Baltic J. Modern Computing, Vol. 4 (2016), No. 4, pp. 860–875
DOI: http://dx.doi.org/10.22364/bjmc.2016.4.4.17

Quantum Fingerprinting and Quantum Hashing. Computational and Cryptographical Aspects

Farid ABLAYEV, Marat ABLAYEV, Alexander VASILIEV, Mansur ZIATDINOV

Kazan Federal University, 35 Kremlyovskaya str., Kazan 420008, Russia

fablayev@gmail.com, mablayev@gmail.com, alexander.ksu@gmail.com, gltronred@gmail.com

Abstract. Rusins Freivalds was one of the first researchers who introduced methods (later called fingerprinting) for constructing efficient classical randomized and quantum algorithms.
Fingerprinting and cryptographic hashing have quite different usages in computer science, but have similar properties. Interpretation of their properties is determined by the area of their usage: fingerprinting methods are methods for constructing efficient randomized and quantum algorithms for computational problems, while hashing methods are one of the central cryptographical primitives.
Fingerprinting and hashing methods are being developed from the mid of the previous century, while quantum fingerprinting and quantum hashing have a short history.
In the paper we present computational aspects of quantum fingerprinting, discuss cryptographical properties of quantum hashing, and present the possible use of quantum hashing for quantum hash-based message authentication codes.

Keywords: quantum computations, quantum cryptography, fingerprinting, hashing

1 Introduction

Fingerprinting and hashing are well-known techniques. Fingerprinting is widely used in various meanings in different areas of computer science. We restrict ourselves to the area of complexity theory where the notion of fingerprinting is more or less formalized. Cryptographic hashing allow to securely present objects and mathematically is more formalized.

Classical and quantum fingerprinting. Fingerprinting in complexity theory is a procedure that maps a large data item to a much shorter string, its fingerprint, that identifies the original data (with high probability). The key properties of classical fingerprinting methods are: i) they allow to build efficient randomized computational algorithms and ii) the resulting algorithms have bounded error (Motwani and Raghavan, 1995).

Rusins Freivalds was one of the first researchers who introduced methods (later called fingerprinting) for constructing efficient randomized algorithms (which are more efficient than any deterministic algorithm) (Freivalds, 1977, 1979).

In quantum case fingerprinting is a procedure that maps classical data to a quantum state that identifies the original data (with high probability). One of the first applications of the quantum fingerprinting method is due to Ambainis and Freivalds (1998): for a specific language they have constructed a quantum finite automaton with an exponentially smaller size than any classical randomized automaton. An explicit definition of the quantum fingerprinting was introduced by Buhrman et al. in (2001) for constructing efficient quantum communication protocol for equality testing.

Cryptographic quantum hashing. Cryptographic hashing has a lot of fruitful applications in cryptography. Note that in cryptography functions satisfying (i) one-way property and (ii) collision resistance property (in different specific meanings) are called hash functions and we propose to do so when we are considering cryptographical aspects of quantum functions with the above properties. So we suggest to call a quantum function that satisfies properties (i) and (ii) (in the quantum setting) a cryptographic quantum hash function or just quantum hash function. Note however, that there is only a thin line between the notions of quantum fingerprinting and quantum hashing. One of the first considerations of a quantum function (that maps classical words into quantum states) as a cryptographic primitive, having one-way property and collision resistance property is due to (Gottesman and Chuang, 2001), where the quantum fingerprinting function from (Buhrman et al., 2001) was used. Another approach to constructing quantum hash functions from quantum walks was considered in (Li et al., 2013a; Li et al., 2013b; Yang et al., 2016), and it resulted in privacy amplification in quantum key distribution and other useful applications.

The paper organization. In Section 3 we consider quantum fingerprinting as a mapping of classical inputs to quantum states that allows to construct efficient quantum algorithms for computing Boolean functions. We consider the quantum fingerprinting function from (Buhrman et al., 2001) as well as the quantum fingerprinting technique from (Ablayev and Vasiliev, 2009). The latter was motivated by the paper (Ambainis and Freivalds, 1998) and its generalization (Ambainis and Nahimovs, 2008).

Section 4 is based on results on quantum hashing developed in our research group. We define a notion of quantum hash function which is quantum one-way function and quantumly collision resistant function. We show that one-way property and collision resistance property are correlated for a quantum hash function. The more the function is one-way the less it is collision resistant and vice versa. We show that such a correlation can be balanced.

We present an approach for quantum hash function constructions by establishing a connection with small biased sets (Naor and Naor, 1990) and quantum hash function constructions: we prove that small sized ϵ-biased sets allow to generate balanced quantum hash functions. Such a connection adds to the long list of small-biased sets' applications.

In particular it was observed in (Naor and Naor, 1990; Ben-Sasson et al., 2003) that the ϵ-bias property is closely related to the error-correcting properties of linear codes.

Note that the quantum fingerprinting function from (Buhrman et al., 2001) is based on a binary error-correcting code and so it solves the problem of constructing quantum hash functions for the binary case. For the general case ϵ-bias does not correspond to Hamming distance. Thus, in contrast to the binary case, an arbitrary linear error correcting code cannot be used directly for quantum hash functions.

Next, recall that any ϵ-biased set gives rise to a Cayley expander graph (Alon and Roichman, 1994). We show how such graphs generate balanced quantum hash functions. Every expander graph can be converted to a bipartite expander graph. The generalization of these bipartite expander graphs is the notion of extractor graphs. Such point of view gives a method for constructing quantum hash functions based on extractors. This construction of quantum hash functions is applied to define the notion of keyed quantum hash functions. The latter is used for constructing quantum hash-based message authentication codes (QMAC). The security proof of QMAC is based on using strong extractors against quantum storage developed by Ta-Shma (2009) .

2 Preliminaries

Recall that mathematically a qubit is described as a unit vector in the two-dimensional Hilbert complex space $\mathcal{H}^2$. Let $s \geq 1$. Let $\mathcal{H}^d$ be the $d = 2^s$-dimensional Hilbert space, describing the states of s qubits. Another notation for $\mathcal{H}^d$ is $(\mathcal{H}^2)^{\otimes s}$, i.e. $\mathcal{H}^d$ is made up of s copies of a single qubit space $\mathcal{H}^2$

$$(\mathcal{H}^2)^{\otimes s} = \mathcal{H}^2 \otimes \cdots \otimes \mathcal{H}^2 = \mathcal{H}^{2^s}.$$

Conventionally, we use notation $|j\rangle$ for the vector from $\mathcal{H}^d$, which has a 1 on the j-th position and 0 elsewhere. An orthonormal basis $|1\rangle$,…,$|d\rangle$ is usually referred to as the *standard computational basis*. For an integer $j \in \{0, \ldots, 2^s - 1\}$ let $\sigma_1 \ldots \sigma_s$ be a binary presentation of j. We use notation $|j\rangle$ to denote quantum state $|\sigma_1\rangle \cdots |\sigma_s\rangle = |\sigma_1\rangle \otimes \cdots \otimes |\sigma_s\rangle$.

We let $\mathbb{Z}_q$ to be finite additive group of Z/qZ, the integers modulo q. Let Σ^k be a set of words of length k over a finite alphabet Σ. Let $\mathbb{X}$ be a finite set. In the paper we let $\mathbb{X} = \Sigma^k$, or $\mathbb{X} = \mathbb{Z}_q$. For $K = |\mathbb{X}|$ and integer $s \geq 1$ we define a $(K; s)$ classical-quantum function (or just quantum function) to be a unitary transformation (determined by an element $w \in \mathbb{X}$) of the initial state $|\psi_0\rangle \in (\mathcal{H}^2)^{\otimes s}$ to a quantum state $|\psi(w)\rangle \in (\mathcal{H}^2)^{\otimes s}$

$$\psi : \{|\psi_0\rangle\} \times \mathbb{X} \to (\mathcal{H}^2)^{\otimes s} \qquad |\psi(w)\rangle = U(w)|\psi_0\rangle,$$

where $U(w)$ is a unitary matrix. We let $|\psi_0\rangle = |0\rangle$ in the paper and use (for short) the following notation (instead of the one above)

$$\psi : \mathbb{X} \to (\mathcal{H}^2)^{\otimes s} \quad \text{or} \quad \psi : w \mapsto |\psi(w)\rangle.$$

3 Quantum fingerprinting

The ideas of the fingerprinting technique in the quantum setting for the first time appeared in (Ambainis and Freivalds, 1998). The authors used a succinct presentation of the classical input by a quantum automata state, which resulted in an exponential improvement over classical algorithm. Later in (Ambainis and Nahimovs, 2008) the ideas were developed further to give an arbitrarily small probability of error. This was the basis for the general quantum fingerprinting framework proposed in (Ablayev and Vasiliev, 2009).

However, the term "quantum fingerprinting" is mostly used in scientific literature to address a seminal paper by Buhrman et al. (2001), where this notion first appeared explicitly. To distinguish between different versions of the quantum fingerprinting techniques, here we call the fingerprinting function from (Buhrman et al., 2001) "binary" (since it uses some binary error-correcting code in its construction), while the fingerprinting from (Ablayev and Vasiliev, 2009) is called "q-ary" for it uses presentation of the input in $\mathbb{Z}_q$.

3.1 Binary quantum fingerprinting function

The quantum fingerprinting function was formally defined in (Buhrman et al., 2001), where it was used for quantum equality testing in a quantum communication model. It is based on the notion of a binary error-correcting code.

An (n, k, d) *error-correcting code* is a map $C : \Sigma^k \to \Sigma^n$ such that, for any two distinct words $w, w' \in \Sigma^k$, the Hamming distance $d(C(w), C(w'))$ between code words $C(w)$ and $C(w')$ is at least d. The code is binary if $\Sigma = \{0, 1\}$.

The construction of the quantum fingerprinting function is as follows.

- Let $c > 2$ and $\epsilon < 1$. Let k be a positive integer and $n = ck$. Let $E : \{0,1\}^k \to \{0,1\}^n$ be an (n, k, d) binary error-correcting code with Hamming distance $d \geq (1-\epsilon)n$.
- Define a family of functions $F_E = \{E_1, \ldots, E_n\}$, where $E_i : \{0,1\}^k \to \mathbb{F}_2$ is defined by the rule: $E_i(w)$ is the i-th bit of the codeword $E(w)$.
- Let $s = \log n + 1$. Define the quantum function $\psi_{F_E} : \{0,1\}^k \to (\mathcal{H}^2)^{\otimes s}$, determined by a word w as

$$|\psi_{F_E}(w)\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle|E_i(w)\rangle.$$

Original paper (Buhrman et al., 2001) used this function to construct a quantum communication protocol that tests equality in the simultaneous message passing (SMP) model with no shared resources. This protocol requires $O(\log n)$ qubits to compare n-bit binary strings which is exponentially smaller than any classical deterministic or even randomized protocol in the SMP setting with no shared randomness. The proposed quantum protocol has one-sided error of $1/2(1+\langle\psi_{F_E}(x)\,|\,\psi_{F_E}(y)\rangle^2)$, where $|\psi_{F_E}(x)\rangle$, $|\psi_{F_E}(y)\rangle$ are two different quantum fingerprints. Their inner product $|\langle\psi_{F_E}(x)\,|\,\psi_{F_E}(y)\rangle|$ is bounded by ϵ, if the Hamming distance of the underlying code is $(1-\epsilon)n$. For instance, Justesen codes mentioned in the paper give $\epsilon < 9/10 + 1/(15c)$ for any chosen $c > 2$.

In the same paper it was shown, that this result can be improved by choosing an error-correcting code with Hamming distance between any two distinct codewords between $(1-\epsilon)n/2$ and $(1+\epsilon)n/2$ for any $\epsilon > 0$ (however, the existence of such codes can only be proved nonconstructively via probabilistic argument).

But even with such code the quantum fingerprinting function above would give

$$|\langle\psi_{F_E}(x)\,|\,\psi_{F_E}(y)\rangle| < (1+\epsilon)/2,$$

which resulted in the following change of construction (Buhrman et al., 2001).

Define the classical-quantum function $\psi : \{0,1\}^k \to (\mathcal{H}^2)^{\otimes s}$, determined by a word w as

$$\psi(w) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{E_i(w)} |i\rangle.$$

This function gives the following bound for the fingerprints of distinct inputs

$$|\langle\psi_{F_E}(x)\,|\,\psi_{F_E}(y)\rangle| < \epsilon.$$

The further research on this topic mostly used this version of quantum fingerprinting.

3.2 q-ary quantum fingerprinting

In this section we show the basic idea of the quantum fingerprinting from (Ablayev and Vasiliev, 2009; Ablayev and Vasiliev, 2011b).

Let $\sigma = \sigma_1 \dots \sigma_n$ be an input string and g be the mapping of $\{0,1\}^n$ onto $\mathbb{Z}_q$ that "encodes" some property of the input we're about to test. We consider g to be the polynomial over $\mathbb{Z}_q$ such that $g(\sigma) = 0 \bmod q$ iff σ has the property encoded by g. For example, if we test the equality of two n-bit binary strings $x_1 \dots x_n$ and $y_1 \dots y_n$, we can choose g equal to the following polynomial over $\mathbb{Z}_{2^n}$:

$$\sum_{i=1}^{n} x_i 2^{i-1} - \sum_{i=1}^{n} y_i 2^{i-1}.$$

To test the property encoded by g we rotate the initial state $|0\rangle$ of a single qubit by an angle $\theta = \pi g(\sigma)/q$:

$$|0\rangle \to \cos\theta|0\rangle + \sin\theta|1\rangle.$$

Then this state is measured and the input σ is accepted iff the result of the measurement is $|0\rangle$.

Obviously, this quantum state is $\pm|0\rangle$ iff $g(\sigma) = 0 \bmod q$. In the worst case this algorithm gives the one-sided error of $\cos^2 \pi(q-1)/q$, which can be arbitrarily close to 1.

The above description can be presented as follows using $\log t + 1 = (\log\log q) + 1$ qubits:

$$\underbrace{|0\rangle \otimes \cdots \otimes |0\rangle}_{\log t} \otimes |0\rangle \longrightarrow \frac{1}{\sqrt{t}} \sum_{i=1}^{t} |i\rangle \Big(\cos\theta_i|0\rangle + \sin\theta_i|1\rangle \Big),$$

where $\theta_i = \frac{2\pi s_i g(\sigma)}{q}$ and the set $S = \{s_1, \ldots, s_t\} \subseteq \mathbb{Z}_q$ is chosen in order to guarantee the small probability of error (Ablayev and Vasiliev, 2009; Ablayev and Vasiliev, 2011b). That is, the last qubit is simultaneously rotated in t different subspaces by corresponding angles.

Summarizing, quantum fingerprinting method may be applied in the following manner:

1. The initial state of the quantum register is $|0\rangle^{\otimes \log t}|0\rangle$.
2. The Hadamard transform creates the equal superposition of the basis states $\frac{1}{\sqrt{t}} \sum\limits_{j=1}^{t} |j\rangle|0\rangle$
3. Based on the input σ it's fingerprint is created: $\frac{1}{\sqrt{t}} \sum\limits_{j=1}^{t} |j\rangle \left(\cos\frac{2\pi s_j g(\sigma)}{q}|0\rangle + \sin\frac{2\pi s_j g(\sigma)}{q}|1\rangle\right)$
4. The Hadamard transform turns the fingerprint into the superposition $\left(\frac{1}{t} \sum\limits_{l=1}^{t} \cos\frac{2\pi s_l g(\sigma)}{q}\right) |0\rangle^{\otimes \log t}|0\rangle + \ldots$
5. The quantum register is measured and the input is accepted iff the result is $|0\rangle^{\otimes \log t}|0\rangle$.

In (Ablayev and Vasiliev, 2009, 2011a, 2011b) we have applied this technique to construct efficient quantum algorithms for a certain class of Boolean functions in the model of read-once quantum branching programs (Ablayev et al., 2001).

4 Quantum hashing

In this section we present recent results on quantum hashing developed in our research group.

4.1 One-way δ-resistance.

We present the following definition of a quantum δ-resistant one-way function. Let "information extracting" mechanism $\mathbf{M}$ be a function $\mathbf{M} : (\mathcal{H}^2)^{\otimes s} \to \mathbb{X}$. Informally speaking, mechanism $\mathbf{M}$ makes some measurement to state $|\psi\rangle \in (\mathcal{H}^2)^{\otimes s}$ and decodes the result of measurement to $\mathbb{X}$.

Definition 1. Let X be a random variable distributed over $\mathbb{X}$ $\{Pr[X = w] : w \in \mathbb{X}\}$. Let $\psi : \mathbb{X} \to (\mathcal{H}^2)^{\otimes s}$ be a quantum function. Let Y be any random variable over $\mathbb{X}$ obtained by some mechanism $\mathbf{M}$ making measurement to the encoding ψ of X and decoding the result of the measurement to $\mathbb{X}$. Let $\delta > 0$. We call a quantum function ψ a one-way δ-resistant function if

1. if it is easy to compute, i.e., a quantum state $|\psi(w)\rangle$ for a particular $w \in \mathbb{X}$ can be determined using a polynomial-time algorithm;
2. for any mechanism $\mathbf{M}$, the probability $Pr[Y = X]$ that $\mathbf{M}$ successfully decodes Y is bounded by δ
$$Pr[Y = X] \leq \delta.$$

For the cryptographic purposes it is natural to expect (and we do this in the rest of the paper) that random variable X is uniformly distributed.

A quantum state of $s \geq 1$ qubits can "carry" an infinite amount of information. On the other hand, the fundamental result of quantum informatics known as the Holevo's Theorem (Holevo, 1973) states that a quantum measurement can only give $O(s)$ bits of information about the state. Here we use the result of (Nayak, 1999) motivated by the Holevo's Theorem.

Property 1. Let X be a random variable uniformly distributed over $\{0,1\}^k$. Let $\psi : \{0,1\}^k \to (\mathcal{H}^2)^{\otimes s}$ be a $(2^k; s)$ quantum function. Let Y be a random variable over $\{0,1\}^k$ obtained by some mechanism $\mathbf{M}$ making some measurement of the encoding ψ of X and decoding the result of measurement to $\{0,1\}^k$. Then the probability of correct decoding is given by

$$Pr[Y = X] \leq \frac{2^s}{2^k}.$$

4.2 Collision ϵ-resistance

The following definition was presented in (Ablayev and Ablayev, 2015b).

Definition 2. Let $\epsilon > 0$. We call a quantum function $\psi : \mathbb{X} \to (\mathcal{H}^2)^{\otimes s}$ a collision ϵ-resistant function if for any pair w, w' of different inputs,

$$|\langle \psi(w) \,|\, \psi(w') \rangle| \leq \epsilon.$$

Testing equality. The crucial procedure for quantum hashing is an equality test for $|\psi(v)\rangle$ and $|\psi(w)\rangle$ that can be used to compare encoded classical messages v and w; see for example (Gottesman and Chuang, 2001). This procedure can be a well-known SWAP-test (Buhrman, Cleve, Watrous, and Wolf, 2001) or something that is adapted for specific hashing function, like REVERSE-test (Ablayev and Vasiliev, 2014).

4.3 Balanced quantum (δ, ϵ)-resistance.

The above two definitions and considerations lead to the following formalization of the quantum cryptographic (one-way and collision resistant) function

Definition 3. Let $K = |\mathbb{X}|$ and $s \geq 1$. Let $\delta > 0$ and $\epsilon > 0$. We call a function $\psi : \mathbb{X} \to (\mathcal{H}^2)^{\otimes s}$ a quantum (δ, ϵ)-Resistant $(K; s)$-hash function (or just quantum (δ, ϵ)-hash function) iff ψ is one-way δ-resistant and is collision ϵ-resistant function.

We present below the following two examples to demonstrate how one-way δ-resistance and collision ϵ-resistance are correlated. The first example was presented in (Ambainis and Freivalds, 1998) in terms of quantum automata.

Example 1. Let us encode numbers v from $\{0, \ldots, 2^k - 1\}$ by a single qubit as follows:

$$\psi : v \mapsto \cos\left(\frac{2\pi v}{2^k}\right)|0\rangle + \sin\left(\frac{2\pi v}{2^k}\right)|1\rangle.$$

Extracting information from $|\psi\rangle$ by measuring $|\psi\rangle$ with respect to the basis $\{|0\rangle, |1\rangle\}$ gives the following result. The function ψ is one-way $\frac{2}{2^k}$-resistant (see Property 1) and collision $\cos\left(\pi/2^{k-1}\right)$-resistant. Thus, the function ψ has good one-way property, but has bad collision resistance property for large k.

Example 2. Let $v = \sigma_1 \dots \sigma_k \in \{0,1\}^k$. We encode v by k qubits: $\psi : v \mapsto |v\rangle = |\sigma_1\rangle \cdots |\sigma_k\rangle$.

Extracting information from $|\psi\rangle$ by measuring $|\psi\rangle$ with respect to the basis $\{|0 \dots 0\rangle, \dots, |1 \dots 1\rangle\}$ gives the following result. The function ψ is one-way 1-resistant and collision 0-resistant. So, in contrast to Example 1 the encoding ψ from Example 2 is collision free, that is, for different words v and w quantum states $|\psi(v)\rangle$ and $|\psi(v)\rangle$ are orthogonal and therefore reliably distinguished; but we lose the one-way property: ψ is easily invertible.

The following result (Ablayev and Ablayev, 2015b) shows that a quantum collision ϵ-resistant $(K; s)$ function needs at least $\log\log K - c(\epsilon)$ qubits.

Property 2. Let $s \geq 1$ and $K = |\mathbb{X}| \geq 4$. Let $\psi : \mathbb{X} \to (\mathcal{H}^2)^{\otimes s}$ be a collision ϵ-resistant quantum hash function. Then

$$s \geq \log\log K - \log\log\left(1 + \sqrt{2/(1-\epsilon)}\right) - 1.$$

Proof. See (Ablayev and Ablayev, 2015b) for the proof. □

Properties 1 and 2 provide a basis for building a "balanced" one-way δ-resistance and collision ϵ-resistance properties. That is, roughly speaking, if we need to hash elements w from the domain $\mathbb{X}$ with $|\mathbb{X}| = K$ and if one can build for an $\epsilon > 0$ a collision ϵ-resistant $(K; s)$ hash function ψ with $s \approx \log\log K - c(\epsilon)$ qubits then the function f is one-way δ-resistant with $\delta \approx (\log K/K)$. Such a function is balanced with respect to Property 2.

To summarize the above considerations we can state the following. A quantum (δ, ϵ)-hash function is a function that satisfies all of the properties that a "classical" hash function should satisfy. Pre-image resistance follows from Property 1. Second pre-image resistance and collision resistance follow, because all inputs are mapped to states that are nearly orthogonal. Therefore, we see that quantum hash functions can satisfy the three properties of a classical cryptographic hash function.

4.4 Quantum (δ, ϵ)-Hash Functions Construction Via Small-Biased Sets.

This section is based on the paper (Vasiliev, 2016). We present here a brief background on ϵ-biased sets as defined in (Chen, Moore, and Russell, 2013) and discuss their connection to quantum hashing. Note that ϵ-biased sets are generally defined for arbitrary finite groups, but here we restrict ourselves to $\mathbb{Z}_q$.

For an $a \in \mathbb{Z}_q$ a character χ_a of $\mathbb{Z}_q$ is a homomorphism $\chi_a : \mathbb{Z}_q \to \mu_q$, where μ_q is the (multiplicative) group of complex q-th roots of unity. That is, $\chi_a(x) = \omega^{ax}$, where $\omega = e^{\frac{2\pi i}{q}}$ is a primitive q-th root of unity. The character $\chi_0 \equiv 1$ is called a trivial character.

Definition 4. A set $S \subseteq \mathbb{Z}_q$ is called ϵ-biased, if for any nontrivial character $\chi \in \{\chi_a : a \in \mathbb{Z}_q\}$

$$\frac{1}{|S|}\left|\sum_{x\in S}\chi(x)\right| \leq \epsilon.$$

These sets are interesting when $|S| \ll |\mathbb{Z}_q|$ (as $S = \mathbb{Z}_q$ is 0-biased). In their seminal paper Naor and Naor (1990) defined these small-biased sets, gave the first explicit constructions of such sets, and demonstrated the power of small-biased sets for several applications.

Remark 1. Note that a set S of $O(\log q/\epsilon^2)$ elements selected uniformly at random from $\mathbb{Z}_q$ is ϵ-biased with positive probability (Alon and Roichman, 1994).

Many other constructions of small-biased sets followed during the last decades.

Vasiliev (2016) showed that ϵ-biased sets generate (δ,ϵ)-resistant hash functions. We present the result of (Vasiliev, 2016) in the following form.

Property 3. Let $S \subseteq \mathbb{Z}_q$ be an ϵ-biased set. Let

$$H_S = \{h_a(x) = ax \pmod q, \quad a \in S, h_a : \mathbb{Z}_q \to \mathbb{Z}_q\}$$

be a set of functions determined by S. Then a quantum function $\psi_S : \mathbb{Z}_q \to (\mathcal{H}^2)^{\otimes \log|S|}$

$$|\psi_S(x)\rangle = \frac{1}{\sqrt{|S|}}\sum_{a\in S}\omega^{h_a(x)}|a\rangle$$

is a (δ,ϵ)-resistant quantum hash function, where $\delta \leq |S|/q$.

Proof. One-way δ-resistance property of ψ_S follows from Property 1: a probability of correct decoding an x from a quantum state $|\psi_S(x)\rangle$ is bounded by $|S|/q$. The efficient computability of such a function follows from the fact that any quantum transformation on s qubits (including the one that creates a quantum hash) can be performed with $O(s^2 4^s)$ elementary quantum gates (Nielsen and Chuang, 2000). Whenever $s = O(\log|S|) = O(\log\log q - \log\epsilon)$, this number of steps is polynomial in $\log q$ (the binary representation of group elements) and $1/\epsilon$.

Collision ϵ-resistance property of ψ_S follows directly from the corresponding property of (Vasiliev, 2016). Note that

$$|\psi_S(x)\rangle = \frac{1}{\sqrt{|S|}}\sum_{a\in S}\omega^{h_a(x)}|a\rangle = \frac{1}{\sqrt{|S|}}\sum_{a\in S}\chi_x(a)|a\rangle.$$

Further proof coincides with the proof of the paper (Vasiliev, 2016). □

Remark 2. It is natural to call the set H_S of functions a *uniform ϵ-biased quantum hash generator* in the context of the definition of quantum hash generator from (Ablayev and Ablayev, 2015a) and the above considerations.

As a corollary of the Property 3 and the above considerations we can state the following.

Property 4. For a small sized ϵ-biased set $S = \{s_1, \dots, s_t\} \subset \mathbb{Z}_q$ with $t = O(\log q/\epsilon^2)$, for $\delta \leq O(\frac{\log q}{\epsilon^2 q})$ a quantum hash generator H_S generates balanced (δ, ϵ)-resistant quantum hash function ψ_S

$$|\psi_S(a)\rangle = \frac{1}{\sqrt{t}} \sum_{j=1}^{t} \omega^{a s_j} |j\rangle.$$

4.5 Quantum fingerprinting functions as hash functions

In this section we give two explicit examples of the quantum hashing for specific finite abelian groups, which turn out to be the known quantum fingerprinting schemas.

Hashing the elements of the Boolean cube. For $G = \mathbb{Z}_2^n$ its characters can be written in the form $\chi_a(x) = (-1)^{(a,x)}$, and the corresponding quantum hash function is the following

$$|\psi_S(a)\rangle = \frac{1}{\sqrt{|S|}} \sum_{j=1}^{|S|} (-1)^{(a,s_j)} |j\rangle.$$

The resulting hash function is exactly the quantum fingerprinting by Buhrman et al. (2001), once we consider an error-correcting code, whose matrix is built from the elements of S. Indeed, as stated in (Ben-Aroya and Ta-Shma, 2009) an ε-balanced error-correcting code can be constructed out of an ε-biased set. Thus, the inner product (a, x) in the exponent is equivalent to the corresponding bit of the codeword, and altogether this gives the quantum fingerprinting function, that stores information in the phase of quantum states (de Wolf, 2001).

Hashing the elements of the cyclic group For $G = \mathbb{Z}_q$ its characters can be written as $\chi_a(x) = \exp(2\pi i a x/q)$, and the corresponding quantum hash function is given by

$$|\psi_S(a)\rangle = \frac{1}{\sqrt{|S|}} \sum_{j=1}^{|S|} \omega^{a s_j} |j\rangle.$$

The above quantum hash function is essentially equivalent to the one we have defined earlier in (Ablayev and Vasiliev, 2014), which is in turn based on the quantum fingerprinting function from (Ablayev and Vasiliev, 2009).

4.6 Quantum hash functions via expander graphs

In this section we show further development of the quantum hashing for finite groups. First, we explore the connection of the small-bias sets to the graph theory and then construct corresponding quantum hash functions.

Let us recall some definitions from graph theory.

A graph Γ is a set V of vertices and a (multi-)set of edges E. Graph Γ is the d-regular graph if all vertices have the same degree d; i.e. each vertex is incident to exactly d edges.

Adjacency matrix of the graph $A = A(\Gamma)$ is an $n \times n$ matrix whose (u, v) entry is the number of edges between vertex u and vertex v. We refer to the eigenvalues of $A(\Gamma)$ as the spectrum of the graph Γ.

Given a d-regular graph Γ with n vertices and spectrum $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ we denote $\lambda(\Gamma) = \max\{|\lambda_2|, |\lambda_n|\}$.

We call the graph Γ a (d, λ)-expander graph if Γ is d-regular and has $\lambda(\Gamma) = \lambda$.

Next we consider the special case of expanders called Cayley graphs. They are defined as follows.

The set of vertices is identified with group elements G. The set of edges is $E = \{(g, gs) : s \in S\}$.

If set S is symmetric (i.e. for any element $s \in S$ of set its inverse s^{-1} is also contained in S), graph $\Gamma(G, S)$ is undirected.

The following fact is true for any finite abelian group G and any symmetric set S.

Property 5. Let χ be a character of a group G. The vector $b = \{\chi(g) : g \in G\}$ is a eigenvector of the matrix $A_S = A(\Gamma(G, S))/|S|$ and corresponding eigenvalue is

$$\frac{1}{|S|} \sum_{s \in S} \chi(s).$$

Proof. Let a_{ij} be elements of matrix A_S. Denote elements of G by $g_1, g_2, \ldots$. Then j-th element of $A_S b$ is

$$\sum_k a_{jk} \chi(g_k) = \frac{1}{|S|} \sum_{s \in S} \chi(g_j s) = \left(\frac{1}{|S|} \sum_{s \in S} \chi(s) \right) \chi(g_j).$$

Therefore, $A_S b = \left(\frac{1}{|S|} \sum_{s \in S} \chi(s) \right) b$.

□

The number of irreducible characters of a group G is equal to the number of conjugacy classes of G, therefore for any abelian group G the following property holds.

Property 6. The Cayley graph $\Gamma(G, S)$ is an $(|S|, \epsilon)$-expander graph if and only if for all nontrivial characters χ

$$\frac{1}{|S|} \left| \sum_{s \in S} \chi(s) \right| \leq \epsilon.$$

Here we note, that any ϵ-biased set S gives rise to an $(|S|, \epsilon)$-expander graph which is Cayley graph, and Ziatdinov (2016) showed that (d, ϵ)-expander graphs generate quantum hash functions using the following construction.

Let $\Gamma = (V, E)$ be a (d, ϵ)-expander graph. We label vertices V of graph Γ with elements of group G.

Let us randomly choose one vertex and perform a random walk of length $t > O\left(\frac{\log |G|}{\epsilon}\right)$ starting from it. Denote vertices that occurred in this walk by s_j. Then the following theorem holds.

Theorem 1. *A quantum function* $\Psi_{\Gamma,t} : G \to (\mathcal{H}^2)^{\log t}$ *defined as*

$$|\Psi_{\Gamma,t}(g)\rangle = \frac{1}{\sqrt{t}} \sum_{k=1}^{t} \chi(g \circ s_k)|k\rangle.$$

is a (δ, ϵ)*-resistant quantum hash function with* $\delta < t/|G|$.

For the proof see (Ziatdinov, 2016).

4.7 Quantum hash functions via extractors

Every expander graph can be converted to a bipartite expander graph. Generalization of these bipartite expander graphs is the notion of extractor graphs. The extractor graph is a bipartite graph where size of components can be different. An extractor can also be defined in terms of function that maps pair of the first component vertex and edge to the second component vertex.

To define extractors we first recall the notions of statistical distance and min-entropy.

Definition 5. We say that two distributions F and G are ϵ-close, if for every event A, $|\Pr[F \in A] - \Pr[G \in A]| \leq \epsilon$.

The support of a distribution X is $\mathrm{Supp}(X) = \{x : \Pr[X = x] > 0\}$.

The uniform distribution over $\{0,1\}^m$ is denoted by U_m and we say that X is ϵ-close to uniform if it is ϵ-close to U_m.

We denote that distribution F is ϵ-close to distribution G by $F \overset{\epsilon}{\approx} G$.

Definition 6. Let X be a distribution. The min-entropy of X is

$$H_\infty(X) = \min_{x \in \mathrm{Supp}(X)} \log \frac{1}{\Pr[X = x]}.$$

Now we recall the definition of extractors.

Definition 7. A function $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k, ϵ)-extractor if for every distribution X over $\{0,1\}^n$ with min-entropy $H_\infty(X) \geq k$, $E(X,Y)$ is ϵ-close to uniform (where Y is distributed like U_d and is independent of X).

The notion of extractor can be used to construct a quantum hash function in the following way.

Let $\mathrm{Ext} : G \times \{0,1\}^d \to H$ be a $(k; \epsilon)$ extractor function. Let X be a distribution with min-entropy $H_\infty(X) \geq k$. Let $t > \frac{\log|H|+1}{2\epsilon^2}||\chi||_\infty$ and $s_i \in G, i \in \{1,\ldots,t\}$ chosen according to distribution X be parameters. Denote $S = \{s_i, i \in \{1,\ldots,t\}\}$.

Let $\Psi_{\mathrm{Ext},t,S} : G \to (\mathcal{H}^2)^{\otimes(d+\log t)}$ be a quantum function defined as

$$|\Psi_{\mathrm{Ext},t,S}(g)\rangle = \frac{1}{\sqrt{t2^d}} \sum_{i=1}^{t} \sum_{j=1}^{2^d} \chi(\mathrm{Ext}(g \circ s_i, j))|j\rangle|i\rangle.$$

The following theorem about $\Psi_{\mathrm{Ext},t,S}$ was proved in (Ziatdinov, 2016).

Theorem 2. $\Psi_{\mathrm{Ext},t,S}$ *is a* (δ, ϵ)*-resistant quantum hash function, where* $\delta \leq 2^d t/|G|$.

Thus, using explicit extractors (like the one of (Guruswami, Umans, and Vadhan, 2009)) we can obtain an explicit quantum hash function with cryptographic properties.

4.8 Message authentication codes via quantum hash functions

Classical message authentication codes (MAC) have a wide range of applications, for more details we refer to (Menezes, Van Oorschot, and Vanstone, 1996). They are defined as a triple of algorithms: G that generates a key, S that uses the key and the message to generate a tag of the message, and V that uses the key, the message and the tag to verify message integrity. This method uses shared secret key, and so parties should trust each other.

Formally, $G : 1^n \to K$, where n is a security parameter and K is a set of all possible keys, $S : K \times X \to T$, where X is a set of messages and T is a set of tags and $V : K \times X \times T \to \{\mathrm{Acc}, \mathrm{Rej}\}$.

We require the following property for MAC to be a sound system:

$$\forall n, \forall x \in X : k = G(1^n), V\big(k, x, S(k, x)\big) = \mathrm{Acc},$$

i.e. that verifier always accepts a generated tag.

We also require that MAC is a secure system and for any adversary A that can query MAC:

$$\forall n, k \notin \mathrm{Query}(A), (x, t) \leftarrow A(S), \Pr\big[V(k, x, t) = \mathrm{Acc}\big] \leq 2^{-n},$$

i.e. any adversary that can query MAC outputs correct tag for some key that was not queried and some message with negligible probability.

One classical construction of MAC is hash-based MAC (also known as keyed hash functions). Basically, keyed hash function is a function $H(k, x)$, such that $H(k, \cdot)$ is a cryptographic hash function for every k. It is easy to see that such function can be used as MAC.

Strong extractors against quantum storage. Ta-Shma (2009) introduced the following definitions.

Definition 8. An (n, b) quantum encoding is a collection $\{\rho(x)\}_{x \in \{0,1\}^n}$ of density matrices $\rho(x) \in (\mathcal{H}^2)^{\otimes b}$.

Definition 9. A boolean test T ϵ-distinguishes a distribution D_1 from a distribution D_2 if

$$\left| \Pr_{x_1 \in D_1}[T(x_1) = 1] - \Pr_{x_2 \in D_2}[T(x_2) = 1] \right| \geq \epsilon.$$

We say D_1 is ϵ-indistinguishable from D_2 if no boolean POVM can ϵ-distinguish D_1 from D_2.

Definition 10. A function $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k, b, ϵ) strong extractor against quantum storage, if for any distribution $X \subseteq \{0,1\}^n$ with $H_\infty(X) \geq k$ and every (n, b) quantum encoding $\{\rho(x)\}$, $U_t \circ E(X, U_t) \circ \rho(X)$ is ϵ-indistinguishable from $U_{t+m} \circ \rho(X)$.

Keyed quantum hash functions. Now we are ready to define keyed quantum hash functions.

Definition 11. A (δ, ϵ) keyed quantum hash function is a quantum function S, such that

- A function S, given a key $k \in K$ and a message $x \in X$, outputs a quantum tag for x: $S : K \times X \to T = (\mathcal{H}^2)^{\otimes t}$.
- S is sound, i.e. tags should be different for different messages under the same key.

  $$\forall k \in K, \forall x \in X, \forall y \neq x : |\langle S(k,x) \,|\, S(k,y)\rangle| < \epsilon.$$

  For $x = y$ we get $\langle S(k,x) \,|\, S(k,x)\rangle = 1$.
- S is unforgeable:

  $$\forall k \in K, k \notin \mathrm{Query}(A), (x,t) \leftarrow A(S), \Pr\left[\langle t \,|\, S(k,x)\rangle \geq \epsilon)\right] \leq \delta,$$

  where A is arbitrary attacker that can query S and $\mathrm{Query}(A)$ is a set of queries made.

Informally, keyed quantum hash function outputs a tag for a message. If someone changes a message, then the verification step fails with high probability. If an attacker Eve can query a keyed quantum hash function, access to the function doesn't help her to forge a tag for some message with some (unqueried) key.

Let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a (k, b, ϵ) extractor against quantum storage and $b > r(d + \log t)$.

Let Ψ_{Ext} be the following quantum function.

$$|\Psi_{\mathrm{Ext}}(\mathsf{key}, g)\rangle = \frac{1}{\sqrt{t2^d}} \sum_{i=1}^{t} \sum_{j=1}^{2^d} \chi(\mathrm{Ext}(g \circ \mathsf{key} \circ s_i, j))|j\rangle|i\rangle$$

Theorem 3. Ψ_{Ext} *is an* $(\epsilon + \epsilon^{2^s+1}, \epsilon)$ *keyed quantum hash function secure against an attacker* A *with access to* r *queries to* Ψ_{Ext}.

For the detailed proof we refer to (Ziatdinov, 2016). Here we note, that using explicit extractor against quantum storage from (De and Vidick, 2009) one can construct a corresponding keyed quantum hash function.

Acknowledgements

The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University. The work was in part supported by the Russian Foundation for Basic Research (grants 14-07-00878, 15-37-21160).

References

Ablayev, F., and Ablayev, M. (2015a). On the concept of cryptographic quantum hashing. *Laser Physics Letters*, *12*(12), 125204.

Ablayev, F., and Ablayev, M. (2015b). Quantum hashing via ϵ-universal hashing constructions and classical fingerprinting. *Lobachevskii Journal of Mathematics*, *36*(2), 89–96.

Ablayev, F., Gainutdinova, A., and Karpinski, M. (2001). On computational power of quantum branching programs. In *FCT* (p. 59-70).

Ablayev, F., and Vasiliev, A. (2009). Algorithms for quantum branching programs based on fingerprinting. *Electronic Proceedings in Theoretical Computer Science*, *9*, 1-11.

Ablayev, F., and Vasiliev, A. (2011a). Classical and quantum parallelism in the quantum fingerprinting method. In V. Malyshkin (Ed.), *11th international conference pact 2011 proceedings* (Vol. 6873, p. 1-13). Springer.

Ablayev, F., and Vasiliev, A. (2011b). On computational power of quantum read-once branching programs. *Electronic Proceedings in Theoretical Computer Science*, *52*, 1-12.

Ablayev, F. M., and Vasiliev, A. V. (2014). Cryptographic quantum hashing. *Laser Physics Letters*, *11*(2), 025202.

Alon, N., and Roichman, Y. (1994). Random cayley graphs and expanders. *Random Structures & Algorithms*, *5*(2), 271–284.

Ambainis, A., and Freivalds, R. (1998). 1-way quantum finite automata: strengths, weaknesses and generalizations. In *Proceeding of the 39th IEEE conference on foundation of computer science* (p. 332-342). Washington, DC, USA: IEEE Computer Society.

Ambainis, A., and Nahimovs, N. (2008). Improved constructions of quantum automata. In Y. Kawano and M. Mosca (Eds.), *Theory of quantum computation, communication, and cryptography* (Vol. 5106, p. 47-56). Springer Berlin / Heidelberg.

Ben-Aroya, A., and Ta-Shma, A. (2009, Oct). Constructing small-bias sets from algebraic-geometric codes. In *Foundations of computer science, 2009. FOCS '09. 50th annual IEEE symposium on* (p. 191-197).

Ben-Sasson, E., Sudan, M., Vadhan, S., and Wigderson, A. (2003). Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *Proceedings of the thirty-fifth annual acm symposium on theory of computing* (pp. 612–621). New York, NY, USA: ACM.

Buhrman, H., Cleve, R., Watrous, J., and Wolf, R. de. (2001, Sep). Quantum fingerprinting. *Phys. Rev. Lett.*, *87*(16), 167902.

Chen, S., Moore, C., and Russell, A. (2013). Small-bias sets for nonabelian groups. In P. Raghavendra, S. Raskhodnikova, K. Jansen, and J. D. Rolim (Eds.), *Approximation, randomization, and combinatorial optimization. algorithms and techniques* (Vol. 8096, p. 436-451). Springer Berlin Heidelberg.

De, A., and Vidick, T. (2009). Near-optimal extractors against quantum storage. Retrieved from http://arxiv.org/abs/0911.4680

Freivalds, R. (1977). Probabilistic machines can use less running time. In *IFIP congress* (Vol. 839, p. 842).

Freivalds, R. (1979). Fast probabilistic algorithms. In J. Becvar (Ed.), *Mathematical foundations of computer science 1979* (Vol. 74, p. 57-69). Springer Berlin / Heidelberg.

Gottesman, D., and Chuang, I. (2001, Nov). *Quantum digital signatures* (Tech. Rep. No. arXiv:quant-ph/0105032). Cornell University Library. Retrieved from http://arxiv.org/abs/quant-ph/0105032

Guruswami, V., Umans, C., and Vadhan, S. (2009). Unbalanced expanders and randomness extractors from parvaresh-vardy codes. *Journal of the ACM*, *56*(4), 1–34.

Holevo, A. S. (1973). Some estimates of the information transmitted by quantum communication channel (Russian). *Probl. Pered. Inform. [Probl. Inf. Transm.]*, *9*(3), 3-11.

Li, D., Zhang, J., Guo, F.-Z., Huang, W., Wen, Q.-Y., and Chen, H. (2013). Discrete-time interacting quantum walks and quantum hash schemes. *Quantum Information Processing*, *12*(3), 1501–1513.

Li, D., Zhang, J., Ma, X.-W., Zhang, W.-W., and Wen, Q.-Y. (2013). Analysis of the two-particle controlled interacting quantum walks. *Quantum Information Processing*, *12*(6), 2167–2176.

Menezes, A. J., Van Oorschot, P. C., and Vanstone, S. A. (1996). *Handbook of applied cryptography*. CRC press.

Motwani, R., and Raghavan, P. (1995). *Randomized algorithms*. Cambridge University Press.

Naor, J., and Naor, M. (1990). Small-bias probability spaces: Efficient constructions and applications. In *Proceedings of the twenty-second annual ACM symposium on theory of computing* (pp. 213–223). New York, NY, USA: ACM.

Nayak, A. (1999). Optimal lower bounds for quantum automata and random access codes. In *Foundations of computer science, 1999. 40th annual symposium on* (p. 369-376).

Nielsen, M. A., and Chuang, I. L. (2000). *Quantum computation and quantum information* (1 ed.). Cambridge University Press.

Ta-Shma, A. (2009). Short Seed Extractors Against Quantum Storage. *Proc. ACM STOC*, 401–408.

Vasiliev, A. (2016). Quantum hashing for finite abelian groups. *Lobachevskii Journal of Mathematics*, *37*(6), 751-754.

Wolf, R. de. (2001). *Quantum computing and communication complexity*. PhD Theses, University of Amsterdam.

Yang, Y.-G., Xu, P., Yang, R., Zhou, Y.-H., and Shi, W.-M. (2016). Quantum hash function and its application to privacy amplification in quantum key distribution, pseudo-random number generation and image encryption. *Scientific Reports*, *6*, 19788.

Ziatdinov, M. (2016). From graphs to keyed quantum hash functions. *Lobachevskii Journal of Mathematics*, *37*(6), 704-711.

Received September 22, 2016 , accepted November 1, 2016