DIMACS Technical Report 2002-09

Pairs of Codes with Prescribed Hamming Distances and Coincidences

by

Vince Grolmusz [1]

[1]Department of Computer Science, Eötvös University, Budapest, Address: Pázmány P. stny. 1/C, H-1117, Budapest, HUNGARY; E-mail: grolmusz@cs.elte.hu

ABSTRACT

The main problem of coding theory is to construct codes with large Hamming-distances between the code-words. In this work we describe a fast algorithm for generating pairs of q-ary codes with prescribed pairwise Hamming-distances and coincidences (for a letter $s \in \{0, 1, \ldots, q-1\}$, the number of s-coincidences between codewords a and b is the number of letters s in the same positions both in a and b). The method is a generalization of a method for constructing set-systems with prescribed intersection sizes (V. Grolmusz: Constructing Set-Systems with Prescribed Intersection Sizes, DIMACS Technical Report No. 2001-03), where only the case $q = 2$ and $s = 1$ was examined.

Keywords: multi-linear polynomials, codes, Hamming-distance, code-generation

1 Introduction

In the theory of codes one of the main questions is to find dense codes with large minimum Hamming-distance. Here we address the problem of generating codes with prescibed pairwise and k-wise Hamming-distances.

First we consider a generalization of the set-intersection for q-ary codes: coincidences. Let $s \in \{0, 1, 2, \ldots, q-1\}$, then the number of s-coincidences of two code-words a and $b \in \{0, 1, 2, \ldots, q-1\}^n$ are the number of coordinates i such that $a_i = b_i = s$, that is, the number of letters s in the same positions in a and b.

We describe a construction in which we apply multi-variate polynomials to codes. Choosing a polynomial f and initial codes A and B we apply f to the codes, getting $f(A)$ and $f(B)$. The most remarkable property of the $A \to f(A)$ mapping is that the s-coincidence of words $f(A)$ and $f(B)$ essentially can be got with applying f to the s-coincidences of the original codes A and B (see Theorems 10 and 13 for the exact statements). This fact yields a tool for manipulating codes in order to get prescribed coincidences. Note, that we can even allow different polynomials for different letters $s \in \{0, 1, \ldots, q-1\}$ in this construction(see Theorem 15).

The construction, presented in this paper is a generalization of a similar construction for set systems appeared in [Gro01].

We also prove that the Hamming-distances between the new code-words are a simple function of the Hamming-distances between the old code-words and f (see Corollaries 22 and 23.) This fact can be used for generating codes with prescribed Hamming distances, if we have a proper polynomial f. Over the integers, our method works only for polynomials with non-negative integer coefficients, which fact forbids most of the interpolating polynomials, containing negative coefficients, but if we aim to set the Hamming-distances only modulo a positive integer, the problem, caused by the negative coefficients, disappears.

In paper [GS01] we proved the k-wise analogue of the classic Delsarte-theorem for codes. Here we define the k-wise coincidences between k code-words (Definition 16), and prove that the k-wise coincidences of codes $f(A_i)$ $i = 1, 2, \ldots, k$ are the f-function of the k-wise coincidences of the codes A_i, $i = 1, 2, \ldots, k$ (Theorem 18). As a corollary of this result, we got a construction for codes with their k-wise Hamming-distances being a simple function of the k-wise coincidences of the original ones (Theorem 25).

2 Preliminaries

2.1 Codes, Hamming-distances and Coincidences

Definition 1 *Set C is a q-ary code of length n if*

$$C \subset \{0, 1, \ldots, q-1\}^n,$$

i.e., it is a set of length-n words formed from symbols $\{0, 1, \ldots, q-1\}$. For any code C, we fix arbitrarily an order of its elements (called code-words), so we can write $C = \{c^1, c^2, \ldots, c^\ell\}$ for $\ell = |C|$.

Definition 2 *Let $A = \{a^1, a^2, \ldots, a^\ell\}$ be a code. Then the matrix of A, denoted by $M(A)$, is an $n \times \ell$ matrix, with column j equal to the code-word a^j, for $j = 1, 2, \ldots, \ell$.*

Let us remark that code A determines $M(A)$ since we have fixed an order of the elements of A in Definition 1.

Definition 3 *Let $A = \{a^1, a^2, \ldots, a^k\}$ and $B = \{b^1, b^2, \ldots, b^\ell\}$ be two q-ary codes. Then the Hamming-distance-matrix of codes A and B, denoted by*

$$H(A, B) = \{h_{ij}\},$$

is a $k \times \ell$ integer matrix, where h_{ij} is the Hamming-distance of words a^i and b^j.

Example 4 *Let*

$$M(A) = \begin{pmatrix} 3 & 1 & 0 \\ 5 & 1 & 2 \\ 2 & 1 & 3 \end{pmatrix}, \; M(B) = \begin{pmatrix} 3 & 1 & 1 & 4 \\ 5 & 0 & 2 & 5 \\ 2 & 1 & 3 & 6 \end{pmatrix}.$$

Then

$$H(A, B) = \begin{pmatrix} 0 & 3 & 3 & 2 \\ 3 & 1 & 2 & 3 \\ 3 & 3 & 1 & 3 \end{pmatrix}.$$

We also need to define a sort of complement of Hamming distance of code-words, or sequences: the number of the same letters in the same positions.

Definition 5 *Let $A = \{a^1, a^2, \ldots, a^k\}$ and $B = \{b^1, b^2, \ldots, b^\ell\}$ be two q-ary codes, and let $s \in \{0, 1, 2, \ldots, q-1\}$. Then the s-coincidence-matrix of codes A and B is a*

$$C_s(A, B) = \{c_{ij}\}$$

$k \times \ell$ integer matrix, where $c_{ij} = c_s(a^i, b^j)$ is the number of coordinates in code-words a^i and b^j, which are both equal to s:

$$c_s(a^i, b^j) = |\{k : 1 \leq k \leq n, a^i_k = b^j_k = s\}|.$$

Similarly, the s-coincidence-position matrix $CP_s(A, B) = \{w_{ij}\}$ is a $k \times \ell$ matrix with length-n vector-elements, where w_{ij} is a 0-1-vector of length n, defined as:

$$w_{ij} = (u_1, u_2, \ldots, u_n), \text{ where } u_t = 1 \text{ iff } a^i_t = b^j_t = s.$$

Example 6 *With the codes A and B of Example 4,*

$$C_1(A, B) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The following lemma describes an easy relation between the Hamming-distance- and the coincidence-matrices.

Lemma 7 *Let J denote the all-1 matrix, and let A and B be two length-n q-ary codes. Then*

$$H(A, B) = nJ - \sum_{i=0}^{q-1} C_i(A, B).$$

□

Definition 8 *Let R and S be two rings, and let $X \in R^{u \times v}$ be a matrix, with elements $\{x_{ij}\}$, and let f be an $f : R \to S$ function. Then we define matrix $f[X] \in S^{u \times v}$ to be a matrix with entries $\{f(x_{ij})\}$.*

Definition 9 *Let $f(x_1, x_2, \ldots, x_n) = \sum_{I \subset \{1,2,\ldots,n\}} a_I x_I$ be a multi-linear polynomial over integers. where $x_I = \prod_{i \in I} x_i$. Let us define its weight as $w(f) = |\{a_I : a_I \neq 0\}|$, and its L_1 norm as $\mathrm{L}_1(f) = \sum_{I \subset \{1,2,\ldots,n\}} |a_I|$.*

3 Generating code-pairs

Our main contribution is a method for generating codes with a prescribed coincidence-matrix with a possibly small code-length. For $q \geq 3$, it is not difficult to give a pair of q-ary codes A and B of length $\Omega(|A||B|)$ with prescribed coincidence-matrices.

Indeed, let us divide each code-word $a \in A$ into $|B|$ segments, and each segments into q sub-segments. The $|B|$ segments in each a will correspond to the elements of B, and the segment, corresponding to $b \in B$ will provide the coincidences with b, and the sub-segment s will provide the prescribed number of s-coincidences between a and b, for any $s \in \{0, 1, \ldots, q-1\}$. If we need r s-coincidences, then the subsegment in

a is given: $\overbrace{s, s, s, s, \ldots, s}^{r}$, in

b is given: $\overbrace{s, s, s, s, \ldots, s}^{r}$, in

$a' \neq a$ is given: $\overbrace{\delta, \delta, \delta, \delta, \ldots, \delta}^{r}$, in

$b' \neq b$ is given: $\overbrace{\sigma, \sigma, \sigma, \sigma, \ldots, \sigma}^{r}$,

for some pairwise different $s, \sigma, \delta \in \{0, 1, \ldots, q-1\}$.

Clearly, this sub-segment will not give rise to any other coincidences just to the s-coincidences and only between a and b.

However, our method yields codes with length not depending on $|A|$, but only on the L_1-norm of a polynomial, which describes the elements of the coincidence-matrix. That means, that if we have a polynomial with small L_1 norm, then our codewords will be short. More exactly, we prove:

Theorem 10 *Let $q > 2$, and let A and B length-n q-ary codes, and let f be an n-variable multi-linear polynomial with non-negative integer coefficients. Then there exist q-ary codes A' and B', such that*

$$C_s(A', B') = f[CP_s(A, B)],$$

a, for $s = 0, \ldots, q-1$, and for $q \geq 4$, the length of codes A' and B' is $2\mathrm{L}_1(f)$;

b, for $s = 0, \ldots, q-1$, and for $q = 3$, where the length of codes A' and B' is $3\mathrm{L}_1(f)$;

c, for $s = 1, 2, \ldots, q-1$ and for $q \geq 2$, where the length of codes A' and B' is $\mathrm{L}_1(f)$.

Moreover, codes A' and B' can be computed from codes A, B, and polynomial f in time $O(L_1(f)\deg(f)(|A| + |B|))$.

For the proof we need to develop some machinery.

3.1 The main lemma

Our main contribution in this work is a method for constructing new codes from given ones, with the values of the coincidence-matrices of the new codes are equal to the function f of the entries of the old coincidence-matrices. Our construction works over the positive integers with f's with non-negative integer coefficients, or over the ring of mod m integers, where the non-negativity assumption is not needed.

Let f be a multi-linear polynomial with n variables and either with non-negative integer coefficients or with coefficients from a ring of mod m integers:

$$f(x_1, x_2, \ldots, x_n) = \sum_{I \subset \{1,2,\ldots,n\}} a_I x_I$$

and let A be a q-ary code of length n. Let $\sigma \in \{0, 1, \ldots, q-1\}$.

Code-generation algorithm Let $\sigma \in \{0, 1, \ldots, q-1\}$, and $\Gamma \subset \{0, 1, \ldots, q-1\}$, $\sigma \notin \Gamma$. Here we describe an algorithm for constructing code-matrix $M(f_\sigma^\Gamma(A))$ with σ-filling and Γ-prohibition:
Consider the matrix $M(A) = \{b_{ij}\}$, and let us correspond its row i to symbol x_i, for $i = 1, 2, \ldots, n$. We will construct a matrix $M(f_\sigma^\Gamma(A))$ with $\mathrm{L}_1(f)$ rows and $k = |A|$ columns as follows. The rows of $M(f_\sigma^\Gamma(A))$ will correspond to the monomials x_I of f for $I \subset \{1, 2, \ldots, n\}$, with $a_I \neq 0$; we will take every monomial x_I a_I times, that is, term $a_I x_I$ will be corresponded to a_I identical rows of matrix $M(f_\Gamma^\delta(A))$. Consequently, matrix $M(f_\sigma^\Gamma(A))$ has $\mathrm{L}_1(f)$ rows. Now we specify the entries in the rows of this matrix. Consider a row, corresponding to a monomial x_I. Let $1 \leq j \leq m$, and consider the jth entry of this row. Let us define this entry to be $y \in \{0, 1, \ldots, q-1\} - \Gamma$ if and only if all the entries $b_{ij}, i \in I$ are equal to y, and let this entry be equal to σ (the filling-element) otherwise.

The code, corresponding to the columns of matrix $M(f_\sigma^\Gamma(A))$ is called $f_\sigma^\Gamma(A)$, if $\Gamma = \emptyset$, then we write $f_\sigma(A)$

Note, that code $f_\sigma^\Gamma(A)$ does not contain any word with any letter from Γ.

Example 11 *Let $f(x_1, x_2, x_3, x_4) = x_1 + x_2 + 2x_3x_4 + x_2x_3x_4$, and let*

$$M(A) = \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 2 & 0 & 2 \\ 2 & 0 & 1 \end{pmatrix}.$$

Then

$$M(f_2(A)) = \begin{array}{l} x_1 \\ x_2 \\ x_3x_4 \\ x_3x_4 \\ x_2x_3x_4 \end{array} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 0 & 2 \\ 2 & 0 & 2 \\ 2 & 0 & 2 \end{pmatrix}, \qquad M(f_2^{\{1\}}(A)) = \begin{array}{l} x_1 \\ x_2 \\ x_3x_4 \\ x_3x_4 \\ x_2x_3x_4 \end{array} \begin{pmatrix} 0 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 0 & 2 \\ 2 & 0 & 2 \end{pmatrix}.$$

Our main lemma describes the most important properties of this code-construction.

Lemma 12 *Let A and B length-n q-ary codes, and let f be a multi-linear polynomial with n variables and either with non-negative integer coefficients or with integer coefficients from the modulo m ring of integers, and let s,σ be different elements of set $\{0,1,\ldots,q-1\}$, and let $\Gamma\subset\{0,1,\ldots,q-1\}$, such that neither s nor σ is in Γ . Then*

a,

$$C_s(f_\sigma^\Gamma(A),f_\sigma^\Gamma(B))=f[CP_s(A,B)];$$

b, If $\Gamma=\{\delta\}$,

$$C_s(f_\sigma^{\{\delta\}}(A),f_\delta^{\{\sigma\}}(B))=f[CP_s(A,B)];$$

c,

$$C_\delta(f_\sigma^{\{\delta\}}(A),f_\delta^{\{\sigma\}}(B))=C_\sigma(f_\sigma^{\{\delta\}}(A),f_\delta^{\{\sigma\}}(B))=0.$$

Proof: Recall, that $A=\{a^1,a^2,\ldots,a^k\}$ and $B=\{b^1,b^2,\ldots,b^\ell\}$.

For proving statement (a), we assume that the element in row i and column j of $CP_s(A,B)$ is $u=(u_1,u_2,\ldots,u_n)\in\{0,1\}^n$. That means, that $a_t^i=b_t^j=s$ exactly when $u_t=1$. Now, $f(u)$ is equal to the number (counting the multiplicities) of monomials x_I in f, such that for all $t\in I$ $u_t=1$. However, this happens exactly when for all $t\in I$ $a_t^i=b_t^j=s$, that is, the coordinates, corresponding to the monomial x_I of the word i of $f(A)$ and word j of $f(B)$ are both s; that means that for this monomial value 1 is contributed to the element in row i and column j of matrix $C_s(f_\sigma^\Gamma(A),f_\sigma^\Gamma(B))$. Note, that since the filling element σ differs from s, only the aforementioned contributions will be counted to matrix $C_s(f_\sigma^\Gamma(A),f_\sigma^\Gamma(B))$.

Proofs of parts (b) and (c) are obvious. $\Box$

Proof of Theorem 10:

The easiest case is (c), so we prove that first.
Part (c): Let the filling-element $\sigma=0$. Then, by Lemma 12 (a), for any $s\neq 0$:

$$C_s(f_\sigma(A),f_\sigma((B))=f[CP_s(A,B)];$$

so $A'=f\sigma(A),B'=f\sigma(B)$ suffice.
Parts (a) & (b): We prove that the length of the codes is

$$\left\lceil\frac{q}{q-2}\right\rceil \mathrm{L}_1(f),$$

and this implies both (a) and (b). First, partition set $\{0,1,\ldots,q-1\}$ into

$$\left\lceil \frac{q}{q-2} \right\rceil$$

classes of size at most $q-2$, let these classes be Ξ_i for $i=1,\ldots,t$ $t=2$ or $t=3$. For some i, take Ξ_i, and $\alpha,\beta\in\{0,1,\ldots,q-1\}-\Xi_i$, $\alpha\neq\beta$. Let

$$A^{(i)'} = f_\beta^{\{\alpha\}}(A),\ B^{(i)'} = f_\alpha^{\{\beta\}}(B).$$

Certainly, by part (b) of Lemma 12, for any $s\in\Xi_i$ the requirements are satisfied. At the end, we get the final A' by concatenating the corresponding code-words of codes $A^{(i)'}$ for $i=1,\ldots,t$ and the final B' by concatenating the corresponding code-words of codes $B^{(i)'}$ for $i=1,\ldots,t$.$\square$

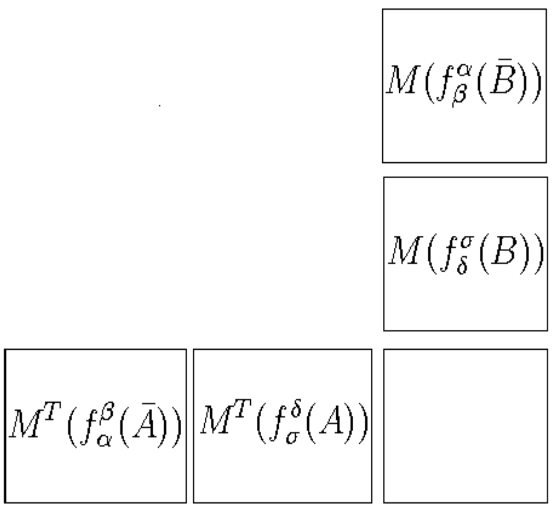

Figure 1: The proof in $q\geq 4$ case.

4 Further results for code-pairs

First we give a variant of the Theorem 10, where $A=B$ and $A'=B'$:

Theorem 13 *Let A be length-n q-ary code, and let f be an n-variable multi-linear polynomial with non-negative integer coefficients. Then there exist explicitly constructible q-ary code A', such that*

$$C_s(A',A') = f[CP_s(A,A)],$$

$s = 1, 2, \dots, q-1$ and for $q \geq 2$, where the length of code A' is $\mathrm{L}_1(f)$. Moreover, code A' can be computed from codes A and polynomial f in time $O(L_1(f)\deg(f)|A|)$.

Proof of Theorem 13: Let $A' = f(A)$, and choose the filling-element $\sigma = 0$, and apply Lemma 12 (a) with no prohibition (i.e., no δ). We get that for all $s \neq 0$ $C_s(A', A') = f[CP_s(A, A)]$, and we are done.□

If f is a symmetric polynomial over a prime-element field, then it can be written as a single-variable polynomial, and we have the following Corollary:

Corollary 14 *Let A and B be length-n q-ary codes, let $q \geq 4$, and for a prime r, let F_r denote the r-element field, and let $f : F_r \to F_r$ be a polynomial. Then we can construct codes A', A'' and B' with a fast polynomial-time algorithm, such that*

$$C_i(A', B') = f[C_i(A, B)], \quad i = 0, 1, \dots, q-1,$$

and

$$C_i(A'', A'') = f[C_i(A, A)], \quad i = 1, \dots, q-1,$$

and where the length of codes A' and B' is at most $2r\sum_{j=0}^{q-1}\binom{n}{j}$, and the length of A'' is at most $r\sum_{j=0}^{q-1}\binom{n}{j}$.

In other words, we can prescribe the values of the coincidency-matrices modulo r.

Proof of Corollary 14: The statement follows from Theorem 10 part (a), using that single-variable functions f over F_r can be interpreted as a symmetric n-variable function $g(z_1, z_2, \dots, z_n) = f(z_1 + z_2 + \cdots + z_n)$. Since f has degree at most $r-1$, g may have at most $\sum_{j=0}^{r-1}\binom{n}{j}$ monomials, consenquently, its L_1-norm is at most $(r-1)\sum_{j=0}^{r-1}\binom{n}{j}$, if the coefficients are represented by the elements of set $\{0, 1, \dots, r-1\}$.□

We can allow different polynomials for different coincidence-matrices, or even different moduli for different polynomials in the following Theorem:

Theorem 15 *Let A and B length-n q-ary codes, let $q \geq 3$, and let $f_{(0)}, f_{(1)}, \dots, f_{(q-1)}$ be n-variable multi-linear polynomials with non-negative integer coefficients, and let $m_0, m_1, \dots, m_{q-1}$ be positive integers, all greater than 1. Then there exist explicitly constructible q-ary codes A' and B' of length*

$$\sum_{i=0}^{q-1} \mathrm{L}_1(f_{(i)}),$$

such that

a, $C_s(A', B') = f_{(s)}[CP_s(A, B)]$, for $s = 0, 1, \dots, q-1$;

b, $C_s(A', B') = f_{(s)}[CP_s(A, B)] \pmod{m_s}$ *for* $s = 0, 1, \ldots, q-1$.

Proof of Theorem 15: We use a similar procedure as in the proof of Part (b) of Theorem 10.

The code-words of A' and B' will be divided to segments, one segment for each $s \in \{0, 1, \ldots, q-1\}$. The length of segment s is at most $L_1(f_{(s)})$. The segment s of the codes is defined as $g_\delta^\sigma(A)$ in A' and $g_\sigma^\delta(B)$, where $g = f_{(s)}$, and $\delta \neq \sigma$ are elements of $\{0, 1, \ldots, q-1\}$, distinct from s. By Lemma 12 both (a) and (b) are satisfied for this s, and by concatenating the segments for each s we get codes A' and B' of length at most

$$\sum_{i=0}^{q-1} \mathrm{L}_1(f_{(i)}),$$

satisfying (a) and (b) for all $s \in \{0, 1, \ldots, q-1\}$. □

5 Generalizations for k-wise coincidences

Our results can be generalized to k-wise coincidence-matrices as well:

Definition 16 *Let $A_1, A_2, \ldots, A_k$ be length-n q-ary codes. Let $s \in \{0, 1, 2, \ldots, q-1\}$. Then the k-wise s-coincidence (k-dimensional) matrix of codes A_ℓ ($\ell = 1, 2, \ldots, k$) is a*

$$C_s(A_1, A_2, \ldots, A_k) = \{c_{i_1, i_2, \ldots, i_k}\}$$

k-dimensional integer matrix, where entry $c_{i_1, i_2, \ldots, i_k}$ is defined to be the number of co-ordinates in the i_1st code word of A_1, in the i_2nd code-word of A_2, ..., in the i_kth code-word of A_k which are all equal to s. Similarly, the k-wise s-coincidence-position matrix $CP_s(A_1, A_2, \ldots, A_k) = \{w_{i_1, i_2, \ldots, i_k}\}$ is a k-dimensional matrix, where $w_{i_1, i_2, \ldots, i_k}$ is a length-n 0-1-vector, defined to be 1 if at that position all the corresponding code-words contain s.

Let us remark that the k-wise coincidence-matrices contain all the k'-wise coincidences for $1 \leq k' \leq k$. Indeed, $c_{i_1, i_2, \ldots, i_k}$ gives a $k' < k$-wise coincidence, when the number of different indeces in set $\{i_1, i_2, \ldots, i_k\}$ is exactly k'. We also need a k-wise version of one of our definitions:

Definition 17 *Let R and S be two rings, and let $X \in R^{u_1 \times u_2 \times \cdots u_k}$ be a k-dimensional matrix, with elements $\{x_{i_1, i_2, \ldots, i_k}\}$, and let f be an $f : R \to S$ function. Then we define k-dimensional matrix $f[X] \in S^{u_1 \times u_2 \times \cdots u_k}$ to be a matrix with entries $\{f(x_{i_1, i_2, \ldots, i_k})\}$.*

Consequently, the following theorem states that choosing a polynomial f will ensure that all the k'-wise coincidences (for $1 \leq k' \leq k$) are set according to f.

Theorem 18 *Let $q > k$, and let $A_1, A_2, \ldots, A_k$ be length-n q-ary codes, and let f be an n-variable multi-linear polynomial with non-negative integer coefficients. Then there exist explicitly constructible q-ary codes $A'_1, A'_2, \ldots, A'_k$, such that*

$$C_s(A'_1, A'_2, \ldots, A'_k) = f[CP_s(A_1, A_2, \ldots, A_k)],$$

a, for $s = 0, \ldots, q-1$, where the length of codes $A'_1, A'_2, \ldots, A'_k$ is $\left\lceil \frac{q}{q-k} \right\rceil \mathrm{L}_1(f)$;

b, for $s = 1, 2, \ldots, q-1$ and for $q \geq 2$, where the length of codes $A'_1, A'_2, \ldots, A'_k$ is $\mathrm{L}_1(f)$.

Moreover, codes $A'_1, A'_2, \ldots, A'_k$ can be computed from codes $A_1, A_2, \ldots, A_k$, and polynomial f in time $O(L_1(f)\deg(f)(|A_1| + |A_2| + \ldots + |A_k|))$.

Proof:

We need a k-wise generalization of Lemma 12:

Lemma 19 *Let $A_1, A_2, \ldots, A_k$ length-n q-ary codes, and let f be a multi-linear polynomial with n variables and either with non-negative integer coefficients or with integer coefficients from the modulo m ring of integers, and let s, σ be different elements of set $\{0, 1, \ldots, q-1\}$, and let $\Gamma \subset \{0, 1, \ldots, q-1\}$, such that neither s nor σ is in Γ . Then*

$$C_s(f_\sigma^\Gamma(A_1), f_\sigma^\Gamma(A_2), \ldots, f_\sigma^\Gamma(A_k)) = f[CP_s(A_1, A_2, \ldots, A_k)];$$

Proof:

Let $a(j)^1, a(j)^2, \ldots$ denote the words of code A_j for $j = 1, 2, \ldots, k$. We assume that the vector with coordinates $(i_1, i_2, \ldots, i_k)$ of the k-dimensional $CP_s(A_1, A_2, \ldots, A_k)$ matrix is $u = (u_1, u_2, \ldots, u_n) \in \{0,1\}^n$. That means, that $a(1)_t^{i_1} = a(2)_t^{i_2} = \cdots = a(k)_t^{i_k} = s$ exactly when $u_t = 1$, for $t = 1, 2, \ldots, n$. Now, $f(u)$ is equal to the number (counting the multiplicities) of monomials x_I in f, such that for all $t \in I$ $u_t = 1$. However, this happens exactly when for all $t \in I$ the coordinates, corresponding to the monomial x_I in the word i_1 of $f_\sigma^\Gamma(A_1)$ and in the word i_2 of $f_\sigma^\Gamma(A_2)$, ...,and in the word i_k of $f_\sigma^\Gamma(A_k)$ are all s; that means that the coordinate, corresponding to monomial x_I, contributes value 1 to element $(i_1, i_2, \ldots, i_k)$ of matrix $C_s(f_\sigma^\Gamma(A_1), f_\sigma^\Gamma(A_2), \ldots, f_\sigma^\Gamma(A_k))$. Note, that since the filling element σ differs from s, only the aforementioned contributions will be counted to this matrix.

□

Now we prove Theorem 18. First we prove statement (b). Here $s \neq 0$, so we can use $\sigma = 0$ for the filling element, and $\Gamma = \emptyset$. Consequently, let $A'_i = f_0(A_i)$, for $i = 1, 2, \ldots, k$, and the length of code A'_i is $\mathrm{L}_1(f)$.

Now we prove part (a). First, partition the set $\{0,1,\ldots,q-1\}$ into

$$h=\left\lceil\frac{q}{q-k}\right\rceil$$

classes $\Xi_1,\Xi_2,\ldots,\Xi_h$, each of size at most $q-k$. Codewords of the codes $A_1',A_2',\ldots,A_k'$ will consist of h segments. Segment i ($i=1,2,\ldots,h$) is constructed as follows: Suppose that $\{\alpha_1,\alpha_2,\ldots,\alpha_k\}=\{0,1,\ldots,q-1\}-\Xi_i$. Then let

$$\begin{array}{rcl}
A_1'^{(i)} & = & f_{\alpha_1}^{\{\alpha_2,\alpha_3,\alpha_4,\ldots,\alpha_k\}}(A_1),\\
A_2'^{(i)} & = & f_{\alpha_2}^{\{\alpha_1,\alpha_3,\alpha_4,\ldots,\alpha_k\}}(A_2),\\
A_3'^{(i)} & = & f_{\alpha_3}^{\{\alpha_1,\alpha_2,\alpha_4\ldots,\alpha_k\}}(A_3),\\
\vdots & \vdots & \vdots\\
A_k'^{(i)} & = & f_{\alpha_k}^{\{\alpha_1,\alpha_2,\ldots,\alpha_{k-1}\}}(A_k).
\end{array}\tag{1}$$

Certainly, in segment i, all the $s\in\Xi_i$ will be set according to the requirements of the Theorem, as follows from Lemma 19. So, concatenating the corresponding words from the segments $i=1,2,\ldots,h$ we will get codes $A_1',A_2',A_3',\ldots,A_k'$, each consist of words of length $h\mathrm{L}_1(f)$.

□

Theorem 20 *Let A be length-n q-ary code, and let f be an n-variable multi-linear polynomial with non-negative integer coefficients. Then there exist explicitly constructible q-ary code A', such that for any $k\geq 2$:*

$$C_s(\overbrace{A',A',\ldots,A'}^{k})=f[CP_s(\overbrace{A,A,\ldots,A}^{k})],$$

for $s=1,2,\ldots,q-1$ and for $q\geq 2$, where the length of code A' is $\mathrm{L}_1(f)$. Moreover, code A' can be computed from codes A and polynomial f in time $O(L_1(f)\deg(f)|A|)$.

Proof: The proof is immediate from the proof of part (a) of Theorem 18, one should choose 0 as the filling element. □

When f is a one-variable function, then it can be applied directly to the coincidence-matrix C_s:

Corollary 21 *Let A be a length-n q-ary code, let $q\geq 2$, and let F_r denote the r-element field, and let $f:F_r\to F_r$ be a polynomial. Then there exist an explicitly constructible q-ary code A', such that*

$$C_s(\overbrace{A',A',\ldots,A'}^{k})=f[C_s(\overbrace{A,A,\ldots,A}^{k})]$$

, for $s=1,2,\ldots,q-1$. The length of A' is at most $r\sum_{j=0}^{q-1}\binom{n}{j}$.

6 Applications for Hamming-Distances

Using Lemma 7 the following Corollary of Theorem 10 is immediate:

Corollary 22 *Let A and B be length-n q-ary codes, where $q \geq 4$, and let f be a multi-linear polynomial with non-negative integer coefficients. Then we can construct codes A' and B' with a $O(L_1(f)\deg(f)(|A|+|B|))$-time algorithm, such that*

$$H(A', B') = 2\mathrm{L}_1(f)J - \sum_{s=0}^{q-1} f[CP_s(A, B)],$$

and where the length of codes A' and B' is at most $2\mathrm{L}_1(f)$, and J denotes the all–1 matrix.

In the case when $A = B$, and we need $A' = B'$, the situation is somewhat harder: By Theorem 13 we cannot prescribe the coincidency-matrix for the filling-element, say for 0. So, we can state:

Corollary 23 *Let A be a length-n q-ary code, where $q \geq 2$, and let f be a multi-linear polynomial with n variables and with non-negative integer coefficients. Then we can construct code A' a $O(L_1(f)\deg(f)(|A|))$-time algorithm, such that*

$$H(A', A') = 2\mathrm{L}_1(f)J - \sum_{s=1}^{q-1} f[CP_s(A, B)] - C_0(A', A'),$$

and where the length of code A' is at most $\mathrm{L}_1(f)$.

In [GS01] we defined the k-wise Hamming-distance of codes; here we also need k-wise Hamming-distance matrices:

Definition 24 *Let $a^i \in A_i$, for $i = 1, 2, \ldots, k$. Their k-wise Hamming distance,*

$$d_k(a^1, a^2, \ldots, a^k)$$

is ℓ, if there exist exactly ℓ coordinates, in which they are not all equal. (Equivalently, their coordinates are all equal on $n-\ell$ positions). The k-wise Hamming-distance matrix $H(A_1, A_2, \ldots, A_k) = \{d_{i_1,i_2,\ldots,i_k}\}$ is a k-dimensional integer matrix, where

$$d_{i_1,i_2,\ldots,i_k} = d_k(a^{i_1}, a^{i_2}, \ldots, a^{i_k}),$$

where $a^{i_j} \in A_j$, $j = 1, 2, \ldots, k$.

As a corollary of Theorem 18 and the k-wise version of Lemma 7, we prove for the k-wise Hamming-distances of the new codes:

Corollary 25 *Let A_i $(i = 1, 2, \ldots, k)$ be length-n q-ary codes, where $q \geq 3$, and let f be a multi-linear polynomial with n variables and with non-negative integer coefficients. Then we can construct codes $A'_1, A'_2, \ldots, A'_k$ with a $O(L_1(f)\deg(f)(|A_1| + |A_2| + \cdots + |A_k|))$-time algorithm, such that*

$$H(A'_1, A'_2, \ldots, A'_k) = \left\lceil \frac{q}{q-k} \right\rceil \mathrm{L}_1(f)J - \sum_{s=0}^{q-1} f[CP_s(A_1, A_2, \ldots, A_k)],$$

and where the length of codes A'_i is at most $\left\lceil \frac{q}{q-k} \right\rceil \mathrm{L}_1(f)$, and J denotes the all–1 matrix.

Acknowledgment.
The author is grateful for Gábor Tardos for comments on this work. The author also acknowledges the partial support of János Bolyai Fellowship, and research grants EU IST FET No. IST-2001-32012 and OTKA T030059.

References

[Gro01] Vince Grolmusz. Constructing set-systems with prescribed intersection sizes. Technical Report DIMACS TR 2001-03, DIMACS, January 2001. ftp://dimacs.rutgers.edu/pub/dimacs/TechnicalReports/TechReports/2001/2001-03.ps.gz.

[GS01] Vince Grolmusz and Benny Sudakov. k-wise set-intersections and k-wise Hamming-distances. Technical Report DIMACS TR 2001-11, DIMACS, February 2001. To appear in JCT-A.