Quantum Hashing

Farid Ablayev∗ Alexander Vasiliev†

Abstract

We present a version of quantum hash function based on non-binary discrete functions. The proposed quantum procedure is “classical-quantum”, that is, it takes a classical bit string as an input and produces a quantum state. The resulting function has the property of a one-way function (pre-image resistance), in addition it has the properties analogous to classical cryptographic hash second pre-image resistance and collision resistance.

This function can be naturally used in a quantum digital signature protocol.

1 Preliminaries and introduction

Hashing has a lot of fruitful applications in computer science, in particular the public-key cryptography relies on cryptographic hash functions. Hash functions are designed to take a string of large length (theoretically any length) as an input and produce a short (in practice a fixed-length) hash value. A cryptographic hash function must be additionally able to withstand all known types of cryptanalytic attack. At least, it must have the following properties:

- **Pre-image resistance** (or equivalently First pre-image resistance).
  
  Given a hash $v$ it should be “computationally difficult to invert” hash function $hash$, that is, to find any message $w$ such that $v = hash(w)$. The pre-image resistance property together with the “easy computation” property (given $w$ it is easy to compute a value $v = hash(w)$) is known as the one-way property.

- **Second pre-image resistance**.
  
  Given an input $w$ it should be “computationally difficult” to find another input $w'$ such that $w \neq w'$ and $hash(w) = hash(w')$. Functions that lack this property are vulnerable to second pre-image attacks.

- **Collision resistance**.
  
  It should be “computationally difficult” to find two different messages $w$ and $w'$ such that $hash(w) = hash(w')$. Such a pair is called a cryptographic hash collision. This property is sometimes referred to as strong collision resistance.
The “computationally difficult (hard) problem” means that for the problem considered there must be no algorithm (oriented for realization in realistic computational model) except enumeration algorithm of possible instances that potentially fit the problem solution. Classical cryptographic functions rely on hardness of certain mathematical problems, such as integer factorization and discrete logarithm. Besides these well known problems several other potentially hard problems were discovered and such investigations are still in progress.

The main problem arrives in this aspect is to proof for a certain candidate, that considered problem is really hard. However, proving for a particular function the one-way property would imply that $P \neq NP$. The latest problem is a modern mathematical challenge of era.

In contrast to classical approach quantum cryptography is based on foundations of quantum mechanics and information properties of quantum systems. In the fall of the last century and last decades several models of quantum one-way functions were proposed. In [1] a family of “classical-classical” functions was considered, whose inputs and outputs are classical binary strings. These functions are candidates to be hardly invertible not only classically but also quantumly. Authors call such functions quantum one-way functions.

Quantum one-way functions defined by Gottesman and Chuang [2] are “classical-quantum” one-way functions, that is, such a function takes a classical bit string as an input and produces a quantum state. Another type of “classical-quantum” one-way function was invented by Buhrman et al. [3] based on binary error-correcting code and is known as quantum fingerprinting. Based on classical-quantum notion of a quantum one-way function several schemes of quantum digital signature were proposed [2, 4, 5].

In this research paper we define a notion of “classical-quantum” hashing function which is a natural extension of the notion of “classical-quantum” one-way function. We present a non-binary variant of quantum hashing function and prove its cryptographic properties. We propose an effective computation scheme of the quantum hashing function based on the quantum branching program model. Finally, as an application of quantum hashing functions we present a digital signature scheme.

**Organization of the paper.** We start the paper with the discussion of quantum one-way function definition and we come to the notion of quantum hash function (section 2). In the section 3 we show that known fingerprinting function [3] is a quantum hashing function. In the section 4 we present our variant of quantum hashing function and prove that it has a desired hashing property.

The section is devoted to the application of quantum hash functions for a digital signature scheme.

## 2 Quantum one-way and hashing functions

The definition of a quantum one-way function is based on [2] and explicitly presented in [4, 5]. Let

$$\psi : \{0, 1\}^n \rightarrow (\mathcal{H}^2)^{\otimes s}$$

be a function (classical-quantum function), where

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

is a $2^s$-dimensional Hilbert space made up of $s$ copies of a single qubit space $\mathcal{H}^2$.

We will also use notation

$$\psi : w \mapsto |\psi(w)\rangle$$

for $\psi$, which is frequently used in different papers.
• Function $\psi$ is called a quantum one-way if it is
  
  – easy to compute: there is quantum polynomial-time algorithm that on input $w$ outputs $|\psi(w)\rangle$.
  
  – hard to invert: given $|\psi(w)\rangle$, it is impossible to invert $w$ by virtue of fundamental quantum information theory.

**Property 2.1** If $n \gg s$ in the definition above, then given $|\psi(w)\rangle$, it is impossible to obtain $w$.

*Proof.* This pre-image resistance property follows from Holevo bound [6], since no more than $O(s)$ classical bits of information can be extracted from $s$ qubits and the original message contains $n \gg s$ bits. □

**Example 2.1 (One-way function)** A word $w \in \{0, 1\}^n$ is encoded by a single qubit:

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

*Here we treat $w = w_0 \ldots w_{n-1}$ also as a number $w = w_0 + w_1 2^1 + \ldots + w_{n-1} 2^{n-1}$.*

Clearly, we have that $\psi$ has the one-way property of the definition and the Property 2 above. What we need in extra and what is implicitly assumed in various papers (see for example [4, 5]) is a collision resistance property. However, there is still no such notion as *quantum collision*. The reason why we need to define it is the observation that in quantum hashing there might be no collisions in the classical sense: since quantum hashes are quantum states they can store arbitrary amount of data and can be different for unequal messages. But the procedure of comparing those quantum states implies measurement, which can lead to collision-type errors.

So, a *quantum collision* is a situation when a procedure that tests an equality of quantum hashes outputs true, while hashes are different. This procedure can be a well-known SWAP-test [3] or something that is adapted for specific hashing function. Anyway, it deals with the notion of distinguishability of quantum states. And since non-orthogonal quantum states cannot be perfectly distinguished, we require them to be “nearly orthogonal”.

To formalize the notion of “nearly orthogonality” we will call states $|\psi_1\rangle$ and $|\psi_2\rangle \delta$-orthogonal if

$$|\langle \psi_1 | \psi_2 \rangle| < \delta.$$ 

Thus, for a quantum hash-function it is important to have an ability to reliably compare quantum hashes of different words and those quantum states need to be distinguishable with high probability, that is, they have to pass non-equality tests.

**REVERSE-test.** Whenever we need to check if a quantum state $|\psi(w)\rangle$ is a hash of a classical message $v$, one can use the procedure that we call a *Reverse-test* (the idea of such test for the case of a quantum message given by $|v\rangle$ was described in [2], but it had not been given its own name).

Essentially the test applies the procedure that inverts the creation of a quantum hash, i.e. it “uncomputes” the hash to the initial state (usually the all-zero state).

Formally, let the procedure of quantum hashing of message $w$ consist of unitary transformation $U(w)$, applied to initial state $|0\rangle$, i.e. $|\psi(w)\rangle = U(w)|0\rangle$. Then the Reverse-test, given $v$ and
\(|\psi(w)\rangle\), applies \(U^{-1}(v)\) to the state \(|\psi(w)\rangle\) and measures the resulting state. It outputs \(v = w\) iff the measurement outcome is \(|0\rangle\). So, if \(v = w\), then \(U^{-1}(v)|\psi(w)\rangle\) would always give \(|0\rangle\), and REVERSE-test would give the correct answer. Otherwise, the resulting state would be \(\delta\)-orthogonal to \(|0\rangle\) since unitary operators preserve inner product:

\[
(|0\rangle, U^{-1}(v)|\psi(w)\rangle) = (U^{-1}(v)|\psi(v)\rangle, U^{-1}(v)|\psi(w)\rangle) = (|\psi(v)\rangle, |\psi(w)\rangle) = (\langle \psi(v) | \psi(w) \rangle < \delta.
\]

Overall, for this test has one-sided error bounded by \(\delta\) if quantum hashes of different messages are \(\delta\)-orthogonal.

**SWAP-test.** A more general test, that checks the equality of two arbitrary states is a well-known SWAP-test [3], given by the following circuit:

![SWAP-test circuit](image)

Applied to quantum hash codes it outputs \(|\psi(w)\rangle = |\psi(v)\rangle\), if the measurement result of the first qubit is \(|0\rangle\).

**Property 2.2** The probability of obtaining \(|0\rangle\) in the SWAP-test is equal to \(\frac{1}{2} (1 + |\langle \psi(w) | \psi(v) \rangle|^2)\).

*Proof.* See [3].

The probability of error of the SWAP-test inherently depends on the value of the inner product of \(|\psi(w)\rangle\) and \(|\psi(v)\rangle\) – it is minimal (close to \(1/2\)), when these states are “nearly orthogonal” [2].

Thus, the property of being \(\delta\)-orthogonal for quantum states is crucial for quantum collision resistance [2]. And at this point we come to a notion of a \(\delta\)-resistance.

**Definition 2.1 (\(\delta\)-resistance)** We call a function \(\psi : w \mapsto |\psi(w)\rangle\) \(\delta\)-resistant if for any pair of inputs \(w, w'\), \(w \neq w'\) their images are \(\delta\)-orthogonal:

\[|\langle \psi(w) | \psi(w') \rangle| < \delta.\]

\(\delta\)-resistance of a hash function is a key property for bounding the probability of error for SWAP-test and REVERSE-test. Note, that the ideas of \(\delta\)-resistance and REVERSE-test were used in [3].

Note, that this \(\delta\)-resistance property also corresponds to the classical *Second pre-image resistance*, since we cannot find two different messages for which the SWAP-test would erroneously output true with probability close to 1.

Finally, we naturally come to the following definition of classical-quantum hash function.

**Definition 2.2 ((\(n, s, \delta\))-quantum hash function)** We call a function

\[\psi : \{0, 1\}^n \rightarrow (\mathcal{H}^2)^\otimes s\]

\((n, s, \delta)\)-quantum hash function, if it is quantum one-way and \(\delta\)-resistant function.

The following property is an immediate implication of Definition 2.2 and Property 2.2.
Property 2.3 If a function $\psi : w \mapsto |\psi(w)\rangle$ is $(n, s, \delta)$-quantum hash function, then the SWAP-test distinguishes the hashes of two messages $w \neq w'$ with probability $\frac{1}{2}(1-\delta^2)$.

Proof. By the Property 2.2 the probability of error for the SWAP-test is $\frac{1}{2} \left( 1 + |\langle \psi(w) | \psi(w') \rangle|^2 \right)$. Since for $\delta$-resistant function $|\langle \psi(w) | \psi(w') \rangle| < \delta$, this test distinguishes quantum hashes of the pair of messages $w \neq w'$ with probability $\frac{1}{2}(1-\delta^2)$. □

Remark 2.1 The error probability of the SWAP-test can be reduced to any $\epsilon > 0$ by standard repetition technique, that is by performing this test upon $k = O(\log 1/\epsilon)$ copies of compared states. In other words, we could have used a function $\psi' : 0^n, 1^n \mapsto |\psi'(u)\rangle$, given by $|\psi'(u)\rangle = |\psi(u)\rangle^\otimes k = |\psi(u)\rangle \otimes \ldots \otimes |\psi(u)\rangle$

In this case, the total number of qubits to encode a word of length $n$ is $O(\log n \log(1/\epsilon))$.

In the next two sections we show that known quantum fingerprinting function is a quantum hashing function and we present our construction of quantum hash function with a slightly different characteristics.

3 Quantum Fingerprinting

In [3] Buhrman et al. defined a quantum one-way function $f_E : u \mapsto |f_E(u)\rangle$ of a bit string $u \in \{0, 1\}^n$, which they have called quantum fingerprinting. Based on the existence of the binary error-correcting code $E : \{0, 1\}^n \mapsto \{0, 1\}^m$ with $m = cn$ and Hamming distance $(1-\delta)m$ (for $c > 2$ and $\delta < 9/10 + 1/(15c)$) they have defined a quantum fingerprint of $u$ as follows:

$$|f_E(u)\rangle = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} (-1)^{E_i(u)} |i\rangle$$

where $E_i(u)$ denotes the $i$-th bit of $E(u)$.

Property 3.1 For a $\delta \approx 9/10 + 1/(15c)$ the quantum fingerprinting function $f_E$ is an $(n, O(\log n), \delta)$-quantum hashing function.

Proof. The function $f_E$ is the quantum one-way and is $\delta$-orthogonal for the $\delta \approx 9/10 + 1/(15c)$ [3]. □
4 Quantum Hashing

In this section we propose a quantum hashing function based on construction from [7].

Let $N = 2^n$. Let $K = \{k_i : k_i \in \{0, \ldots, N - 1\}\}$ and $d = |K|$. We define a classical-quantum function

$$h_K : \{0, 1\}^n \rightarrow (\mathbb{C}^2)^{\otimes (d+1)}$$

as follows. For a message $M \in \{0, 1\}^n$ we let

$$|h_K(M)\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle \left( \cos \frac{2\pi k_i M}{N} |0\rangle + \sin \frac{2\pi k_i M}{N} |1\rangle \right).$$

**Theorem 4.1** For arbitrary $\epsilon > 0$ there exists a set $K$ with $|K| = \lceil (2/\epsilon^2) \ln(2N) \rceil$ such that quantum function $h_K$ is an $(n, O(\log n + \log 1/\epsilon), \epsilon)$-quantum hashing function.

4.1 Proof of the Theorem 4.1

To prove the Theorem 4.1 we will show that the function $h_K$ is a one-way function and has $\delta$-resistance.

4.1.1 $\delta$-resistance of $h_K$

To prove $\delta$-resistance $h_K$ we recall some definitions and statements that we will use later.

**Definition 4.1** The discrete Fourier transform of the characteristic function of a set $K \subset \mathbb{Z}_N$ is the function

$$f_K(l) = \sum_{k \in K} e^{-2\pi k l / N}.$$ 

Let $\lambda(K) = \max_{l \neq 0} \frac{|f_K(l)|}{|K|}$ and let $\delta(K) = \max_{l \neq 0} \frac{|\text{Re}(f_K(l))|}{|K|}$. As it was mentioned in [8], $\lambda(K)$ gives some measure of randomness of the set $A$: the smaller it is, the more “random” $A$ is. $\delta(K)$ is defined similarly, but it uses the real part of $f_K(l)$. Clearly, $\delta(K) \leq \lambda(K)$.

For our technique will need a set $K$ with $\delta(K)$ as small as possible. In [8] a construction is given that for $\epsilon = \left(\frac{1}{\log N}\right)^{O(1)}$ yields a set $K$ with $|K| = (\log N)^{O(1)}$ and $\lambda(K) \leq \epsilon$.

Additionally, in [9] a proof of the following lemma is given.

**Lemma 4.1** For any $\epsilon \in (0, 1)$ there exists a set $K$ with $|K| = \lceil \frac{2}{\epsilon^2} \ln(2N) \rceil$ and $\delta(K) < \epsilon$.

Note, that in [9] we did not use the notation $\delta(K)$, the sum $\frac{1}{|K|} \sum_{k \in K} \cos \frac{2\pi k}{N}$ was used instead.

Let $\delta \in (0, 1)$ and pick a set $K$ satisfying Lemma 4.1 for $\epsilon = \delta^2$. In this case $h_K$ is $\delta$-resistant, since for any pair of messages $M_1 \neq M_2$

$$|\langle h_{M_1} | h_{M_2} \rangle| = \left| \frac{1}{|K|} \sum_{i=1}^{|K|} \left( \cos \frac{2\pi k_i M_1}{N} \cos \frac{2\pi k_i M_2}{N} + \sin \frac{2\pi k_i M_1}{N} \sin \frac{2\pi k_i M_2}{N} \right) \right| =$$
Thus, $h_K$ hashes $n$-bit messages into quantum states of $O(\log n + \log 1/\delta^2)$ qubits and provides $\delta$-orthogonality of quantum hashes.

### 4.1.2 $h_k$ is a quantum one-way function

**Irreversibility.** Since to hash an $n$-bit message we use about $O(\log n)$ qubits, then by Holevo bound no more than $O(\log n)$ bits of information can be extracted from it, i.e. one cannot restore all of $n$ bits.

**Effective computation.** The proposed hashing function can be efficiently implemented in quantum computational models that allow classical control, such as Quantum Branching Programs [10].

Below is a read-once quantum branching program (a quantum OBDD) that hashes an $n$-bit string $M = b_1 b_2 \ldots b_n$ into $|h_K(M)\rangle$ using $O(\log n)$ qubits:

![Quantum Branching Program Diagram]

Here, $R(\theta_{i,j})$ denotes a rotation by the angle $\frac{4\pi k_i 2^j}{N}$ around the $\hat{y}$ axis of the Bloch sphere, and $d = |K|$.

### 4.2 Numerical Results

The aforementioned methods for computing the set $K$ of hashing parameters make sense for comparatively large $N$. For smaller $N$ the influence of $\epsilon$ results in a quite big sizes of the set $K$. This problem is especially important for quantum digital signature protocol, where the value of $N$ is not very large.

To deal with this problem we have developed a genetic algorithm that gives good results in acceptable time. Here are some examples of its work:
The proposed quantum hashing is a suitable one-way function for quantum digital signature protocol from [2] and below we describe its basic structure modified for the specific hashing function.

To sign a single message bit $b$ Alice picks from $\{1, \ldots, L\}$ uniformly at random a pair of keys $(K_0, K_1)$. This pair constitutes her private key.

Using her private key pair Alice creates a sufficient numbers of public key pairs

\[
(h_{K_0}), (h_{K_1})
\]

and sends them to potential recipients. It can be easily verified that quantum states in each pair are nearly orthogonal and thus distinguishable with high probability.

Now, given a message bit $b$, Alice sends it to Bob together with a part of her private key $K_b$, which constitutes her signature.
Finally, Bob, the recipient of a signed message, validates the signature by “uncomputing” $|h_{K_b}\rangle$ the same way it was created, i.e. by a sequence of controlled rotations by the negative angles and the Hadamard transform. If the signature is correct Bob will always obtain the all-zero state out of it. Otherwise, the probability of error will be bounded by $\varepsilon$ due to $\varepsilon$-collision resistance of the hashing function.

This protocol uses $O(\log \log L)$ qubits for public keys, where $L$ is a security level parameter and it should be chosen to deal with the $1/L$ probability of guessing what the private key is by the possible forger.

Acknowledgements. Research was supported by the Russian Fund for Basic Research (under the grants 11-07-00465, 12-01-31216).

References

[1] Elham Kashefi and Iordanis Kerenidis. Statistical zero knowledge and quantum one-way functions. Theoretical Computer Science, 378(1):101–116, June 2007.

[2] Daniel Gottesman and Isaac Chuang. Quantum digital signatures. Technical Report arXiv:quant-ph/0105032, Cornell University Library, Nov 2001.

[3] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. Phys. Rev. Lett., 87(16):167902, Sep 2001.

[4] Xin Lu and Dengguo Feng. Quantum digital signature based on quantum one-way functions. In The 7th International Conference on Advanced Communication Technology, 2005, ICACT 2005, volume 1, pages 514–517, 2005.

[5] JingXian Zhou, YaJian Zhou, XinXin Niu, and YiXian Yang. Quantum proxy signature scheme with public verifiability. Science China Physics, Mechanics and Astronomy, 54(10):1828–1832, 2011.

[6] Alexander S. Holevo. Some estimates of the information transmitted by quantum communication channel (russian). Probl. Pered. Inform. [Probl. Inf. Transm.], 9(3):311, 1973.

[7] Farid Ablayev and Alexander Vasiliev. Algorithms for quantum branching programs based on fingerprinting. Electronic Proceedings in Theoretical Computer Science, 9:1–11, 2009.

[8] Alexander A. Razborov, Endre Szemeredi, and Avi Wigderson. Constructing small sets that are uniform in arithmetic progressions. Combinatorics, Probability & Computing, 2:513–518, 1993.

[9] Farid Ablayev and Alexander Vasiliev. On the computation of boolean functions by quantum branching programs via fingerprinting. Electronic Colloquium on Computational Complexity (ECCC), 15(059), 2008.

[10] Farid Ablayev, Aida Gainutdinova, and Marek Karpinski. On computational power of quantum branching programs. In FCT, pages 59–70, 2001.