On codes for multiple access adder channel with noise and feedback

Vladimir Gritsenko, Grigory Kabatiansky, Vladimir Lebedev, Alexey Maevskiy

On codes for multiple access adder channel with noise and feedback

Vladimir Gritsenko, Grigory Kabatiansky, Vladimir Lebedev, and Alexey Maevskiy *

Institute for Information Transmission Problems RAS, Moscow, Russia
gritsenko.vld@gmail.com, kaba@iitp.ru, lebedev37@mail.ru, a.maevskiy@iitp.ru

Abstract. We prove a better than previously known lower bound on the rate of codes for multiple access adder channel with noise and discuss its application to well-known coin weighing problem with not exact measurements.

1 Introduction

A q-ary code C of the length r and cardinality N over alphabet $A = \{0, 1, \ldots, q-1\} \subset \mathbb{Z}$ is called s-signature code if sums of any s or less codevectors (as vectors over the field of real numbers) are distinct, [1]. A binary s-signature code is the same as columns of a "parity" $r \times N$-matrix H of *non-adaptive* search of s or less counterfeit coins among N coins on "spring scale", see [2]. Note that the problem of detecting s counterfeit coins is elder (see [3]) than codes for multiple access adder channel.

Denote by $r(s; N)$ the minimal number of "weighings" to detect s or less counterfeit coins among N coins and by $N(s; r)$ the maximal cardinality of s-signature code of length r. Denote also by $R(s; r) = r^{-1} \log_q N(s; r)$ the rate of optimal s-signature code of length r.

The case of arbitrary number of counterfeit coins (namely, $s = N$) was solved in [4] and [5], where it was proved that

$$r(N; N) = \frac{2N}{\log_2 N}(1 + o(1)) \tag{1}$$

Investigation of the case s-fixed was initiated in [6], where the notion of B_s sequences of vectors was introduced. In particular, for the case $s = 2$ it was proved in [6] that

$$5/3 \log_2 N \le r(2; N) \le 2 \log_2 N$$

or, equivalently,

$$0.6 \ge R(2; r) \ge 0.5$$

and much later it was improved in [7] to $R(2; r) \le 0.5753$ or, equivalently, $r(2; N) \ge 1.738 \log_2 N$. Note that for the case of two counterfeit coins adaptive

* The research was supported in part by RFBR grants 13-07-00978 and 14-01-93108.

search appears to be much more effective, namely, the minimal number of weighings is not more than $1.4 \log_2 N$ [8].
It is obvious that a parity check matrix of a linear binary code correcting s errors is a "parity" matrix of non-adaptive algorithm for detecting s or less counterfeit coins and therefore binary BCH codes provide $r(s;N) \le s \log_2 N(1+o(1))$ with explicit construction of matrix H and effective algorithm of "counterfeit coins detection" (i.e. decoding) of complexity $O(sN \log N)$.
By random coding technique the bound $r(s;N) \le s \log_2 N$ was improved in [9] for all $s \ge 5$, and later it was further improved in [10] to the following form

$$r(s;N) \le r_{rand}(s;N) = \frac{2s-1}{P_s} \log_2 N(1+o(1)), \tag{2}$$

where $P_s = 2s - \log_2 C_{2s}^s$. The latter bound is better than BCH-codes for all $s \ge 3$, but for $s = 2$ binary BCH-codes, correcting two errors, give the best for today result.

In this paper we consider two generalizations of the original problems detecting counterfeit coins or codes for MAA channel. First, we consider measurements with errors or equivalently MAA channel with noise (*noisy MAAC*). Considering noise, i.e., measurements with errors, is a natural extension of group testing models with different types of measurements. For traditional group testing model it was first considered by M.Malyutov in 70-s of last century, see [11], but it became much more popular after discovery of compressed sensing in [12],[13]. For instance, the model of detecting counterfeit coins on "spring scale with noise", see [2], is equivalent to compressed sensing model with restriction that unknown real s-sparse vector $x = (x_1, \dots, x_N)$ as well as the measurement matrix H consist of only 0 and 1, see also [14]. Note, that we consider below a model of noisy MAAC which is different from model of [15] where errors in l_1 metric are considered as well as our model is different from model in [2], where a noise in s-sparse vector x was considered. Moreover, we propose constructions which are very different from known before.

As second extension of traditional problem of counterfeit coins detection we consider two-stage adaptive detection algorithms as a "middle point" between adaptive and deterministic detection algorithms. Note that it is equivalent to sending messages over MAA channel with once feedback.

2 Signature codes for noisy MAAC or toward to compressed sensing over $\mathbb{Z}$

Recall that a binary code C of length r and cardinality N is called a binary *s-signature code if sums of any s or less codevectors are distinct, where sums of vectors are taken over the field $\mathbb{R}$ of real numbers* [1].
Denote by $H = H_C$ a binary $r \times N$ matrix, which columns are all vectors of the code C. Then the property of s-signature code is equivalent to the property

that for any two different s-sparse vectors x and y their "syndroms" $S = Hx^T$ and $S' = Hy^T$ are distinct also (here matrix, vectors and their products are considered over $\mathbb{R}$). We call such matrix as a "parity" matrix of non-adaptive algorithm detecting s or less counterfeit coins on exact scale. Indeed, consider N coins which are enumerated as $1, \dots, N$ and let X be a subset of $1, \dots, N$ corresponding to counterfeit coins. We assume as usual that it is known that the weight of a counterfeit coin is 1 gram less than for a right coin. If we put a set $A \subset \{1, 2, \dots, N\}$ of coins on an exact scale and the scale shows that the total weight is on S_A grams less than for $|A|$ right coins then it means that there are S_A counterfeit coins among chosen set A. Equivalently, $S_A = |X \cap A| = (x, \chi(A))$, where the scalar product is taken over $\mathbb{R}$, $\chi(A)$ denotes the characteristic vector of the set A and $x = \chi(X)$. Let sets $A_1, \dots, A_r$ be chosen in such way that $\chi(A_i) = h_i$. Hence, if we put sets $A_1, \dots, A_r$ on the scale then we know (x, h_i) for all i, i.e., we know Hx^T, and therefore can uniquely recover x if the Hamming weight $wt(x) \le s$.

Consider the noisy MAAC. We assume that for any s codewords $c^{(1)}, \dots, c^{(s)}$ of a signature code C being transmitted through the noisy MAAC the corresponding output vector $z = (z_1, \dots, z_r)$ may differ from the vector $\sum_{i=1}^{s} c^{(i)}$ in at most l coordinates, i.e.,

$$d(z, \sum_{i=1}^{s} c^{(i)}) \le l, \tag{3}$$

where $d(a, b) = |\{i : a_i \ne b_i\}$ is the Hamming distance between vectors a and b (and $wt(a) = d(a, 0)$ is the Hamming weight of a). For detecting counterfeit coins it is equivalent to the assumption that some but not more than l of r measurements have been distorted by noise. Therefore we use the following

Definition 1. *We shall say that a code C is (s, l)-signature code, or s-signature code correcting l errors, if*

$$d(\sum_{c \in I \subset C} c, \sum_{c \in J \subset C} c) \ge 2l + 1, \tag{4}$$

for any two different subsets I and J both cardinality not more than s.

Let again $H = H_C$ be an $r \times N$-matrix which columns are codewords of (s, l)-signature code code C, where $N = |C|$. Then (4) is equivalent to the property that for any $\hat{S}$ the following equation

$$\hat{S} = Hx^T + e, \tag{5}$$

has not more than a single solution among pairs $\{x, e\}$ such that $wt(x) \le s$ and $wt(e) \le l$. Equivalently, for any "distorted syndrom" $\hat{S}$ there is at most one binary s-sparse vector x such that $d(\hat{S}, Hx^T) \le l$. A very similar problem was considered in [16] as a discrete variant of compressed sensing problem. Later classes of optimal and asymptotically optimal "codes" were constructed in [17].

Let us describe that construction and apply it for the considered problem.
Denote by $N(s,l;r)$ the maximal cardinality of (s,l)-signature code of length r and by $r(s,l;N)$ the minimal number of "weighings" to detect s or less counterfeit coins among N coins in presence of at most l wrong "weighings".
Let a binary $\tilde{r} \times N$ matrix $\tilde{H}$ be a parity-check matrix of an $(N, N-\tilde{r})$-code over $\mathbb{F}_2$, correcting s errors, i.e. any $2s$ columns of $\tilde{H}$ are linear independent. And let G be a generator matrix of an $(r,\tilde{r})$-code over $\mathbb{F}_2$ of length r, correcting l errors. Let matrix H consists of columns $h_1,\ldots,h_N$, where

$$h_j^T = \tilde{h}_j^T G \tag{6}$$

and transposition ^{T} means, that vectors h_j and $\tilde{h}_j$ are considered in (6) as row vectors, i.e.

$$H = G^T \tilde{H} \tag{7}$$

Saying in words, we encode columns of parity-check matrix $\tilde{H}$, which is already capable to correct s errors, by a linear binary code, correcting l errors, in order to restore correctly syndrom of $\tilde{H}$. It was proved [17] that the constructed matrix has a desirable property even in a bit stronger sense, namely, over the binary field $\mathbb{F}_2$. Since sum of any s (or less) columns of H differs by $\mod 2$ from sum of any other s columns of H in at least l positions the same is true for sums as real numbers, hence we have the desirable property.
Let us choose both constituent codes as binary BCH-codes, for which it is known that the redundancy $r_{BCH}(t,n)$ of binary BCH t error-correcting code of length n is at most $t\lceil \log_2(n+1)\rceil$, see [18]. Hence this construction provides $r \times N$ parity matrices for detecting s counterfeit coins in presence of at most l wrong measurements, or an s-signature code of length r and cardinality N for the noisy MAAC with at most l errors during the transmission, with

$$r \le (s \log_2 N + l \log_2 \log_2 N)(1+o(1)) \tag{8}$$

It is clear that this construction does not use in full the property that addition of vectors is over real numbers field. The next construction exploits this property more effective.

We will use the same idea as above, namely, we construct a parity $r \times N$ binary matrix H consisting of two submatrices - upper $r_1 \times N$ matrix H_1 and lower $r_2 \times N$ matrix H_2. Let matrix H_1 be capable to detect up to s counterfeit coins among N, i.e. all sums of s or less column vectors of H_1 are distinct. Let p be a minimal prime number such that $p > s$ and let C be a systematic linear (r, r_1)-code of length r over the field $\mathbb{Z}_p$ with r_1 information symbols and $r_2 = r - r_1$ parity symbols, which is capable to correct l errors (over $\mathbb{Z}_p$).
Encode columns of the matrix H_1 by the code C. We cannot use directly these vectors as columns of H since the result of encoding gives vectors which r_2 parity symbols aren't binary vectors, but belong to $\mathbb{Z}_p$. Instead we replace each parity symbol of these vectors by the corresponding m-bit column based on the following simple mapping $\Psi(a) : \mathbb{Z}_p \to \{0,1\}^m$, where $m = \lceil \log_2(p+1)\rceil$ and

$\Psi(a) = (a_0, \ldots, a_{m-1})$ is the binary representation of integer number a.
Now consider decoding procedure. Note that from sum

$$\Psi(a) + \Psi(b) = (a_0 + b_0, \ldots, a_{m-1} + b_{m-1}),$$

taken over $\mathbb{R}$, we can recover $a + b$ as $a + b = \sum_{i=0}^{m-1}(a_i + b_i)2^i$. Of course, the same is true for sums of s or less elements. Hence from the sum of parity parts of column vectors of H, namely,

$$(\Psi(a^{(1)}), \ldots, \Psi(a^{(r_2)})) + (\Psi(b^{(1)}), \ldots, \Psi(b^{(r_2)})) + (\Psi(c^{(1)}), \ldots, \Psi(c^{(r_2)})) + \ldots$$

we can recover the corresponding vector

$$(a^{(1)} + b^{(1)} + c^{(1)} + \ldots, \ldots, a^{(r_2)} + b^{(r_2)} + c^{(r_2)} + \ldots)$$

and then we take residues of these coordinates by $\mod p$. Therefore in presence of l or less errors we will correct them since columns of H are taken from linear code and therefore their sum is also a codevector. Hence decoding gives us a correct sum vector of s columns vectors and we can recover x by the property that the matrix H_1 is capable to detect up to s counterfeit coins.
The redundancy r_p of p-ary BCH code of length n, correcting l errors, is

$$r_p \leq 1 + (2l - 1 - \lceil \frac{2l-1}{p} \rceil)\lceil \log_p(n+1) \rceil \tag{9}$$

Note, that there are codes with asymptotically better redundancy for $d > p$, see[19], but the corresponding gain is not very large. Therefore for the second term r_2 of redundancy we have that $r_2 \leq mr_p$ and therefore

$$r(s, l; N) \leq r(s, N) + 2l\lceil \log_2(p+1) \rceil \frac{(p-1)\log_2 \log_2 N}{p \log_2 p}(1 + o(1)) \tag{10}$$

Rather straightforward calculations for random codes with expurgation give the following lower bound

$$r(s, l; N) \leq r_{rand}(s, N) + 2l \log_2 \log_2 N(1 + o(1)) \tag{11}$$

It is clear from (10) and (11) that "semi-constructive" codes of (10) are better than random codes (even with expurgation) when $p \log_2 p > (p-1)\lceil \log_2(p+1) \rceil$, for instance, for $p = 7$ (and $s \leq 6$).

3 Two-stage adaptive detection algorithms

Consider the case of two counterfeit coins. Let $2^{m-1} \leq N < 2^m$. For the first stage we use $m \times N$ binary matrix H_1 which columns are distinct binary m-tuples. Let x be a binary vector of weight 2 corresponding to the positions of two counterfeit coins. Vector $S = H_1 x^T$ is the result of measurements at first stage. It is clear that different vectors x of weight 2 such that $H_1 x^T = S$ have

nonintersecting supports, and there are $2^{W_1(S)-1}$ such vectors, where $W_1(S)$ is the number of coordinates in S which are equal to 1. Let us enumerate these vectors $\{x\}$ by ternary vectors of length $L = \lceil \log_3 2^{W_1(S)-1} \rceil$ and let ternary vector $c_{i,j} = (c_1, \ldots, c_L)$ corresponds to the vector x which has 1 on positions i and j. Now we form parity matrix H_2 for the second stage by choosing as i-th and j-th columns of H_2 binary vectors h_i and h_j such than $h_i + h_j = c_{i,j}$. For example, if $c_{i,j} = (0, 2, 1, 2, 1)$ then we choose $h_i = (0, 1, 0, 1, 0)$ and $h_j = (0, 1, 1, 1, 1)$. Then $H_2 x^T = c_{i,j}$ and different vectors x of weight 2 such that $H_1 x^T = S$ are distinguished by the second stage. The total redundancy

$$r_{two} \le \lceil \log_2(N+1) \rceil + \lceil \log_3 2^{m-1} \rceil = (1 + \log_3 2) \log_2 N(1 + o(1)) \qquad (12)$$

and hence the redundancy of two-stage search is asymptotically $r_{two} \le 1.631 \log_2 N$. It is better than any deterministic detection because $r(2; N) \ge 1.738 \log_2 N$, but worse than adaptive search with $1.4 \log_2 N$ weighings [8] .

Consider the case of three counterfeit coins. Let for simplicity of notations $N = 2^m - 1$. For the first stage we use parity-check $2m \times N$ matrix H_1 of binary BCH-code, correcting two errors, plus extra row consisting of all ones. Because of this row after first stage we know exactly the number of counterfeit coins and if this number is less than 3 then we can find them from "syndrom" $S = H_1 x^T$. Let x be a binary vector of weight 3 which has ones on positions i, j, k, and set $\{i, j, k\}$ is the support set of x. And let y be another binary vector of weight 3 which has ones on positions i', j', k' and $H_1 x^T = H_1 y^T$. Then these two vectors have disjoint supports. Indeed, let $k = k'$, then $h_1^{(i)} + h_1^{(j)} + h_1^{(k)} = h_1^{(i')} + h_1^{(j')} + h_1^{(k)}$ and hence $h_1^{(i)} + h_1^{(j)} = h_1^{(i')} + h_1^{(j')}$ but all pairwise sums of columns of H_1 are different (since the code correct two errors). Now we have similar problem as above - we enumerate all vectors $x : wt(x) = 3, H_1 x^T = S$ by quaternary vectors of length $L \le \lceil \log_4(N/3) \rceil$ and repeat all aforementioned arguments by replacing pairs on triples. The total redundancy of this two-stage algorithm doesn't exceed $2.5 \log_2 N(1 + o(1))$ what is much better than random codes having $r \approx 2.97 \log_2 N$, see (2).
It is an interesting open question to establish lower bound on the redundancy of two-stage detection.

References

1. S.C.Chang and E.J.Weldon,"Coding for T-user multiple access channels", IEEE Trans. Inform. Theory, v. 25 (6), pp. 684-691, 1979.
2. N.H.Bshouty, H.Mazzawi, "Algorithms for the Coin Weighing Problem with the Presence of Noise", Electronic Colloquim on Computational Complexity, Rep. 124, 2011.
3. P.Erdos and A.Renyi, "On two problems of information theory", Publ.Math. Inst. Hung. Acad.Sci, V.8, pp. 241-254, 1963.
4. B. Lindstrom, "On a combinatorial detection problem,I", Publ.Math. Inst. Hung. Acad.Sci, V.9, pp. 195-207, 1964.

5. D.Cantor, W.Mills " Determining a subset from a certain combinatorial properties", Canadian J. Math., V.18, pp. 42-48, 1966.
6. B. Lindstrom, "On B_2-sequences of vectors", J. Number Theory, V.4, pp. 261-265, 1972.
7. G.Cohen, S. Litsyn, G.Zemor,"Binary B_2-sequences: a new upper bound", Journal of Combinatorial Theory, Ser. A, vol. 94, No.1, pp. 152-155, 2001.
8. L. Gargano, V.Montuori, G.Setaro and U.Vaccaro, "An improved algorithm for quantative group testing", Discrete Applied Mathematics, v.36, pp. 299–306, 1992.
9. A.A.Dyachkov, V.V.Rykov, "On a coding model for a multiple-access adder channel", Problems of Information Transmission, v. 17 (2), pp. 94-104, 1981.
10. G.S. Poltyrev, "Improved upper bound on the probability of decoding error for codes of complex structure", Problems of Information Transmission, v. 23 (4), pp. 251-262, 1987.
11. M. Malyutov, "Search for Sparse Active Inputs: a Review", Information Theory, Combinatorics and Search Theory, LNCS 7777, pp. 609-647, 2014.
12. D. L. Donoho, "Compressed sensing", *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.
13. E. J. Candes, T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies? ", *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 5406 - 5425, 2006.
14. C.L. Chan, S.Jaggi, V.Saligrama and S.Agnihotri, "Non-adaptive Groupt Testing: Explicit Bounds and Novel Algorithms", *IEEE Trans. Inform. Theory*, vol. 60, no. 5, pp. 3019 - 3035, 2014.
15. S.Lu, J.Cheng,W.Hou, and Y.Watanabe, "Generalized Construction of Signature Code for Multiple-Access Adder Channel", in *Proc. 2013 IEEE Int. Symp. Informatiomn Theory*, pp. 1655-1659, 2013.
16. G.Kabatiansky, S.Vladuts,"What to do if syndromes are corrupte also", in *Proc. Int.Workshop Optimal Codes*, Albena, Bulgaria, 2013.
17. G.Kabatiansky, V. Lomakov, S.Vladuts, " On error-correction under distortions in channel and syndrom", Problems of Information Transmission, v. 51, 2015 (in print)
18. F.J. MacWilliams, N.J.A. Sloane, The Theory of Error-Correcting Codes (North-Holland Mathematical Library)
19. S. Yekhanin and I. Dumer, "Long nonbinary codes exceeding the Gilbert-Varshamov bound for any fixed distance", *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2357 - 2362, 2004.