Quantum hashing via ϵ-universal hashing constructions, Freivald's fingerprinting schemas and error correcting codes

Farid Ablayev Marat Ablayev

Theory Days at Ratnieki 2014

October, 2014

The talk is based on papers:

1. F. Ablayev, A. Vasiliev : Cryptographic quantum hashing, Laser Physics Letters Volume 11 Number 2, 2014
2. F. Ablayev, M. Ablayev : Quantum Hashing via Classical epsilon-universal Hashing Constructions, 2014 arXiv:1404.1503 [quant-ph] (2014)
3. F. Ablayev, M. Ablayev : Quantum Hashing via ϵ-universal Hashing Constructions and Freivalds Fingerprinting Schemas, Springer LNCS volume 8614, Proceedings of the 16th International Workshop on Descrip-tional Complexity of Formal Systems, (DCFS), 42-52 p., 2014.

Motivation

- P. Shor 1994. Quantum algorithm for integer factorization.
- "Post-quantum cryptography" http://pqcrypto.org/
  The book: Daniel J. Bernstein, Johannes Buchmann, Erik Dahmen (editors). Post-quantum cryptography. Springer, 2009.
  - ...
  - Hash-based signature schemes such as L. Lamport signatures and R. Merkle signature schemes.
- Hashing itself is an important basic concept for the organization transformation and reliable transmission of information.
  - In 1995 A. Wigderson characterizes universal hashing as being a tool which "should belong to the fundamental bag of tricks of every computer scientist".

ϵ-Universal Hash Family

Σ^k, Σ^m – sets of words. $|\Sigma^k| = K$, $|\Sigma^m| = M$.

A hash function is a map $f : \Sigma^k \to \Sigma^m$ with $K > M$.

ϵ-Universal hash family

A hash family $F = \{f_1, \ldots, f_N\}$ is called ϵ-Universal, if for any two distinct words w, w':

$$|\{f \in F : f(w) = f(w')\}| \leq \epsilon N.$$

F – ϵ-U $(N; K, M)$

ϵ-Universal Hash Family. Probabilistic approach

Property

Let F be an ϵ-Universal hash family. If the $f \in F$ is chosen uniformly at random, then the probability that any two distinct words collide under f is at most ϵ.

- The parameter ϵ is often referred to as the collision probability of the hash family F.
- The case of $\epsilon = 1/N$ is known as universal hashing.

Quantum function

- Mathematically. Qubit

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle, \qquad |||\psi\rangle||^2 = |a_0|^2 + |a_1|^2 = 1$$

  is a unit vector in the two-dimensional Hilbert complex space $\mathcal{H}^2$.
- $(\mathcal{H}^2)^{\otimes s} = \mathcal{H}^2 \otimes \cdots \otimes \mathcal{H}^2$ – (2^s)-dimensional Hilbert space of s qubits

$$|\psi\rangle = \sum_{i=1}^{2^s} a_i|i\rangle, \qquad |||\psi\rangle||^2 = 1.$$

Quantum (classical-quantum) function

$$\psi : \mathbb{X} \to (\mathcal{H}^2)^{\otimes s}, \qquad \psi : w \mapsto |\psi(w)\rangle \qquad (\psi : |0\rangle, w \mapsto |\psi(w)\rangle).$$

Quantum function. Example

- Word (binary) $w = w_0 \dots w_{k-1}$.
- Number $w = \sum_{i=0}^{k-1} w_i 2^i$.

$$0 \leq w \leq 2^k - 1.$$

Example

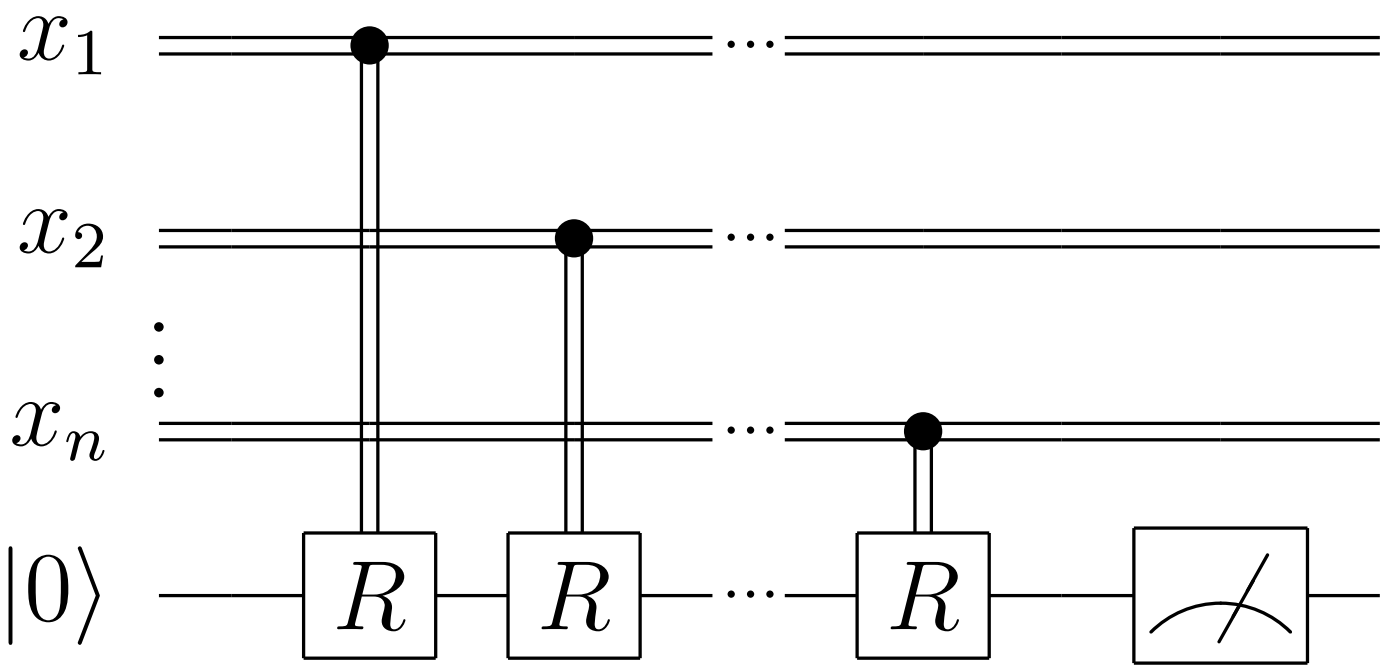
$$\psi : \{0,1\}^k \to \mathcal{H}^2$$

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

$$|\psi(w)\rangle = a_0(w)|0\rangle + a_1(w)|1\rangle =$$

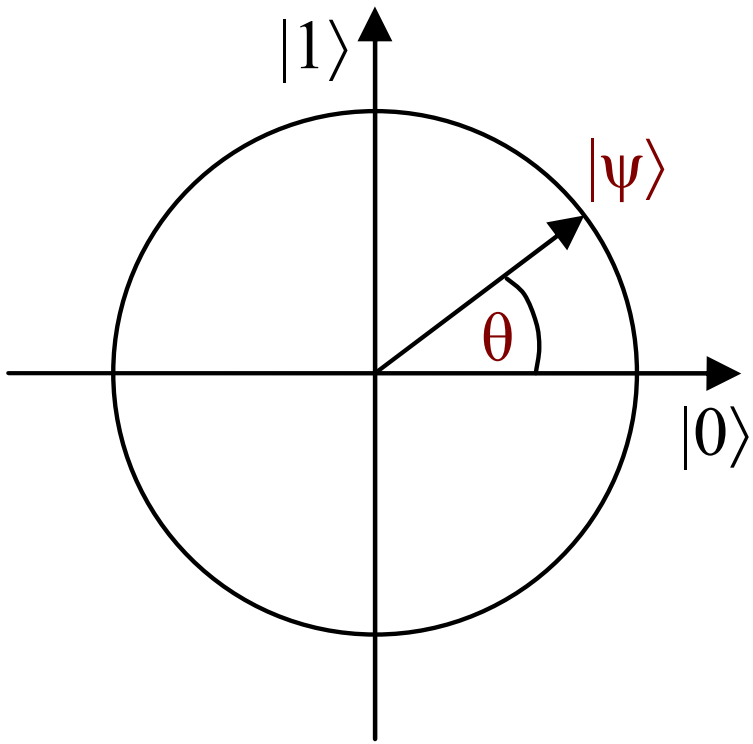
$$|\psi(w)\rangle = \cos\left(\frac{2\pi w}{2^k}\right)|0\rangle + \sin\left(\frac{2\pi w}{2^k}\right)|1\rangle,$$

Computational model



$|\psi(w)\rangle$ – one qubit

$$|\psi(w)\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle = \cos\left(\frac{2\pi w}{2^k}\right)|0\rangle + \sin\left(\frac{2\pi w}{2^k}\right)|1\rangle,$$



δ-Resistant $(|\Sigma^k|, s)$ quantum function

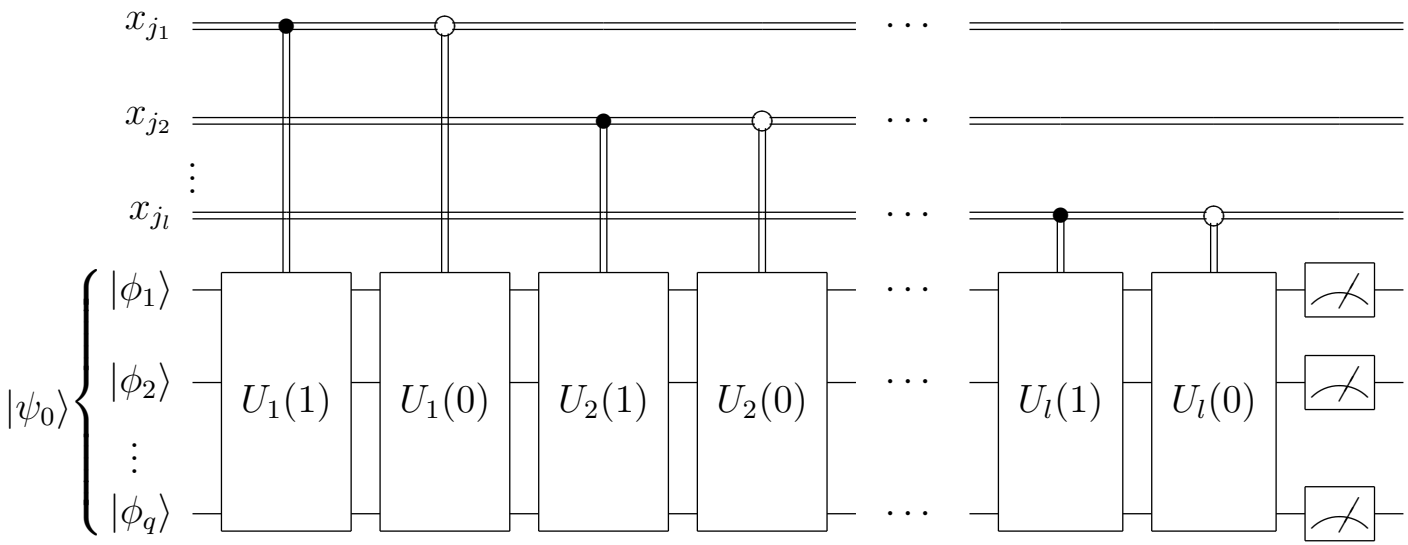
Definition

$$\psi : \Sigma^k \to (\mathcal{H}^2)^{\otimes s}$$

δ-Resistant $(|\Sigma^k|, s)$ function, if for different words $w, w' \in \Sigma^k$

$$\left|\langle \psi(w) \mid \psi(w') \rangle\right| \leq \delta.$$

Computational model – Quantum Branching Program – quantum case of Algebraic Branching Program



Lamport digital one bit signature (Quantum variant)

1. Alice secretly selects
   - a word $w = \sigma_1 \dots \sigma_k$ for the bit 0
   - a word $v = \sigma'_1 \dots \sigma'_k$ for the bit 1.
2. Alice prepares two pairs (quantum state and a classical bit):

$$(|\psi(w)\rangle, 0) \qquad \text{and} \qquad (|\psi(v)\rangle, 1)$$

   by preparing states $\psi : |0\rangle, w \mapsto |\psi(w)\rangle$ and $\psi : |0\rangle, v \mapsto |\psi(v)\rangle$
3. Alice sends pairs $(|\psi(w)\rangle, 0)$ and $(|\psi(1)\rangle, 1)$ to Bob.
   Bob keeps these pairs.
4. - Alice decided to sign the bit 1. Then
   - Alice sends (classical) pair $(v, 1)$ to Bob.
5. Bob Reverse $|\psi(v)\rangle$ using the word v and get a quantum state $|\psi\rangle$.
   Bob verify whether $|\psi\rangle = |0\rangle$.

Lower bound for s for δ-Resistant $(|\Sigma^k|, s)$ quantum function

Theorem (Lower Bound)

If $\psi : \Sigma^k \to (\mathcal{H}^2)^{\otimes s}$ is δ-Resistant $(|\Sigma^k|, s)$ quantum function then

$$s \geq \log k + \log\log|\Sigma| - \log\log\left(1 + \sqrt{2/(1-\delta)}\right) - 1.$$

$$|||\psi\rangle - |\psi'\rangle||^2 = |||\psi\rangle||^2 + |||\psi'\rangle||^2 - 2\langle\psi \mid \psi'\rangle = 2 - 2\langle\psi \mid \psi'\rangle.$$

Property

If ψ is δ-Resistant, then for w, w'

$$\rho(|\psi(w)\rangle, |\psi(w')\rangle) = |||w\rangle - |w'\rangle|| \geq \sqrt{2(1-\delta)}.$$

δ-R $(|\Sigma^k|, s)$ Quantum hash function

Definition

We call δ-Resistant $(|\Sigma^k|, s)$ quantum function ψ a δ-R $(|\Sigma^k|, s)$ Quantum hash function

Quantum function generated by classical function. Example

Word $w = w_0 \dots w_{k-1}$, number $w = \sum_{i=0}^{k-1} w_i 2^i$

$$h : \{0,1\}^k \to \mathbb{F}_{2^k}, \qquad h(w) = w$$

Function h generates quantum function

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

$$|\psi_h(w)\rangle = a_0(w)|0\rangle + a_1(w)|1\rangle =$$

$$= \cos\left(\frac{2\pi h(w)}{2^k}\right)|0\rangle + \sin\left(\frac{2\pi h(w)}{2^k}\right)|1\rangle$$

Quantum function generated by classical function

Discrete function $g : \Sigma^k \to \mathbb{F}_q$
Let $\ell \geq 1$. Quantum function

$$\psi_g : \Sigma^k \to (\mathcal{H}^2)^{\otimes \ell}$$

determined by the rule

$$\psi_g : w \mapsto |\psi_g(w)\rangle = \sum_{i=1}^{2^\ell} \alpha_i(g(w))|i\rangle,$$

amplitudes $\alpha_i(g(w))$, $i \in \{1, \ldots, 2^\ell\}$ of the state $|\psi_g(w)\rangle$ determined by $g(w)$,

$$\sum_{i=1}^{2^l} |\alpha_i(g(w))|^2 = 1.$$

Quantum function generated by a family of functions. Example

Binary word $w = w_0 \dots w_{k-1}$, number $w = \sum_{i=0}^{k-1} w_i 2^i$, $b_j \in \mathbb{F}_q$.
Family $H = \{h_1, \dots, h_T\}$

$$h_j(w) = b_j w \pmod q.$$

Quantum function $\psi_{h_j} : \{0,1\}^k \to \mathcal{H}^2$ generated by $h \in H$

$$|\psi_{h_j}(w)\rangle = \cos \frac{2\pi h_j(w)}{q} |0\rangle + \sin \frac{2\pi h_j(w)}{q} |1\rangle$$

Quantum function $\psi_H : \{0,1\}^k \to (\mathcal{H}^2)^{\otimes(\log T+1)}$ generated by H

$$|\psi_H(w)\rangle = \frac{1}{\sqrt{T}} \sum_{j=1}^{T} |j\rangle |\psi_{h_j}(w)\rangle =$$

$$\frac{1}{\sqrt{T}} \sum_{j=1}^{T} |j\rangle \left(\cos \frac{2\pi h_j(w)}{q} |0\rangle + \sin \frac{2\pi h_j(w)}{q} |1\rangle \right).$$

Quantum fingerprinting (2001) = binary quantum hash function

H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf

- Let $E : \{0,1\}^k \to \{0,1\}^n$ be an (n,k,d) error correcting code with Hamming distance $d \geq n - \delta n$.
- Family $E = \{E_1, \dots E_n\}$, here $E_i(w)$ – i-th bit of code word.

Quantum hash function $\psi_{F_E} : \{0,1\}^k \to (\mathcal{H}^2)^{\otimes s}$,

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

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle \left(\cos \frac{\pi E_i(w)}{2} |0\rangle + \sin \frac{\pi E_i(w)}{2} |1\rangle \right),$$

Property

For $s = \log n + 1$, $\delta \geq (1 - d/n)$ function ψ_{F_E} is an δ-R $(2^k, s)$ quantum hash function.

"Non binary" quantum hash function (2008)

F. Ablayev, A. Vasiliev

$\mathbb{F}_q$ – finite field, q – prime power. $H = \{h_1, \dots h_T\}$ where

$$h_j : \mathbb{F}_q \to \mathbb{F}_q \qquad h_j(w) = b_j w \pmod q.$$

For $s = \log T + 1$ Quantum function $\psi_H : \mathbb{F}_q \to (\mathcal{H}^2)^{\otimes s}$,

$$|\psi_H(w)\rangle = \frac{1}{\sqrt{T}} \sum_{j=1}^{T} |j\rangle \left(\cos \frac{2\pi h_j(w)}{q} |0\rangle + \sin \frac{2\pi h_j(w)}{q} |1\rangle \right).$$

Property (Ablayev, Vasiliev 2013)

For $\delta > 0$, for $T = \lceil (2/\delta^2) \ln(2q) \rceil$, for $s = \log T + 1$ there exists a family

$$H_{\delta,q} = \{h_1, \dots, h_T\}$$

such that $\psi_{H_{\delta,q}}$ is an δ-R (q, s) quantum hash function.

Quantum hash generator

Let $G = \{g_1, \ldots, g_D\}$ be a family of functions $g_j : \Sigma^k \to \mathbb{F}_q$. Let $\ell \geq 1$ be an integer and ψ_{g_j}, $j \in \{1, \ldots, D\}$, be a quantum functions

$$\psi_{g_j} : \Sigma^k \to (\mathcal{H}^2)^\ell,$$

determined by $g_j \in G$. Let $d = \log D$. We define a quantum function

$$\psi_G : \Sigma^k \to (\mathcal{H}^2)^{\otimes(d+\ell)}$$

by the rule

$$\psi_G(w) = \frac{1}{\sqrt{D}} \sum_{j=1}^{D} \underbrace{|j\rangle}_{d} \underbrace{|\psi_{g_j}(w)\rangle}_{\ell}.$$

We call G a δ-R $(|\Sigma^k|, d + \ell)$ quantum hash generator, if ψ_G is an δ-R $(|\Sigma^k|, d + \ell)$ quantum hash function.

Examples of quantum hash generator

Binary

For binary (n, k, d) error correcting code $E : \{0,1\}^k \to \{0,1\}^n$ with Hamming distance d the following is true.
For $\delta = 1 - d/n$ The family

$$E = \{E_1, \ldots E_n\}$$

is δ-R $(2^k, \log n + 1)$ quantum hash generator

Non binary

For $\delta > 0$, for q prime power, for $T = \lceil (2/\delta^2) \ln(2q) \rceil$ there exists a set

$$H_{\delta,q} = \{h_1, \ldots, h_T\}$$

which is an δ-R $(q, \log T + 1)$ quantum hash generator.

Quantum hashing via classical hashing constructions

- Let $F = \{f_1, \ldots, f_N\}$ be an ϵ-U $(N; |\Sigma^k|, q)$ hash family

$$f_i : \Sigma^k \to \mathbb{F}_q.$$

- Let $H = \{h_1, \ldots, h_T\}$

$$h_j : \mathbb{F}_q \to \mathbb{F}_q.$$

  be an δ-R $(q, \log T + \ell)$ quantum hash generator.
- Define composition $G = F \circ H$ of families F and H

$$G = \{g_{ij}(w) = h_j(f_i(w)) : i \in \{1, \ldots, N\}, j \in \{1, \ldots, T\}\},$$

Theorem

$G = F \circ H$ is an Δ-R $(|\Sigma^k|, s)$ quantum hash generator, where

$$\Delta \leq \epsilon + \delta \quad \text{and} \quad s = \log N + \log T + \ell.$$

Quantum hashing based on Freivalds' fingerprinting 1979

For $w \in \{0,1\}^k$ (also $w \in \mathbb{F}_{2^k}$), for the i-th prime p_i a function

$$f_i : \{0,1\}^k \to \mathbb{F}_{p_i} \qquad f_i(w) = w \pmod{p_i}.$$

is a fingerprint of w.

Freivalds 1979

- Pick $c > 1$, pick $M = ck \ln k$.
- $\pi(M)$ – the number of primes less than or equal to M.
- $\pi(M) \sim M/\ln M$ as $M \to \infty$.
- The set

$$F_M = \{f_1, \ldots, f_{\pi(M)}\}$$

of fingerprints is a $(1/c)$-U $(\pi(M); 2^k, M)$ hash family.

Quantum hashing based on Freivalds' fingerprinting

Theorem

1. Let $c > 1$, let $M = ck \ln k$. Let $F_M = \{f_1, \ldots, f_{\pi(M)}\}$ be a $(1/c)$-U $(\pi(M); 2^k, M)$ hash family.
2. Let $q \in \{M, \ldots, 2M\}$ be a prime, let $\delta > 0$. Let $H_{\delta,q} = \{h_1, \ldots, h_T\}$ be an δ-R $(q, \log T + 1)$ quantum hash generator.

Then family $G = F_M \circ H_{\delta,q}$ is a Δ-R $(2^k; s)$ quantum hash generator, where

$$\Delta \leq \frac{1}{c} + \delta \qquad s \leq \log ck + \log\log k + \log\log q + 2\log 1/\delta + 3.$$

Lower bond

$$s \geq \log k + \log\log q - \log\log\left(1 + \sqrt{2/(1-\delta)}\right) - 1.$$

Quantum hashing from universal linear hash family

1979-1980

Let $k > 0$ – integer, q – prime power, $\mathbb{X} = (\mathbb{F}_q)^k \backslash \{(0, \ldots, 0)\}$.
For every vector $a \in (\mathbb{F}_q)^k$ define hash function $f_a : \mathbb{X} \to \mathbb{F}_q$ by the rule

$$f_a(w) = \sum_{i=1}^{k} a_i w_i.$$

Then

$$F_{lin} = \{f_a : a \in (\mathbb{F}_q)^k\}$$

is an $(1/q)$-U $(q^k; (q^k - 1); q)$ hash family (universal hash family).

Quantum hashing from universal linear hash family

Theorem

For arbitrary $\delta \in (0,1)$ composition $G = F_{lin} \circ H_{\delta,q}$ is a Δ-R $(q^k; s)$ quantum hash generator with $\Delta \leq (1/q) + \delta$ and

$$s \leq k \log q + \log\log q + 2\log 1/\delta + 3.$$

Lower bound

$$s \geq \log k + \log\log q - \log\log\left(1 + \sqrt{2/(1-\delta)}\right) - 1.$$

This lower bound shows that the quantum hash function ψ_G is not asymptotically optimal in the sense of number of qubits used for the construction.

Quantum hash functions based on error correcting codes

Theorem

Let $\mathcal{C}$ – be a linear $[n, k, d]_q$ ECC

$$\mathcal{C} : \mathbb{F}_q^k \to \mathbb{F}_q^n.$$

Then for arbitrary $\delta \in (0, 1)$ there exists Δ-R $(q^k; s)$ quantum hash generator G, where

$$\Delta = (1 - d/n) + \delta,$$

$$s \leq \log n + \log\log q + 2\log 1/\delta + 4.$$

Proof idea. Having $[n, k, d]_q$ ECC $\mathcal{C}$ one can construct $(1 - d/n)$-U $(n; q^k; q)$ hash family $F_{\mathcal{C}}$. J. Bierbrauer, T. Johansson, G. Kabatianskii, B. Smeets 1994

Quantum hash functions based on $[n, k, d]_q$ RS-code

q – prime power, $k \leq n \leq q$. A common special case is $n = q - 1$. Each word $w \in (\mathbb{F}_q)^k$, $w = w_0 w_1 \dots w_{k-1}$ associated with the polynomial

$$P_w(x) = \sum_{i=0}^{k-1} w_i x^i.$$

$$C_{RS} : (\mathbb{F}_q)^k \to (\mathbb{F}_q)^n \quad w \mapsto C_{RS}(w) = (P_w(1) \dots P_w(n))$$

$(k-1)/q$-U $(q; \mathbb{F}_q^k; q)$ hash family $F_{RS} = \{f_a : a \in A\}$ For $a \in \mathbb{F}_q \backslash 0$ define f_a

$$f_a : (\mathbb{F}_q)^k \to \mathbb{F}_q \quad f_a(w_0 \dots w_{k-1}) = \sum_{i=0}^{k-1} w_i a^i.$$

Quantum hash functions based on Reed-Solomon codes

Theorem.

Let q be a prime power and let $1 \leq k \leq q$. Then for arbitrary $\theta \in (0,1)$ there is a δ-R (q^k, s) quantum hash generator G_{RS} such that $\delta \leq \frac{k-1}{q} + \theta$ and $s \leq \log(k \log q) + 2 \log 1/\theta + 4$.

- If we select $n \in [ck, c'k]$ for constants $c < c'$, then $\Delta \leq 1/c + \delta$ for $\delta \in (0,1)$ and

$$s \leq \log(q \log q) + 2 \log 1/\Delta + 4.$$

Lower bound

$$s \geq \log(q \log q) - \log\log\left(1 + \sqrt{2/(1-\Delta)}\right) - \log c'/2$$

Thus, Reed Solomon codes provides good enough parameters for resistance value Δ and for a number s of qubits we need to construct quantum hash function ψ_{RS}.

Explicit constructions of G_{RS} and $\psi_{G_{RS}}$.

Let $H_{\delta,q} = \{h_1, \ldots, h_T\}$, where $h_j : \mathbb{F}_q \to \mathbb{F}_q$ and $T = \lceil (2/\delta^2) \ln 2q \rceil$. composition

$$G_{RS} = F_{RS} \circ H_{\delta,q} = \{g_{j\,i} = h_j(f_{a_i}) : j \in [T], i \in [n]\}$$

For $s = \log n + \log T + 1$ defines function $\psi_{G_{RS}}$ for a word $w \in (\mathbb{F}_q)^k$ by the rule.

$$\psi_{G_{RS}}(w) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle \otimes \left(\frac{1}{\sqrt{T}} \sum_{j=1}^{T} |j\rangle |\psi_{g_{ji}}(w)\rangle \right).$$

$$= \frac{1}{\sqrt{nT}} \sum_{i=1,j=1}^{n,T} \underbrace{|i\rangle|j\rangle}_{\log n + \log T} \otimes \underbrace{\left(\cos \frac{2\pi h_j(f_{a_i}(w))}{q} |0\rangle + \sin \frac{2\pi h_j(f_{a_i}(w))}{q} |1\rangle \right)}_{|\psi_{g_{ji}}(w)\rangle \text{ – one qubit}}.$$