Systematic Error-Correcting Codes for Rank Modulation

S. Buzaglo¹ E. Yaakobi ² T. Etzion ¹ J. Bruck²

¹Technion – Israel Institute of Technology

²California Institute of Technology

December 24, 2013

Outline

- Introduction
 - Rank Modulation for Flash Memories
 - Error Correction for Rank Modulation
- Systematic Codes for Rank Modulation
 - Definitions and Known Constructions
 - Construction of Systematic Error-Correcting Codes
 - Analysis for the Number of Redundancy Symbols
- 3 Conclusion

Flash Memory

- Flash memory is a Non-Volatile Memory technology that is both electrically programmable and electrically erasable.
- Programming is easy to perform on single cells. Erasure can only be done on large blocks of cells.
- Charge is slowly injected into the cell over several iterations.
- Common error factors: charge leakage and read disturbance.

Rank Modulation for Flash Memories

In rank modulation¹ data is represented by permutations.

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \rightarrow \sigma = [\sigma(1), \sigma(2), \dots, \sigma(n)],$$

where $x_{\sigma(1)} < x_{\sigma(2)} < \dots < x_{\sigma(n)}.$

Example

$$\mathbf{x} = (0.8, 1.5, 2.3, 1) \rightarrow \sigma = [1, 4, 2, 3].$$
 $x_1 < x_4 < x_2 < x_3$

¹A. Jiang, R. Mateescu, M. Schwartz, and J. Bruck, "Rank modulation for flash memories," *IEEE Trans. on Inform. Theory*, 2009.

Why Rank Modulation?

• Better programming efficiency.

Example

Programming $\sigma = [1, 4, 2, 3]$:

$$\mathbf{x} = (0.1, , ,) \rightarrow \mathbf{x} = (0.1, , , 0.5) \rightarrow$$

$$\mathbf{x} = (0.1, 1.2, ,0.5) \rightarrow \mathbf{x} = (0.1, 1.2, 2, 0.5).$$

 The ranking of the cell's charge levels is more robust to charge leakage.

Kendall's τ -Metric

Let S_n be the set of all permutations on n elements.

For $\sigma, \pi \in S_n$, the Kendall's τ -distance, $d_K(\sigma, \pi)$, is the minimum number of adjacent transpositions required to change σ into π .

Example

If
$$\sigma = [3, 2, 4, 1]$$
 and $\pi = [2, 1, 3, 4]$:

$$\sigma = [3, 2, 4, 1] \rightarrow [2, 3, 4, 1] \rightarrow [2, 3, 1, 4] \rightarrow [2, 1, 3, 4] = \pi.$$

$$d_{\mathcal{K}}(\sigma,\pi)=3.$$

Kendall's τ -Metric

$$d_K(\sigma,\pi) = |\{(i,j) \ : \ \sigma^{-1}(i) < \sigma^{-1}(j) \land \pi^{-1}(i) > \pi^{-1}(j)\}|.$$

Example

$$\sigma = [3, 2, 4, 1]$$
 and $\pi = [2, 1, 3, 4]$

$$d_{\mathcal{K}}([3,2,4,1],[2,1,3,4]) = |\{(3,1),(3,2),(4,1)\}| = 3.$$

Why Kendall's τ -Metric?

Small error corresponds to small Kendall's τ -distance².

$$\mathbf{x} = (0.8, 1.5, 2.3, 1)$$
 \xrightarrow{Error} $\mathbf{y} = (1.1, 1.2, 2.3, 1).$ \downarrow \downarrow $\sigma = [1, 4, 2, 3]$ $\pi = [4, 1, 2, 3]$ $d_{\mathcal{K}}(\sigma, \pi) = 1.$

²A. Jiang, M. Schwartz, and J. Bruck, "Correcting charge-constrained errors in the rank-modulation scheme," *IEEE Trans. on Inform. Theory*, 2010.

Systematic Codes for Permutations

Definition

A code $C \subseteq S_n$ is an (n, k) systematic code if for every $\sigma \in S_k$ there exists exactly one $\alpha \in C$ such that σ is a sub-permutation of α . |C| = k!.

The number of *redundancy symbols* of an (n, k) systematic code is $r \stackrel{\text{def}}{=} n - k$.

Factoradic Representation

Definition

For a permutation $\sigma \in S_n$, the *insertion vector*

$$\mathbf{g}_{\sigma} = (g_{\sigma,1}, g_{\sigma_2}, \dots, g_{\sigma,n-1})$$
 is defined by

$$g_{\sigma,i} \stackrel{\text{def}}{=} |\{j : j < i+1, \sigma^{-1}(j) > \sigma^{-1}(i+1)\}|, \qquad 1 \le i \le n-1.$$

Example

If $\sigma = [5, 2, 1, 4, 3]$ then $\mathbf{q}_{\sigma} = (1, 0, 1, 4)$. Conversely:

$$[1] \to [2,1] \to [2,1,3] \to [2,1,4,3] \to [5,2,1,4,3]$$

The mapping $\sigma \to \mathbf{g}_{\sigma}$ is a bijection of S_n into

$$\mathbb{Z}_n! = \mathbb{Z}_2 \times \mathbb{Z}_3 \times \ldots \times \mathbb{Z}_n.$$

Metric Embedding

The Manhattan distance between $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_n!$:

$$d_{M}(\mathbf{x},\mathbf{y}) \stackrel{\mathrm{def}}{=} \sum_{i=1}^{n-1} |x_{i}-y_{i}|.$$

The Lee distance between $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_q^N$:

$$d_L(\mathbf{x},\mathbf{y}) \stackrel{\text{def}}{=} \sum_{i=1}^N \min\{|x_i-y_i|, q-|x_i-y_i|\}.$$

Lemma (Jiang, Schwartz, and Bruck, 2010)

For every $\sigma, \pi \in S_n$ and q > n

$$d_{\mathcal{K}}(\sigma,\pi) \geq d_{\mathcal{M}}(\mathbf{g}_{\sigma},\mathbf{g}_{\pi}) \geq d_{\mathcal{L}}(\mathbf{g}_{\sigma},\mathbf{g}_{\pi})$$

Known Constructions of Systematic Codes

- Using a greedy approach, Zhou et al.³ proved the existence of an (n, k) systematic code with minimum distance d and $r = n k \le d$ redundancy symbols.
- Using BCH codes over the factoradic representation, Zhou et al. constructed an (n, k) systematic t-error-correcting code, where $n \ge 6t + 5$ and $r \le 2t + 1$.

³H. Zhou, M. Schwartz, A. Jiang, and J. Bruck, "Systematic error-correction codes for rank modulation," 2013

Systematic Single-Error-Correcting Code

A perfect single-error-correcting code in S_n does not exist, where n is a prime⁴.

In that case if C is an (n, k) single-error-correcting code then

$$k! = |\mathcal{C}| < (n-1)! \Rightarrow n \ge k+2.$$

Zhou et al. constructed a (k + 2, k) systematic single-error-correcting codes for all $k \ge 2$.

 $^{^4}$ S. Buzaglo and T. Etzion, "Perfect permutations codes with the Kendall's $\tau\text{-Metric.}$ " 2013.

Systematic Single-Error-Correcting Code

Construction (Zhou et al., 2013)

Let $m \in \{k, k+1\}$ be a prime and define C as follows. For all $\sigma \in S_k$, define $\alpha \in C$ where,

$$g_{\alpha,i}=g_{\sigma,i}, \quad 1\leq i\leq k-1$$

$$g_{\alpha,k} \equiv \sum_{i=1}^{k} (2i-1)\sigma(i) \ (mod \ m)$$

$$g_{\alpha,k+1} \equiv \sum_{i=1}^k (2i-1)^2 \sigma(i) \pmod{m}.$$

C is (k + 2, k) systematic single-error-correcting code.

Multi-Permutations

A multi-set $\mathcal{M} = \{v_1^{m_1}, v_2^{m_2}, \cdots, v_\ell^{m_\ell}\}$ is a collection of the elements $\{v_1, v_2, \dots, v_\ell\}$ in which every v_i appears m_i times.

A multi-permutation on \mathcal{M} is an ordering of the elements of \mathcal{M} .

Let $S(\mathcal{M})$ be the set of all multi-permutations on \mathcal{M} .

The Kendall's τ can be extended to $S(\mathcal{M})$.

Systematic Codes and Multi-Permutations

Let
$$\mathcal{M}_{k,r} = \{0^k, k+1, k+2, \dots, k+r\}.$$

For $\alpha \in S_{k+r}$ let $\alpha_{k \mapsto 0} \in S(\mathcal{M}_{k,r})$ obtained from α by replacing every element of $\{1, 2, \dots, k\}$ by 0.

Example

If
$$\alpha = [2, 5, 4, 1, 3, 6]$$
 and $k = 3$ then $\alpha_{k \mapsto 0} = [0, 5, 4, 0, 0, 6]$.

Systematic Codes and Multi-Permutations

For $\sigma \in S_k$, $\rho \in S(\mathcal{M}_{k,r})$, denote by $\sigma * \rho$ the permutation in S_{k+r} obtained by substituting σ in ρ .

Example

If $\rho = [0,6,0,0,5,7,0]$ and $\sigma = [2,4,1,3]$, then $\sigma * \rho = [2,6,4,1,5,7,3]$.

Systematic Codes and Multi-Permutations

An (n, k) systematic code C is equivalent to a mapping

$$\phi: \mathcal{S}_k \to \mathcal{S}(\mathcal{M}_{k,n-k}).$$

If σ is a sub-permutation of $\alpha \in \mathcal{C}$ then $\phi(\sigma) = \alpha_{k \mapsto 0}$ and $\sigma * \phi(\sigma) = \alpha$.

Lemma

For every $\sigma, \pi \in S_k$, $\rho_1, \rho_2 \in S(\mathcal{M}_{k,r})$

$$d_K(\sigma * \rho_1, \sigma * \rho_2) \ge d_K(\sigma, \pi) + d_K(\rho_1, \rho_2).$$

Construction of Systematic Codes

Ingredients:

1) Integers $h_1, h_2, \ldots, h_{k-1}$, and M_t , s.t.

$$\sum_{i=1}^{k-1} e_i \cdot h_i \text{ (mod } M_t), \quad \mathbf{e} \in \mathbb{Z}^{k-1}, \text{ } ||\mathbf{e}||_1 \leq t$$

are all distinct.

2) A code $C_r \subset S(\mathcal{M}_{k,r})$ of size M_t and with minimum Kendall's τ -distance 2t.

Construction of Systematic Codes

Recipe:

 $\overline{\text{Let }\rho_0,\rho_1,\ldots,\rho_{M_t-1}}$ be the M_t codewords in C_r .

Define $C \subset S_{k+r}$ as follows.

$$C = \{\sigma * \rho_j : \sigma \in S_k, \sum_{i=1}^{k-1} \mathbf{g}_{\sigma,i} h_i \equiv j \bmod M_t \}.$$

C is a (k + r, k) systematic t-error-correcting code.

Proof

C is a (k + r, k) systematic code.

Let $\sigma, \pi \in S_k$ and let $\rho_h, \rho_h \in C_r$ s.t. $\sigma * \rho_h, \pi * \rho_h \in C$.

If $d_K(\sigma, \pi) \geq 2t + 1$ then

$$d_{K}(\sigma * \rho_{i_{1}}, \pi * \rho_{i_{2}}) \geq 2t + 1.$$

We claim that if $1 \le d_K(\sigma, \pi) \le 2t$ then $j_1 \ne j_2$ and therefore

$$d_K(\sigma * \rho_{j_1}, \pi * \rho_{j_2}) \ge d_K(\sigma, \pi) + d_K(\rho_{j_1}, \rho_{j_2}) \ge 2t + 1.$$

Proof

$$1 \leq d_K(\sigma,\pi) \leq 2t \Rightarrow 1 \leq d_L(\mathbf{g}_{\sigma},\mathbf{g}_{\pi}) \leq 2t.$$

There exist **e**, **f** $\in \mathbb{Z}^{k-1}$, $||\mathbf{e}||_1$, $||\mathbf{f}||_1 \le t$, s.t.

$$\mathbf{g}_{\sigma} + \mathbf{e} = \mathbf{g}_{\pi} + \mathbf{f}$$
.

Assume to the contrary that $j_1 = j_2$, then

$$\sum_{i=1}^{k-1} \mathbf{g}_{\sigma,i} h_i \equiv \sum_{i=1}^{k-1} \mathbf{g}_{\rho,i} h_i \bmod M_t$$

and

$$\sum_{i=1}^{k-1} e_i h_i \equiv \sum_{i=1}^{k-1} f_i h_i \bmod M_t$$

a contradiction.

Example

Example (t=1)

Let k be integer and r = 2.

1) Let $M_1 = 2(k-1) + 1$ and let $h_i = i$, $1 \le i \le k-1$. Then the sums $\sum_{i=1}^{k-1} e_i h_i$, $||\mathbf{e}||_1 \le 1$, are all distinct modulo M_1 .

Example Continues

Example (t=1)

2) Fix $\rho \in S(\mathcal{M}_{k,2})$ and consider the codes

$$\mathcal{C}_2^{\text{e}} = \{ \gamma \in \textit{S}(\mathcal{M}_{\textit{k},2}) \; : \; \textit{d}_{\textit{K}}(\rho,\gamma) \equiv 0 \; (\text{mod 2}) \},$$

$$C_2^o = \{ \gamma \in S(\mathcal{M}_{k,2}) : d_K(\rho, \gamma) \equiv 1 \pmod{2} \}.$$

The minimum distance of both \mathcal{C}_2^e and \mathcal{C}_2^o is 2.

The size of either C_2^e or C_2^o is at least

$$\frac{|S(\mathcal{M}_{k,2})|}{2} = \frac{(k+2)!}{k! \cdot 2} = \frac{(k+2)(k+1)}{2} \ge 2(k-1)+1 = M_1.$$

Then we can construct a (k + 2, k)-systematic single-error-correcting code.

Getting Ingredient 1

Theorem (Barg & Mazumdar, 2010)

Let q be a power of a prime and $M = (q^{t+1} - 1)/(q - 1)$. Let

$$M_t = \left\{ egin{array}{ll} t(t+1)M, & t \mbox{ is odd} \\ t(t+2)M, & t \mbox{ is even} \end{array}
ight.$$

Then there exist integers $h_1, h_2, \ldots, h_{q+1}$ s.t. for all $\mathbf{e} \in \mathbb{Z}^{q+1}$, $||\mathbf{e}||_1 \le t$, the sums $\sum_{i=1}^{q+1} \mathbf{e}_i h_i$ are all distinct modulo M_t .

Getting Ingredient 2

Theorem (Sala, Gabris, & Dolecek, 2013),

Let $M = \frac{q^{t+1}-1}{q-1}$, where $q = \sum_{i=2}^{\ell} m_i - 1$ is a power of a prime. There exists a t-error-correcting code $\mathcal{C} \subset S(\mathcal{M})$ in the Kendall's τ -metric, whose size satisfies

$$|\mathcal{C}| \geq \left\{ egin{array}{l} rac{|S(\mathcal{M})|}{t(t+1)M}, & t ext{ is odd} \ rac{|S(\mathcal{M})|}{t(t+2)M}, & t ext{ is even} \end{array}
ight.$$

Example for t = 2

Example

Let k be an integer s.t. k-2 is a power of a prime and let r=3.

1) Let $M_2 = 8((k-2)^3 - 1))/(k-3) = 8((k-2)^2 + k - 1)$. There exist $h_1, h_2, \ldots, h_{k-1}$ s.t. the sums $\sum_{i=1}^{k-1} e_i h_i$, $||\mathbf{e}||_1 \le 2$, are all distinct modulo M_2 .

Example

2) There exists a single-error-correcting code $C_K \subset S(\mathcal{M}_{K,3})$ of size $|\mathcal{C}_K| \geq \frac{|S(\mathcal{M}_{k,3})|}{2^{2}+1}$. Fix $\rho \in S(\mathcal{M}_{k,3})$ and consider the

codes
$$C^{\theta} = \{x \in \mathcal{C} \text{ and } (x,y) = 0 \pmod{2}\}$$

 $\mathcal{C}_{\mathbf{3}}^{e} = \{ \gamma \in \mathcal{C}_{K} : d_{K}(\rho, \gamma) \equiv 0 \pmod{2} \},$ $\mathcal{C}_3^0 = \{ \gamma \in \mathcal{C}_K : d_K(\rho, \gamma) \equiv 1 \pmod{2} \}.$

One of these codes must be of size at least $\frac{|\mathcal{C}_{\mathcal{K}}|}{2} \geq \frac{|S(\mathcal{M}_{k,3})|}{14} = \frac{(k+3)!}{k!} = \frac{(k+3)(k+2)(k+1)}{14}.$

The minimum distance of the codes C_3^e and C_3^o is 4.

If
$$k \ge 113$$
 then $\frac{(k+3)(k+2)(k+1)}{14} \ge 8((k-2)^2 + k - 1)$ and we can construct a $(k+3,k)$ systematic double-error-correcting code.

The Number of Redundancy Symbols

Theorem

Let k be a sufficiently large integer, let $t = k^{\epsilon}$ be an integer, and let $r = \lceil \mu t \rceil$. If

$$\left\{ \begin{array}{ll} \mu > 1 + \epsilon & \textit{for} \quad 0 \leq \epsilon \leq 1 \\ \mu > 1 + \frac{1}{\epsilon} & \textit{for} \quad 1 < \epsilon. \end{array} \right.$$

There exists a (k + r, k)-systematic t-error-correcting.

Conclusion

- A construction for (k + r, k)-systematic t-error-correcting codes, was presented.
- For most values of t, the construction provides less redundancy symbols than the number of redundancy symbols of the known constructions. In particular, for a fixed t and for sufficiently large k the number of redundancy symbols is r = t + 1.
- Do there exist codes with less redundancy symbols?

The End!

Thank You!