On Codes for Flash Memory

H. Kostadinov[1], N. L. Manev[2]

[1]Institute of Mathematics and Informatics, BAS

[2]Higher School of Civil Engineering "Lyuben Karavelov"

15-21 June 2012

Introduction

- Nonvolatile memory is computer memory that maintains stored information without a power supply. With the rise of portable electronic devices like cell phones, mp3 players, digital cameras, and PDAs, nonvolatile memory is increasingly important.
- Flash memory is currently the dominant nonvolatile memory. It is cheap because it does not contain any moving parts, consumes less power, and can be electrically programmed and erased with relative ease. Reading and writing are very fast ($\sim$ 100 times faster than disk).

Flash memory consists of cells that store one or more bits by electrical charge of two or more voltage levels. The cells are organized in blocks. Flash memory has two very specific features:

- The voltage of the charge can easily be increased, but can only be decreased by an erasure operation. Only whole blocks can be erased.
- Erasures are very slow. Each block has a limited number of erase cycles it can handle. After 10,000 - 100,000 erasures, the block cannot be reliably be used.

Physical characteristics of flash memory, namely impossibility to decrease the voltage only in a cell results in asymmetry both in reading and writing processes. This leads to two main problems concerning flash memories

- Development of methods for writing in (re-programming) cells with as minimum as possible erasures (WOM codes, floating codes, flash code, rewriting codes, etc.).
- Development of suitable error correcting codes. Errors occur during the process of reading are with limited magnitude and in one dominant direction (the "read voltage level" is less than the actual one).

Example. Track 4 bits with 8 cells having 4 states

3 2 2 0 3 0 3 1	$\longrightarrow$	1 0 1 0	Change bit 3
3 2 2 0 3 1 3 1	$\longrightarrow$	1 0 0 0	Change bit 2
3 2 3 0 3 1 3 1	$\longrightarrow$	1 1 0 0	Change bit 1
3 3 3 0 3 1 3 1	$\longrightarrow$	0 1 0 0	Change bit 1
1 0 1 0 0 0 0 0	$\longrightarrow$	1 1 0 0	

Correcting asymmetric errors:

- Conventional (symmetric) error correcting codes as BCH, Reed-Solomon, LDPC codes was first used.
- Asymmetric error correcting codes were considered by Varshamov and Tenengolz in 1965.
- Multilevel flash memories renew interest in codes correcting asymmetric errors. At ACCT'02 Ahlswede et al. introduced a q-ary asymmetric channel.
- Codes over ring $\mathbb{Z}_q$ of integers modulo q specially constructed for correcting asymmetric errors have been proposed by Cassuto et al. (2008)
- Codes based on B sets described by Klove et al. (2011)

Notations and Definitions

Let $\mathcal{A}$ be an alphabet of size q and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be a codeword. $\mathbf{y} = \mathbf{x} + \mathbf{e}$, where $\mathbf{e} = (e_1, e_2, \ldots, e_n) \in \mathcal{A}^n$ is an error vector.

Definition 1. An error vector $\mathbf{e} = (e_1, e_2, \ldots, e_n)$ is called a **t-asymmetric λ-limited-magnitude error** if $\mathrm{wt}(\mathbf{e}) = |\{i : e_i \neq 0\}| \leq t$ and $0 \leq e_i \leq \lambda$, for all $i = 1, 2, \ldots, n$. A code $\mathcal{C}$ is called a **t-asymmetric λ-limited-magnitude error correcting code** if it can correct all t-asymmetric λ-limited-magnitude errors.

Definition 2. A q-ary ***integer code*** of length n with parity check matrix $\mathbf{H} \in \mathbb{Z}_q^{r \times n}$, is referred to be the subset of $\mathbb{Z}_q^n$, defined by

$$\mathcal{C}(\mathbf{H}, \mathbf{d}) = \{\mathbf{c} \in \mathbb{Z}_q^n \,|\, \mathbf{c}\mathbf{H}^T = \mathbf{d} \pmod{q}\}$$

where $\mathbf{d} \in \mathbb{Z}_q^r$.

Let $E(\lambda, n, t)$ denote the set of all possible t asymmetric λ-limited-magnitude error vectors of length n over $\mathcal{A}$.

$$|E(\lambda, n, t)| = \sum_{i=1}^{t} \binom{n}{i} \lambda^i$$

The code $\mathcal{C}(\mathbf{H}, \mathbf{d})$ is a t-asymmetric λ-limited-magnitude error correcting code if syndromes of all elements of $E(\lambda, n, t)$ are distinct, that is, all the vectors of the set

$$\left\{ \mathbf{e}\mathbf{H}^T \mid \mathbf{e} \in E(\lambda, n, t) \right\}$$

are distinct.

Hamming bound: If $\mathcal{C}(\mathbf{H}, \mathbf{d})$ is a q-ary t-asymmetric λ-limited-magnitude error correcting code then

$$q^r \geq \sum_{i=0}^{t} \binom{n}{i} \lambda^i$$

Construction of t-asymmetric λ-limited-magnitude error correcting codes over $\mathbb{Z}_q$ (T. Klove et al.).

Definition 3. Let us consider the set $B = \{b_1, b_2, \ldots, b_n\}$ of distinct positive integers. We say that B is a $B_t[\lambda](q)$ set if all the syndrome values

$$S = \left\{ \sum_{j=1}^{n} e_j b_j \mod (q) \mid (e_1, e_2, \ldots, e_n) \in E(\lambda, n, t) \right\}$$

are distinct.

Let $M_\lambda(q)$ is the maximal size of a $B_1[\lambda](q)$ set and P_o be the set of odd primes p such that $\mathrm{ord}_p(2)$ is odd.

$$q = p_1^{t_1} p_2^{t_2} \dots p_s^{t_s}, \qquad q_o = \prod_{\substack{1 \le i \le s \\ p_i \in P_o}} p_i^{t_i}$$

Theorem. If q is odd, then

$$\omega_q = \sum_{d | q_o, d>1} \frac{\varphi(d)}{\mathrm{ord}_p(2)}$$

$$M_2(q) = \frac{q-1}{2} - \frac{1}{2} \sum_{d | q_o, d>1} \frac{\varphi(d)}{\mathrm{ord}_p(2)}.$$

Perfect $B_1[2](q)$ sets exist if and only if none of the primes dividing q belongs to P_o.

New Result

Proposition. A single asymmetric 2-limited-magnitude error correctable code $\mathcal{C}$ of length n over $\mathbb{Z}_q$ has the following parity-check matrix H

- $H = (1, 3, 5, \ldots, n-1, n+3, n+5, \ldots, 2n+1)$, where $q = 2n+2$ and n is even
- $H = (1, 3, 5, \ldots, n-3, n+3, n+5, \ldots, 2n+1)$, where $q = 2n+4$ and n is odd

In the case when n is even and $p|(2n+1), p \in P_o$ the code is quasi-perfect (optimal).

Let $q = 2n+2$ and $p|(q-1)$ where $p \in P_o$. From the Theorem we have that it doesn't exist a perfect single asymmetric 2-limited-magnitude error correctable code of length n over Z_{q-1}.

Open Problems

- Codes correcting single λ-asymmetric error were $\lambda \geq 3$
- Codes correcting multiple asymmetric errors

THANK YOU FOR
YOUR ATTENTION!