT^{\otimes n+1}\frac{1}{\sqrt{q^n}}\sum_{x_1,\dots,x_n\in\mathbb{F}_q}|x_1\rangle\dots|x_n\rangle\left|\sum_{i=1}^{n}a_ix_i+b_x(mod\,q)\right\rangle\\&=\frac{1}{q^{n+\frac{1}{2}}}\sum_{x_1,\dots,x_n\in\mathbb{F}_q}\sum_{j_1,\dots,j_{n+1}\in\mathbb{F}_q}\omega^{b_xj_{n+1}+\sum_{i=1}^{n}x_i(j_i+j_{n+1}a_i)}|j_1\rangle\dots|j_n\rangle|j_{n+1}\rangle\end{aligned}$$

where ω is the q-th root of unity.

If we measure all the qudits in the computational basis, the probability the last qudit is $|0\rangle$ is

$$\left\|\frac{1}{q^{n+\frac{1}{2}}}\sum_{j_1,\dots,j_n\in\mathbb{F}_q}\sum_{x_1,\dots,x_n\in\mathbb{F}_q}\omega^{\sum_{i=1}^{n}x_ij_i}|j_1\rangle\dots|j_n\rangle|0\rangle\right\|^2=\left\|\frac{q^n}{q^{n+\frac{1}{2}}}|0\rangle\dots|0\rangle|0\rangle\right\|^2=\frac{1}{q}.$$

This proves item (i).

Assuming the last qubit is not $|0\rangle$ and renormalizing the state, we now calculate the probability that $o=c$ for any c, i.e., the probability that for all $i\in[n], j_i=-j_{n+1}c_i(mod\,q)$. We denote $d_i=-j_{n+1}c_i+j_{n+1}a_i$ and have

$$\begin{aligned}&\left\|\frac{q^{\frac{1}{2}}}{(q-1)^{\frac{1}{2}}}\frac{1}{q^{n+\frac{1}{2}}}\sum_{j_{n+1}\in\mathbb{F}_q^*}\sum_{x\in\mathbb{F}_q^n}\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}|-j_{n+1}c_1(mod\,q)\rangle\dots|-j_{n+1}c_n(mod\,q)\rangle|j_{n+1}\rangle\right\|^2\\&=\frac{1}{q^{2n}(q-1)}\sum_{j_{n+1}\in\mathbb{F}_q^*}\left(\Re\left(\sum_{x\in\mathbb{F}_q^n}\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)^2+\Im\left(\sum_{x\in\mathbb{F}_q^n}\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)^2\right)\\&=\frac{1}{q^{2n}(q-1)}\sum_{j_{n+1}\in\mathbb{F}_q^*}\left(\left(\sum_{x\in\mathbb{F}_q^n}\Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right)^2+\left(\sum_{x\in\mathbb{F}_q^n}\Im\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right)^2\right)\quad(3)\end{aligned}$$

To prove item (ii), we provide a lower bound for the case $c=a$, in other words, for the case where for all $i\in[n], d_i=0$. Note also that $\Re(\omega^a)=\cos\left(\frac{2\pi}{q}a\right)$. Starting from eq. (3) we have

$$\begin{aligned}&\frac{1}{q^{2n}(q-1)}\sum_{j_{n+1}\in\mathbb{F}_q^*}\left(\left(\sum_{x\in\mathbb{F}_q^n}\Re\left(\omega^{j_{n+1}b_x}\right)\right)^2+\left(\sum_{x\in\mathbb{F}_q^n}\Im\left(\omega^{j_{n+1}b_x}\right)\right)^2\right)\\&\geq\frac{1}{q^{2n}(q-1)}\sum_{j_{n+1}\in\mathbb{F}_q^*}\left(\sum_{x\in\mathbb{F}_q^n}\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right)^2\\&=\frac{1}{q^{2n}(q-1)}\sum_{\substack{j_{n+1}\in\mathbb{F}_q^*\\j_{n+1}\leq\frac{q}{6k}}}\left(\sum_{x\in\mathbb{F}_q^n}\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right)^2+\sum_{\substack{j_{n+1}\in\mathbb{F}_q^*\\j_{n+1}>\frac{q}{6k}}}\left(\sum_{x\in\mathbb{F}_q^n}\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right)^2\\&\geq\frac{1}{q^{2n}(q-1)}\sum_{\substack{j_{n+1}\in\mathbb{F}_q^*\\j_{n+1}\leq\frac{q}{6k}}}\left(\sum_{x\in\mathbb{F}_q^n}\frac{1}{2}\right)^2\\&=\frac{q}{24(q-1)k}.\end{aligned}$$

where we have removed some positive quantities and the last inequality follows from the fact that $|b_x| \leq k$ and for $j_{n+1} \leq \frac{q}{6k}$, we have that $\left|\frac{2\pi}{q} j_{n+1} b_x\right| \leq \frac{\pi}{3}$. This proves item (ii).

We now prove item (iii). For this, we will look at the expected value of the probability that $o = c$ for some $c \neq a$, where the expectation is over the noise random variables b_x and the probability is over the outcome of the quantum measurement.

We will need the following technical lemma whose proof appear in the appendix.

Lemma 5.2. *Let $j_{n+1}, d_1, ..., d_n \in \mathbb{F}_q$ and b_x be independent random variables over $\mathbb{F}_q$ whose distributions are symmetric around 0. It follows that*

$$\mathbb{E}\left[\left(\sum_{x\in\mathbb{F}_q^n} \Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]} d_i x_i}\right)\right)^2 + \left(\sum_{x\in\mathbb{F}_q^n} \Im\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]} d_i x_i}\right)\right)^2\right] = \sum_{x\in\mathbb{F}_q^n}\left(1 - \mathbb{E}\left[\cos\left(\frac{2\pi}{q} j_{n+1} b_x\right)\right]^2\right)$$

Starting from eq. (3) and using Lemma 5.2 we have

$$\begin{aligned}
&\frac{1}{q^{2n}(q-1)} \sum_{j_{n+1}\in\mathbb{F}_q^*} \sum_{x\in\mathbb{F}_q^n} \left(1 - \mathbb{E}\left[\cos\left(\frac{2\pi}{q} j_{n+1} b_x\right)\right]^2\right) \\
&\leq \frac{1}{q^{2n}(q-1)} \sum_{j_{n+1}\in\mathbb{F}_q^*} q^n \\
&= \frac{1}{q^n}.
\end{aligned}$$

$\square$

Corollary 5.3. *For the dimension n, let q be a prime in the interval $[2^{n^\gamma}, 2\cdot 2^{n^\gamma})$. Let*

$$|\psi\rangle = \frac{1}{\sqrt{q^n}} \sum_{x\in\mathbb{F}_q^n} |x_1\rangle \dots |x_n\rangle \left|\sum_{i=1}^n a_i x_i + b_x (mod\, q)\right\rangle,$$

where the b_x are independent random variables drawn from a noise distribution that is symmetric around 0 and the noise magnitude is at most $\frac{n}{24}$.

Let o be the output of Algorithm 1. We have the following properties: (i) $Pr[o = \perp] = \frac{1}{q}$; (ii) $Pr[o = a | o \neq \perp] \geq \frac{1}{n}$; and (iii) for $c \neq a$, $\mathbb{E}[Pr[o = c | o \neq \perp]] \leq \frac{1}{q^n}$. Moreover, Algorithm 1 runs in time $poly(n)$.

5.1 Amplifying the success probability

In this section we show how to amplify the probability of success of the learning and output a with high probability. The idea is very straightforward: repeat the above process and output the most frequent of the retrieved strings.

Algorithm 2 for LWE

1. **For** $i = 1...t$ **run** $o^i \leftarrow$ Algorithm 1
2. Output most frequent among $(o^1, ..., o^t)$

Theorem 5.4. *For the dimension n, let q be a prime in the interval $[2^{n^\gamma}, 2 \cdot 2^{n^\gamma})$. Let*

$$|\psi\rangle = \frac{1}{\sqrt{q^n}} \sum_{x \in \mathbb{F}_q^n} |x_1\rangle \dots |x_n\rangle \left| \sum_{i=1}^{n} a_i x_i + b_x \right\rangle,$$

where the b_x are independent random variables drawn from a noise distribution that is symmetric around 0 and the noise magnitude is at most $\frac{n}{24}$.

Algorithm 2 outputs a with probability $1-\eta$ with sample complexity $t = O(n \log \frac{1}{\eta})$ and running time $poly(n, \log \frac{1}{\eta})$.

Proof. By Corollary 5.3, we know that for each o^i, the probability it is $\perp$ is $\frac{1}{q}$, the probability it is a is at least $\frac{q-1}{qn}$, and the probability it is $c \neq a$ is at most $\frac{q-1}{q^{n+1}}$.

We bound now the probability that the output of Algorithm 2 is different from a.

Let I_i be the indicator variable of the event $o^i = a$. We have that $\mathbb{E}[I_i] \geq \frac{q-1}{qn}$. Therefore, using the Chernoff bound, it follows that

$$P_1 = Pr\left[\left|i : o^i = a\right| \leq (1-\delta)\frac{(q-1)t}{qn}\right] = Pr\left[\sum_{i \in [t]} I_i \leq (1-\delta)\frac{(q-1)t}{qn}\right] \leq e^{-\delta^2 t(q-1)/2qn}$$

We can do a similar analysis and have that for any $c \neq a$ and for $\delta' = \sqrt{(2nq \log q/(t(q-1)) + \delta^2/n)q^n}$

$$Pr\left[\left|i : o^i = c\right| \geq (1+\delta')\frac{(q-1)t}{q^{n+1}}\right] \leq q^{-n} e^{-\delta^2 t(q-1)/2qn}$$

and by union bound,

$$P_2 = Pr\left[\exists c \neq a, \left|i : o^i = c\right| \geq (1+\delta')\frac{(q-1)t}{q^{n+1}}\right] \leq e^{-\delta^2 t(q-1)/2qn}.$$

Again, we can do the analysis for $\perp$ and $\delta'' = \delta\sqrt{(q-1)/n}$

$$P_3 = Pr\left[\left|i : o^i = \perp\right| \geq (1+\delta'')\frac{t}{q}\right] \leq e^{-\delta^2 t(q-1)/2qn}$$

Therefore, from the fact that $(1-\delta)\frac{(q-1)t}{qn}$ is greater than $(1+\delta')\frac{(q-1)t}{q^{n+1}}$ and $(1+\delta'')\frac{t}{q}$, we can use union bound again and have that the probability that the output is not a is at most

$$P_1 + P_2 + P_3 \leq 3e^{-\delta^2 t(q-1)/2qn}$$

Therefore if we want Algorithm 2 to outputs a with probability at least $1-\eta$, we can pick $t = \frac{2nq}{\delta^2(q-1)}\left(\ln 3 + \ln \frac{1}{\eta}\right) = O(n \log \frac{1}{\eta})$. The running time follows from the number of samples and the running time of Algorithm 1. $\square$

Acknowledgments

AG and IK thank Ronald de Wolf for helpful discussions. AG thanks also Lucas Boczkowski, Brieuc Guinard, Alexandre Nolin for helpful discussions. Supported by ERC QCC.

References

[ABB10] Shweta Agrawal, Dan Boneh, and Xavier Boyen. Efficient lattice (H)IBE in the standard model. In *Advances in Cryptology - EUROCRYPT 2010*, pages 553–572, 2010.

[AdW16] Srinivasan Arunachalam and Ronald de Wolf. Optimal quantum sample complexity of learning algorithms. *CoRR*, abs/1607.00932, 2016.

[AdW17] Srinivasan Arunachalam and Ronald de Wolf. A survey of quantum learning theory. *CoRR*, abs/1701.06806, 2017.

[AG11] Sanjeev Arora and Rong Ge. New algorithms for learning in presence of errors. In *Automata, Languages and Programming - 38th International Colloquium, ICALP 2011*, pages 403–415, 2011.

[AIK$^+$04] Andris Ambainis, Kazuo Iwama, Akinori Kawachi, Hiroyuki Masuda, Raymond H. Putra, and Shigeru Yamashita. Quantum identification of boolean oracles. In *21st Annual Symposium on Theoretical Aspects of Computer Science, STACS 2004*, pages 105–116, 2004.

[AS05] Alp Atici and Rocco A. Servedio. Improved bounds on quantum learning algorithms. *Quantum Information Processing*, 4(5):355–386, 2005.

[AS07] Alp Atici and Rocco A. Servedio. Quantum algorithms for learning and testing juntas. *Quantum Information Processing*, 6(5):323–348, 2007.

[ATS03] Dorit Aharonov and Amnon Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 20–29, 2003.

[Bel15] Aleksandrs Belovs. Quantum algorithms for learning symmetric juntas via the adversary bound. *Comput. Complex.*, 24(2):255–293, June 2015.

[BJ95] Nader H. Bshouty and Jeffrey C. Jackson. Learning dnf over the uniform distribution using a quantum example oracle. In *Proceedings of the Eighth Annual Conference on Computational Learning Theory*, COLT '95, pages 118–127, 1995.

[BKW03] Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *J. ACM*, 50(4):506–519, July 2003.

[BV97] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. *SIAM J. Comput.*, 26(5), October 1997.

[BV14] Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. *SIAM Journal on Computing*, 43(2):831–871, 2014.

[BZ13] Dan Boneh and Mark Zhandry. Secure signatures and chosen ciphertext security in a quantum computing world. In *Advances in Cryptology – CRYPTO 2013*, pages 361–379, 2013.

[CHKP12] David Cash, Dennis Hofheinz, Eike Kiltz, and Chris Peikert. Bonsai trees, or how to delegate a lattice basis. *J. Cryptology*, 25(4):601–639, 2012.

[CSS15] Andrew W Cross, Graeme Smith, and John A Smolin. Quantum learning robust against noise. *Physical Review A*, 92(1):012327, 2015.

[DFNS14] Ivan Damgård, Jakob Funder, Jesper Buus Nielsen, and Louis Salvail. Superposition attacks on cryptographic protocols. In *Information Theoretic Security: 7th International Conference, ICITS 2013*, 2014.

[GPV08] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, STOC 2008*, pages 197–206, 2008.

[HMP$^+$10] Markus Hunziker, David A. Meyer, Jihun Park, James Pommersheim, and Mitch Rothstein. The geometry of quantum learning. *Quantum Information Processing*, 9(3):321–341, 2010.

[KLLP16] Marc Kaplan, Gaëtan Leurent, Anthony Leverrier, and María Naya-Plasencia. Breaking symmetric cryptosystems using quantum period finding. In *Advances in Cryptology – CRYPTO 2016*, 2016.

[Lyu05] Vadim Lyubashevsky. The parity problem in the presence of noise, decoding random linear codes, and the subset sum problem. In *Approximation, Randomization and Combinatorial Optimization, Algorithms and Techniques APPROX-RANDOM 2005*, pages 378–389, 2005.

[MR08] Daniele Micciancio and Oded Regev. Lattice-based cryptography. In *Post Quantum Cryptography*, 2008.

[NC11] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[Pei16] Chris Peikert. A decade of lattice cryptography. *Foundations and Trends in Theoretical Computer Science*, 10(4):283–424, 2016.

[PVW08] Chris Peikert, Vinod Vaikuntanathan, and Brent Waters. A framework for efficient and composable oblivious transfer. In *Advances in Cryptology - CRYPTO 2008*, pages 554–571, 2008.

[RdSR$^+$15] Diego Rist, Marcus P. da Silva, Colm A. Ryan, Andrew W. Cross, John A. Smolin, Jay M. Gambetta, Jerry M. Chow, and Blake R. Johnson. Demonstration of quantum advantage in machine learning. *CoRR*, abs/1512.0606G9, 2015.

[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, STOC 2005*, pages 84–93, 2005.

[SG04] Rocco A. Servedio and Steven J. Gortler. Equivalences and separations between quantum and classical learnability. *SIAM J. Comput.*, 33(5):1067–1092, 2004.

[Val84] L. G. Valiant. A theory of the learnable. *Commun. ACM*, 27(11):1134–1142, 1984.

[Zha12] Mark Zhandry. How to construct quantum random functions. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012*, pages 679–687, 2012.

A Proof of Lemma 5.2

Proof. We start by calculating the expectation for the square of the real part, decomposing it into the sum of the variance and the square of the expectation

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{x\in\mathbb{F}_q^n}\Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right)^2\right]\\
&= var\left[\sum_{x\in\mathbb{F}_q^n}\Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right]+\mathbb{E}\left[\sum_{x\in\mathbb{F}_q^n}\Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right]^2\\
&= \sum_{x\in\mathbb{F}_q^n} var\left[\Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right]+\left(\sum_{x\in\mathbb{F}_q^n}\mathbb{E}\left[\Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right]\right)^2,
\end{aligned}
$$

where the last equality holds since the random variables b_x are independent.

Since $\Re\left(\omega^a\right)=\cos\left(\frac{2\pi}{q}a\right)$ we have that for a fixed x

$$
\begin{aligned}
&\mathbb{E}\left[\Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right]\\
&=\mathbb{E}\left[\cos\left(\frac{2\pi}{q}\left(j_{n+1}b_x+\sum_{i\in[n]}d_ix_i\right)\right)\right]\\
&=\mathbb{E}\left[\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\cos\left(\frac{2\pi}{q}\sum_{i\in[n]}d_ix_i\right)-\sin\left(\frac{2\pi}{q}j_{n+1}b_x\right)\sin\left(\frac{2\pi}{q}\sum_{i\in[n]}d_ix_i\right)\right]\\
&=\cos\left(\frac{2\pi}{q}\sum_{i\in[n]}d_ix_i\right)\mathbb{E}\left[\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right]-\sin\left(\frac{2\pi}{q}\sum_{i\in[n]}d_ix_i\right)\mathbb{E}\left[\sin\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right]\\
&=\cos\left(\frac{2\pi}{q}\sum_{i\in[n]}d_ix_i\right)\mathbb{E}\left[\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right]
\end{aligned}
$$

where the last equality holds since b_x is a random variable that is symmetric around 0. We will now show that the above is equal to 0 when summing over all $x\in\mathbb{F}_q^n$.

Let $\tilde{x}=x_1x_2...x_{i^*-1}x_{i^*+1}...x_n$. We will now split the sum for $\tilde{x}$ and x_{i^*}:

$$
\begin{aligned}
&\sum_{x\in\mathbb{F}_q^n}\mathbb{E}\left[\Re\left(\omega^{j_{n+1}b_x+\sum_{i\in[n]}d_ix_i}\right)\right]\\
&=\mathbb{E}\left[\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right]\sum_{x\in\mathbb{F}_q^n}\cos\left(\frac{2\pi}{q}\sum_{i\in[n]}d_ix_i\right)\\
&=\mathbb{E}\left[\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right]\sum_{\tilde{x}\in\mathbb{F}_q^{n-1}}\sum_{x_{i^*}\in\mathbb{F}_q}\cos\left(\frac{2\pi}{q}x_{i^*}+\frac{2\pi}{q}\sum_{i\in[n]/\{i^*\}}d_ix_i\right)\\
&=\mathbb{E}\left[\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right]\sum_{x\in\mathbb{F}_q^n}\sum_{\tilde{y}\in\mathbb{F}_q^{n-1}}\cos\left(\frac{2\pi}{q}\sum_{i\in[n]/\{i^*\}}d_xx_i\right)\sum_{x_{i^*}\in\mathbb{F}_q}\cos\left(\frac{2\pi}{q}x_{i^*}\right)\\
&-\mathbb{E}\left[\cos\left(\frac{2\pi}{q}j_{n+1}b_x\right)\right]\sum_{x\in\mathbb{F}_q^n}\sum_{\tilde{y}\in\mathbb{F}_q^{n-1}}\sin\left(\frac{2\pi}{q}\sum_{i\in[n]/\{i^*\}}d_ix_i\right)\sum_{x_{i^*}\in\mathbb{F}_q}\sin\left(\frac{2\pi}{q}x_{i^*}\right)
\end{aligned}
$$

and since $\sum_{x_{i^*} \in \mathbb{F}_q} \cos\left(\frac{2\pi}{q} x_{i^*}\right) = \sum_{x_{i^*} \in \mathbb{F}_q} \sin\left(\frac{2\pi}{q} x_{i^*}\right) = 0$, we have that

$$\sum_{x \in \mathbb{F}_q^n} \mathbb{E}\left[\Re\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right] = 0$$

For the variance, we have by its definition that

$$\begin{aligned}
&var\left[\Re\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right] \\
&= \mathbb{E}\left[\cos^2\left(\frac{2\pi}{q}\left(j_{n+1} b_x + \sum_{i\in[n]} d_i x_i\right)\right)\right] - \cos^2\left(\frac{2\pi}{q}\sum_{i\in[n]} d_i x_i\right)\mathbb{E}\left[\cos\left(\frac{2\pi}{q} j_{n+1} b_x\right)\right]^2
\end{aligned}$$

As in the real case, for the imaginary part we can decompose it in

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{x \in \mathbb{F}_q^n} \Im\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right)^2\right] \\
&= \sum_{x \in \mathbb{F}_q^n} var\left[\Im\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right] + \left(\sum_{x \in \mathbb{F}_q^n} \mathbb{E}\left[\Im\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right]\right)^2,
\end{aligned}$$

and achieve similar bounds

$$\sum_{x \in \mathbb{F}_q^n} \mathbb{E}\left[\Im\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right] = 0$$

and

$$\begin{aligned}
&var\left[\Im\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right] \\
&= \mathbb{E}\left[\sin^2\left(\frac{2\pi}{q}\left(j_{n+1} b_x + \sum_{i\in[n]} d_i x_i\right)\right)\right] - \sin^2\left(\frac{2\pi}{q}\sum_{i\in[n]} d_i x_i\right)\mathbb{E}\left[\cos\left(\frac{2\pi}{q} j_{n+1} b_x\right)\right]^2.
\end{aligned}$$

We can then simplify the expressions for a fixed x

$$var\left[\Re\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right] + var\left[\Im\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right] = 1 - \mathbb{E}\left[\cos\left(\frac{2\pi}{q} j_{n+1} b_x\right)\right]^2.$$

Therefore, we have that

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{x \in \mathbb{F}_q^n} \Re\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right)^2 + \left(\sum_{x \in \mathbb{F}_q^n} \Im\left(\omega^{j_{n+1} b_x + \sum_{i\in[n]} d_i x_i}\right)\right)^2\right] \\
&= \sum_{x \in \mathbb{F}_q^n}\left(1 - \mathbb{E}\left[\cos\left(\frac{2\pi}{q} j_{n+1} b_x\right)\right]^2\right)
\end{aligned}$$

$\square$