Speeding up deciphering by hypergraph ordering

Peter Horak [1,*] Zsolt Tuza [2,3,†]

[1] University of Washington, Tacoma, USA
[2] Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences
[3] University of Pannonia, Veszprém, Hungary

Latest update on September 18, 2013

Abstract

The "Gluing Algorithm" of Semaev [Des. Codes Cryptogr. 49 (2008), 47–60] — that finds all solutions of a sparse system of linear equations over the Galois field $GF(q)$ — has average running time $O(mq^{\max|\cup_1^k X_j|-k})$, where m is the total number of equations, and $\cup_1^k X_j$ is the set of all unknowns actively occurring in the first k equations. Our goal here is to minimize the exponent of q in the case where every equation contains at most three unknowns. The main result states that if the total number $|\cup_1^m X_j|$ of unknowns is equal to m, then the best achievable exponent is between $c_1 m$ and $c_2 m$ for some positive constants c_1 and c_2.

1 Introduction

Sparse objects such as sparse matrices, sparse system of (non-)linear equations occur frequently in science or engineering. For example, huge sparse matrices often appear when solving partial differential equations. It seems that [7] was the first monograph on the subject, see [3] for a more a recent one, and [4] for a monograph on solving sparse linear systems of equations.

Nowadays sparse systems are frequently studied in algebraic cryptoanalysis. First, given a cipher system, one converts it into a system of equations. Second, the system of equations is solved to retrieve either a key or a plaintext. As pointed in [2], this system of equations will be sparse, since efficient implementations of real-word systems require a low gate count. Also, as mentioned in [1], the cryptanalysis of several modern ciphers reduces to finding the common zeros of m quadratic polynomials in n unknowns over F_2. In the paper [1] an algorithm reducing the problem to a combination of exhaustive search and sparse linear algebra in given.

* Supported by a SPIRE grant from University of Bergen, Norway, and by a research grant from IAS, University of Washington.
† Supported in part by the Hungarian Scientific Research Fund, OTKA grant T-81493.

There are plenty of papers on methods for solving a sparse system of equations. In [5] a so called Gluing Algorithm was designed to solve such systems over a finite field $GF(q)$. If the set S_k of solutions of the first k equations together with the next equation $f_{k+1} = 0$ is given then the algorithm constructs the set S_{k+1}. It is shown there that the average complexity of finding all solutions to the original system is $O(mq^{\max|\cup_1^k X_j|-k})$, where m is the total number of equations, and $\cup_1^k X_j$ is the set of all unknowns actively occurring in the first k equations. Clearly, the complexity of finding all solutions to the system by the Gluing Algorithm depends on the order of equations. Therefore one is interested to find a permutation π that minimizes the average complexity, and also in the worst case scenario, i.e., the system of equation for which the average complexity is maximum. Therefore I. Semaev [6] suggested to study the following combinatorial problem.

Let $\mathcal{S}_{n,m,c}$ be a family of all collections of sets $\mathcal{X} = \{X_1, ..., X_m\}$, where $X_i \subset X, |X| = n$, and $|X_i| \leq c$ for all $i = 1, ..., m$; we allow that some set may occur in $\mathcal{X}$ more than once. Further, let π be a permutation on $[m] = \{1, ..., m\}$, and $1 \leq k \leq m$. Then we set $\Delta(\mathcal{X}, \pi, k) := \left|\bigcup_{i=1}^k X_{\pi(i)}\right| - k$, and $\Delta(\mathcal{X}, \pi) := \max_{1 \leq k \leq m} \Delta(\mathcal{X}, \pi, k)$, and $\Delta(\mathcal{X}) := \min_\pi \Delta(\mathcal{X}, \pi)$, where the minimum runs over all permutations π on $[m]$. Finally, $f_c(n, m) := \max_{\mathcal{X}} \Delta(\mathcal{X})$, where the maximum is taken over all families $\mathcal{X}$ in $\mathcal{S}_{n,m,c}$.

In this paper we confine ourselves to the case $|X_i| \leq 3$ for all $i \in [m]$, that is, to the case when each equation of the sparse system contains at most three active variables. We determine $f_2(n, m)$ for $n \geq 2$ and all m, and also $f_3(n, n)$ for $n \leq 9$. The main result of the paper claims that $f_3(n, n)$ grows linearly. More precisely we show that

Theorem 1 *For all n sufficiently large, $f_3(n, n) \geq 0.0818757697n \doteq \frac{n}{12.2137}$, while for all $n \geq 3, f_3(n, n) \leq \lceil \frac{n}{4} \rceil + 2$.*

Conjecture 2 *The quotient $\frac{f_3(n,n)}{n}$ tends to a constant as $n \to \infty$.*

We point out that after we obtained the above upper bound, an asymptotically better inequality $f_3(n, n) \leq \frac{n}{5} + 1 + \log_2 n$ has been proved in [6]. For small n the bound in Theorem 1 is slightly better. However, the main reason why we include it in the paper is that it applies different techniques, and we hope they may have the potential to obtain even a better bound.

2 Preliminaries

In this section we introduce some more needed notions and notation. Several auxiliary lemmas and observations will be stated as well.

We start with a lemma that allows one to confine to a special type of families in $\mathcal{S}_{n,m,c}$.

Lemma 3 *Let $n \geq c$, there exists a family $\mathcal{X} \in \mathcal{S}_{n,m,c}$ so that $\Delta(\mathcal{X}) = f_c(n,m)$ and $|X_i| = c$ for each $i = 1, ..., m$.*

Proof. Let $\mathcal{X} = \{X_1, ..., X_m\}$ and $\mathcal{X}' = \{X_1', ..., X_m'\}$ be in $\mathcal{S}_{n,m,c}$ and $X_i \subseteq X_i'$ for all i. Then $\Delta(\mathcal{X}) \leq \Delta(\mathcal{X}')$ and the statement follows. ■

The next observations follow directly from the definition of $\Delta(\mathcal{X}, \pi)$.

Lemma 4 *Let $\mathcal{X} = \{\mathcal{X}_1, ..., X_{k+s}\}$, $\mathcal{X}_k = \{\mathcal{X}_1, ..., X_k\}$, $\mathcal{Y}_s = \{Y_1, ..., Y_s\}$, $Y_i = X_{k+i} - \bigcup_{i=1}^{k} X_i$, and π_k and π' be the restriction of an ordering π of $[k+s]$ to $[k]$ and $[k+s] - [k]$, respectively. Then*

(a) $\boldsymbol{\Delta}(\mathcal{X}, \pi, k+s) = \boldsymbol{\Delta}(\mathcal{X}, \pi, k) + \boldsymbol{\Delta}(\mathcal{Y}_s, \pi', s) = \boldsymbol{\Delta}(\mathcal{X}, \pi, k) + \left| \bigcup_{i=k+1}^{s} X_i - \bigcup_{i=1}^{k} X_i \right| - s$

(b) $\boldsymbol{\Delta}(\mathcal{X}, \pi) = \max\{\boldsymbol{\Delta}(\mathcal{X}_k, \pi_k), \boldsymbol{\Delta}(\mathcal{X}, \pi, k) + \boldsymbol{\Delta}(\mathcal{Y}_s, \pi')\}$.

Clearly, for each $\mathcal{X} \in \mathcal{S}_{n,m,c}$ and all $k \leq m-1$, we get

$$-1 \leq \Delta(\mathcal{X}, \pi, k+1) - \Delta(\mathcal{X}, \pi, k) \leq c - 1.$$

The following observation will be frequently used.

Lemma 5 *Let $1 \leq s \leq c$. Then $\Delta(\mathcal{X}, \pi, k+1) - \Delta(\mathcal{X}, \pi, k) = s - 1$ iff $\left| X_{\pi(k+1)} - \bigcup_{i=1}^{k} X_{\pi(i)} \right| = s$.*

The notions of a connected/disconnected family of sets as well as a connectivity component will be transferred from the corresponding graph. More precisely:

Definition 6 *Let $\mathcal{X} = \{X_1, ..., X_m\}$. Then by $G_{\mathcal{X}} = (V, E)$ we denote a graph with the vertex set $V = \bigcup_{i=1}^{m} X_i$, and $\{i, j\}$ is an edge in E if there is a set X in $\mathcal{X}$ so that $\{i, j\} \subset X$. The family $\mathcal{X}$ will be called connected/disconnected if $G_{\mathcal{X}}$ is connected. If $\mathcal{X}$ is disconnected, and $C = (V_C, E_C)$ is a component of $G_{\mathcal{X}}$ then the set V_C will be called a component of $\mathcal{X}$. By the order $|C|$ of C we mean $|V_C|$, while by the size $e(C)$ of C we understand the number of sets X in $\mathcal{X}$ such that $X \subset V_C$.*

The following inequality is well known and easy to see.

Lemma 7 *Let $\mathcal{X} \in S_{n,m,c}$ be connected. Then $m \geq \left\lceil \frac{n-1}{c-1} \right\rceil$.*

A standard ordering π of sets in $\mathcal{X}$ will be defined recursively. Choose $X_{\pi(1)}$ in an arbitrary way. After $t \geq 1$ sets have been ordered (that is, when $\pi(1), ..., \pi(t)$ have been set) we choose $\pi(t+1)$ so that $\left| X_{\pi(t+1)} - \bigcup_{i=1}^{t} X_{\pi(i)} \right|$

is minimum. If $\mathcal{X}$ is connected, we have $\left|X_{\pi(t+1)} - \bigcup_{i=1}^{t} X_{\pi(i)}\right| \leq c-1$ for all $t \geq 1$. This in turn implies, see Lemma 5, that

$$\text{for all } t \leq n-1,\ \Delta(\mathcal{X},\pi,t+1) - \Delta(\mathcal{X},\pi,t) \leq c-2 \tag{1}$$

For a disconnected family $\mathcal{X}$ we get that in this case a standard ordering is obtained by first ordering the components of $\mathcal{X}$ and then the sets in the individual components are ordered in a standard way.

3 Families with 2-sets

In this section we determine the value of $f_2(n,m)$ for all m, n. It is obvious that for a connected family $\mathcal{X} \in S_{n,m,2}$, it is $\Delta(X) = 1$. The proof in the case of $\mathcal{X}$ disconnected is more involved. We note that following key claim is true only for families of 2-sets.

Lemma 8 *Let $\mathcal{X} \in S_{n,m,2}$. Then there is a standard ordering π so that $\Delta(\mathcal{X},\pi) = \Delta(\mathcal{X})$.*

Proof. Let τ be an ordering of sets in $\mathcal{X}$ such that $\Delta(\mathcal{X},\tau) = \Delta(\mathcal{X})$. We construct a desired ordering π in a recursive way. First we set $\pi(1) = \tau(1)$. After $\pi(t)$ has been set (and $t < m$), we define $\pi(t+1)$ as follows. If possible choose $\pi(t+1)$ such that

$$\left|X_{\pi(t+1)} \cap \bigcup_{i=1}^{t} X_{\pi(i)}\right| \leq 1 \tag{2}$$

is satisfied, otherwise we set $\pi(t+1) = \tau(s)$, where s is the smallest number such that $X_{\tau(s)}$ has not been ordered yet in the permutation π. It is not difficult to check that for all $k \leq m$ we have $\Delta(\mathcal{X},\pi,k) \leq \Delta(\mathcal{X},\tau,k)$. ■

We recall that a component of a graph comprising a single vertex is called a singleton, or trivial.

Theorem 9 *For $n \geq 2$ and all m, $f_2(n,m)$ equals the maximum number of non-trivial components in a simple graph on n vertices with m edges; i.e., $f_2(n,m) = m$ for $m \leq \frac{n}{2}$, $f_2(n,m) = n-m$ for $\frac{n}{2} < m < n-1$, and $f_2(n,m) = 1$ for $m \geq n-1$.*

Proof. Let $\mathcal{X} = \{X_1, ..., X_m\}$ be a family of sets so that $\Delta(\mathcal{X}) = f_2(n,m)$. By Lemma 3, we assume that $|X_i| = 2$ for all $i \in [n]$. Consider first the case when $\mathcal{X}$ is connected; clearly in this case we have $m \geq n-1$. Let π be a standard ordering of sets in $\mathcal{X}$. Then $\Delta(\mathcal{X},\pi,1) = 1$, and, by (1), $\Delta(\mathcal{X},\pi,t+1) - \Delta(\mathcal{X},\pi,t) \leq 0$ for all $k \leq m-1$. Thus $\Delta(\mathcal{X}) = f_2(n,m) = 1$.

Suppose now that $\mathcal{X}$ is disconnected. With respect to Lemma 8, we can confine ourselves to standard orderings. As mentioned in Preliminaries, a standard ordering of a disconnected family $\mathcal{X}$ is an ordering where first the components of $\mathcal{X}$ are ordered and then the sets in individual components are ordered in a standard way. For each component C of $G_{\mathcal{X}}$ we set $d(C) := |C| - e(C)$. Obviously, $d(C) \leq 1$ for each component C, and if $C_1, ..., C_t$ are all components of G then

$$\sum_{i=1}^{t} d(C_i) = \sum_{i=1}^{t} |C_i| - e(C_i) = \sum_{i=1}^{t} |C_i| - \sum_{i=1}^{t} e(C_i) = n - m. \tag{3}$$

Claim 10 *Let π be a standard ordering of $\mathcal{X}$. Then*

$$\Delta(\mathcal{X}, \pi) = \max_{0 \leq s \leq t-1} \{1 + \sum_{i=1}^{s} d(C_{\pi(i)})\}.$$

We have shown above that if $\mathcal{X}$ is connected then $\Delta(\mathcal{X}) = 1$. To show the statement it suffices to note that, for $k < m$, $\Delta(\mathcal{X}, \pi, k) = 1 + \sum_{i=1}^{s} d(C_{\pi(i)})$, where s is the number for which $\sum_{i=1}^{s} e(C_{\pi(i)}) \leq k < \sum_{i=1}^{s+1} d(C_{\pi(i)})$, and $\Delta(\mathcal{X}, \pi, m) = n - m$. From the above claim we immediately get one of key observations:

Claim 11 *Let π be a standard ordering such that the components of $G_{\mathcal{X}}$ are ordered in the increasing way with respect to the invariant $d(C)$, and let τ be any standard ordering of $\mathcal{X}$. Then $\Delta(\mathcal{X}, \tau) \geq \Delta(\mathcal{X}, \pi) = \Delta(\mathcal{X})$.*

Thus, we can confine ourselves to the ordering π. We assume without loss of generality that $C_1, ..., C_t$ is the order of components in this ordering. Let $m \geq n - 1$. Then, by (3), $\sum_{i=1}^{s} d(C_i) \leq 0$ for each $s < t$, and, by Lemma 10, $\Delta(\mathcal{X}) = 1$. Assume now $m < n - 1$. Then, again by (3) and Lemma 10, $\Delta(\mathcal{X}, \pi)$ is maximized by a family $\mathcal{X}$ with all components C of $\mathcal{X}$ satisfying $d(C) = 1$, thus $\Delta(\mathcal{X}, \pi)$ is maximized by a family $\mathcal{X}$ where the corresponding graph $G_{\mathcal{X}}$ possesses the maximum possible number of non-trivial components among all graphs on n vertices and m edges. ■

4 Families with 3-sets

For the rest of the paper we deal only with families of 3-sets. Thus, in $f_3(n, m)$ we will drop the subscript and write $f(n, m)$; in addition, for the most interesting case of $n = m$, we write only $f(n)$.

4.1 Exact values

There are only a few values of $f(n)$ that we are able to determine analytically. Here we state only values for $n \leq 9$, as otherwise determining the value $f(n)$ is too elaborate as it requires considering a large number of cases. We start with a rather obvious result that will simplify the proof of the next theorem.

Lemma 12 *For all $n \geq 3$, $f(n) \leq f(n+1)$.*

Proof. Let $\mathcal{X} = \{X_1, ..., X_n\} \in \mathcal{S}_{n,n,3}$ be such that $\Delta(\mathcal{X}) = f(n)$ and $z \notin \bigcup_{i=1}^n X_i$. Set $\mathcal{X}' = \mathcal{X} \cup \{z\}$. Let π' be an ordering of sets in $\mathcal{X}'$. Consider the ordering π of sets in $\mathcal{X}$, obtained by dropping the set $\{z\}$ from this order. Then $\Delta(\mathcal{X}', \pi') = \Delta(\mathcal{X}, \pi)$, and the statement follows. ■

Theorem 13 *$f(3) = 2$, and $f(n) = \left\lceil \frac{n}{3} \right\rceil$ for $4 \leq n \leq 9$.*

Proof. The statement is obvious for $n = 3$. First we show that, for $4 \leq n \leq 9$, $f(n) \geq \left\lceil \frac{n}{3} \right\rceil$. By $f(3) = 2$ and Lemma 12, $f(n) \geq 2$ for all n; this proves the lower bound for $4 \leq n \leq 6$. To see that $f(n) \geq 3$ for $n = 7, 8, 9$ it is sufficient to take for $\mathcal{X}$ a family such that for any two triples X, X' in $\mathcal{X}$ it is $|X \cap X'| \leq 1$. Then, for any permutation π we get $\Delta(\mathcal{X}, \pi, 2) \geq 3$, that is, $\Delta(\mathcal{X}) \geq 3$. We note that, for $n = 7$, the Fano plane, and for $n = 9$, any 9 triples of the unique Steiner triple system $STS(9)$ have the property. For $n = 8$, to get the desired family of 8 triples it suffices to remove from $STS(9)$ all triples incident with a fixed element x_0.

We note that we are able to prove that $f(n) \leq \left\lceil \frac{n}{3} \right\rceil$ for all $n \geq 4$. This bound is better than the bound $f(n) \leq \left\lceil \frac{n}{4} \right\rceil + 2$, proved in this paper, for a few small values of n. We have not included the proof of $f(n) \leq \left\lceil \frac{n}{3} \right\rceil$ to this paper as it is quite long. To have our paper self-contained we prove here $f(n) \leq \left\lceil \frac{n}{3} \right\rceil$ only for $n \leq 9$. In view of Claim 12, it suffices to show that $f(6) \leq 2$, and $f(9) \leq 3$.

For $n \in \{6, 9\}$, let $\mathcal{X} \in \mathcal{S}_{n,n,,3}$ be such that $\Delta(\mathcal{X}) = f(n)$, and $|X_i| = 3$, $i \in [n]$, see Lemma 3. For $\mathcal{X}$ disconnected, the inequality $\Delta(\mathcal{X}) \leq \left\lceil \frac{n}{3} \right\rceil$ follows from Lemma 17, as the order of the largest component $G_{\mathcal{X}}$ is at most $n - 3$. So now we assume that $\mathcal{X}$ is connected. We will construct in a recursive way an ordering π of sets in $\mathcal{X}$ such that $\Delta(\mathcal{X}, \pi) \leq \left\lceil \frac{n}{3} \right\rceil$. Let $e = \{x, y\}$ be an edge with maximum multiplicity $m(e) = M$ in $G_{\mathcal{X}}$. At the beginning of the order π come all sets X_i with $\{x, y\} \subset X_i$. Thus, $\Delta(\mathcal{X}, \pi, k) = 2$ for all $k \leq m$ and $\left|\bigcup_{i=1}^M X_{\pi(i)}\right| = M + 2$. For $n = 6$, $G_{\mathcal{X}}$ has 18 edges, thus there is an edge e in $G_{\mathcal{X}}$ with multiplicity $m(e) > 1$. Assume that $t < n$ sets in $\mathcal{X}$ have been ordered. As e has the maximum multiplicity, for $n = 6$, the set $X_{\pi(t+1)}$ can be chosen such that (2) is satisfied. Thus $\Delta(\mathcal{X}, \pi, k) = 2$ for all $k \leq n$, i.e., $\Delta(\mathcal{X}) \leq 2$.

So we are left with the case $n = 9$. After M sets containing x, y we order, in a recursive way, sets, if any, satisfying (2). If we are able to order in this way all sets of $\mathcal{X}$, then even $\Delta(\mathcal{X}, \pi, k) = 2$ for all $k \leq n$, and we are done. Otherwise, as $\mathcal{X}$ is connected, we are able to choose as $X_{\pi(t+1)}$ a set satisfying $\left|X_{\pi(t+1)} \cap \bigcup_{i=1}^t X_{\pi(i)}\right| = 2$. Then $\Delta(\mathcal{X}, \pi, k) \leq 3$ for all $k \leq t + 1$. We note that in all cases, including $M = 1$, we have at this moment t sets ordered with $\left|\bigcup_{i=1}^t X_{\pi(i)}\right| \geq 5$. We leave it to the reader to check that the remaining sets can be ordered to satisfy (2). The proof is complete. ■

4.2 Lower bound

Theorem 14 *For n sufficiently large, we get $f(n) > 0.0818757697\,n$.*

Proof. We will prove the existence of a family $\mathcal{X} = \{X_1, ..., X_n\} \in \mathcal{S}_{n.n,3}$, $X_i \subset \{x_1, ..., x_n\}$, with the required property $\Delta(\mathcal{X}) \geq 0.0818757697\,n$ using the following probabilistic model: Select two permutations π and τ on $[n]$ randomly and independently; that is, any permutation on $[n]$ coincides with π and with τ with probability $1/n!$, and any ordered pair of permutations of $[n]$ coincides with (π, τ) with probability $(1/n!)^2$. Set

$$X_i := \{i, \pi(i), \tau(i)\} \qquad i = 1, \ldots, n.$$

We will prove that for n sufficiently large $\mathcal{X}$ satisfies $\Delta(\mathcal{X}) \geq 0.0818757697\,n$ with a positive probability. Hence there exists at least one set system having $\Delta(\mathcal{X})$ sufficiently large. More precisely, we shall prove that there exist positive constants c and ε with the following property: The union of any cn members of $\mathcal{X}$ have cardinality at least $(c+\varepsilon)n$ with positive probability as n gets large. For simplicity, but without loss of generality we assume here and also below that cn and εn are integers. This implies that for any ordering λ of members of $\mathcal{X}$,where n is sufficiently large, we have $\Delta(\mathcal{X}, \lambda, cn) \geq \varepsilon n$; that is $\Delta(\mathcal{X}) \geq \varepsilon n$. Computation will show that the requirement is satisfied if we put $c = 0,4590625$ and $\varepsilon = 0.0818757697$. To prove the statement we will show that $\mathcal{X}$ contains, with the probability strictly less than 1, a subfamily $\{X_{i_1}, ..., X_{i_m}\}$ of $k := cn$ members such that their union $Y = X_{i_1} \cup ... \cup X_{i_k}$ is of cardinality at most $(c+\varepsilon)n$.

A subfamily of m members can be chosen in $\binom{n}{cn}$ ways. Clearly, by definition of X_i, $x_{i_j} \in Y$ for $j = 1, ..., m$. Therefore, there are $\binom{n-cn}{\varepsilon n}$ ways how to choose additional εn elements in Y. Let $M = \{i_1, ..., i_m\}$. Then $\pi(M)$ can be chosen in $\binom{(c+\varepsilon)n}{cn}$ ways, and π can be defined on M in $(cn)!$ ways, while π can be defined on $[n] - M$ in $((1-c)n)!$ ways. Since the permutations π and τ have been chosen independently, the same is valid for τ. Finally, the pair (π, τ) has been chosen with probability $(n!)^2$. Thus, in aggregate, the probability p that $\mathcal{X}$ contains a subfamily of cn elements with their union being of cardinality at most $(c+\varepsilon)n$ is

$$p \leq \frac{\binom{n}{cn}\binom{n-cn}{\varepsilon n}\binom{(c+\varepsilon)n}{cn}^2 (cn)!^2((1-c)n)!^2}{(n!)^2}$$

which in turn equals

$$p \leq \frac{\binom{n}{cn}\binom{n-cn}{\varepsilon n}\binom{(c+\varepsilon)n}{cn}^2}{\binom{n}{cn}^2} = \frac{\binom{(1-c)n}{\varepsilon n}\binom{(c+\varepsilon)n}{cn}^2}{\binom{n}{cn}}$$

We will calculate c and ε so that $p < 1$. It is well known that from Stirling formula we get

$$\frac{\log_2 \binom{x}{ax}}{x} \to H(a) \text{ for } x \to \infty, \ \ a \text{ fixed, where } H(a) = -a \log_2 a - (1-a) \log_2 (1-a).$$

Thus, taking binary logarithm of $p < 1$ we get that the inequality holds for every sufficiently large n if

$$\log_2 \binom{(1-c)n}{\varepsilon n} + 2 \log_2 \binom{(c+\varepsilon)n}{cn} - \log_2 \binom{n}{cn} < 0,$$

that is, if

$$\frac{1}{n}\left[\frac{1-c}{1-c} \log_2 \binom{(1-c)n}{\frac{\varepsilon}{1-c}(1-c)n} + \frac{c+\varepsilon}{c+\varepsilon} 2 \log_2 \binom{(c+\varepsilon)n}{\frac{c}{c+\varepsilon}(c+\varepsilon)n} - \log_2 \binom{n}{cn}\right] < 0,$$

hence

$$(1-c)H(\frac{\varepsilon}{1-c}) + 2(c+\varepsilon)H(\frac{c}{c+\varepsilon}) - H(c) < 0.$$

Substituting the values $c = 0.4590625$ and $\varepsilon = 0.0818757697241$ one can check that the left side of the above inequality is approximately -0.0000000000005 and hence strictly negative, consequently the probability $p < 1$ for n large enough. The proof is complete. ∎

4.3 Upper bound

In this section we will prove that, for all $n \geq 4$, $f(n) \leq \lceil \frac{n}{4} \rceil + 2$. We will start with a series of auxiliary upper bounds. The first one looks to be fairly crude but for m small with respect to n it is sharp.

Lemma 15 *For all n, m we get $f(n,m) \leq 2 \lceil \frac{n}{3} \rceil$.*

Proof. Let $\mathcal{X} \in \mathcal{S}_{n,m,3}$ and π be a permutation on $[m]$. For any $k \leq \lceil \frac{n}{3} \rceil$, $\Delta(\mathcal{X}, \pi, k) = \left|\bigcup_{i=1}^{k} X_{\pi(i)}\right| - k \leq 3k - k \leq 2 \lceil \frac{n}{3} \rceil$. Otherwise, $\left|\bigcup_{i=1}^{k} X_{\pi(i)}\right| - k \leq n - k \leq n - (\lceil \frac{n}{3} \rceil + 1) \leq 2 \lceil \frac{n}{3} \rceil$. ∎

A better bound can be obtained if m is sufficiently large. Also in this case the bound is sharp for some values of m.

Lemma 16 *For $m \geq \lceil \frac{n-1}{2} \rceil$, we have $f(n,m) \leq \lfloor \frac{n+1}{2} \rfloor$.*

Proof. Let $\mathcal{X} = \{X_1, ..., X_m\}$ be a family of sets such that $\Delta(\mathcal{X}) = f(n,m)$. By Lemma 3, we assume that $|X_i| = 3$ for all $i \in [m]$. Consider first a case when $\mathcal{X}$ is connected, and let π be a standard ordering of $\mathcal{X}$. Then $\Delta(\mathcal{X}, \pi, 1) = 2$, and by (1), $\Delta(\mathcal{X}, \pi, t+1) - \Delta(\mathcal{X}, \pi, t) \leq 1$, for all $t \geq 1$. As $\Delta(\mathcal{X}, \pi, t+1) - \Delta(\mathcal{X}, \pi, t) = 1$

implies $\left|X_{\pi(t+1)} - \bigcup_{i=1}^{t} X_{\pi(i)}\right| = 2$, we have that $\Delta(\mathcal{X},\pi) \leq 2 + \left\lfloor \frac{n-3}{2} \right\rfloor = \left\lfloor \frac{n+1}{2} \right\rfloor$, thus $\Delta(\mathcal{X}) \leq \left\lfloor \frac{n+1}{2} \right\rfloor$.

Now let $\mathcal{X}$ be disconnected and $C_1, ..., C_s$ be components of the graph $G_{\mathcal{X}}$. By Lemma 7, a connected family $\mathcal{Y} \in \mathcal{S}_{n,m,3}$ has to contain at least $\left\lceil \frac{n-1}{2} \right\rceil$ triples. We define $\gamma(C_i) = e(C_i) - \left\lceil \frac{|C_i|-1}{2} \right\rceil$. Thus, $\gamma(C_i) \geq 0$ for all $i \in [s]$. Moreover, $\sum_{i=1}^{s} \gamma(C_i) = \sum_{i=1}^{s} (e(C_i) - \left\lceil \frac{|C_i|-1}{2} \right\rceil) = m - \sum_{|C_i| \text{ odd}} \left\lceil \frac{|C_i|-1}{2} \right\rceil - \sum_{|C_i| \text{ even}} \left\lceil \frac{|C_i|-1}{2} \right\rceil = m - \sum_{|C_i| \text{ odd}} \frac{|C_i|-1}{2} - \sum_{|C_i| \text{ even}} \frac{|C_i|}{2} = m - \frac{n}{2} + \frac{odd}{2} \geq \left\lceil \frac{n-1}{2} \right\rceil - \frac{n}{2} + \frac{odd}{2} = \left\lfloor \frac{odd}{2} \right\rfloor$, where *odd* is the number of the odd order components in $G_{\mathcal{X}}$. Let π be a standard ordering on $\mathcal{X}$ where the components C_i are ordered in the decreasing manner with respect to γ; without loss of generality we assume that $C_1, ..., C_s$ is this ordering. Then $\sum_{i=1}^{s} \gamma(C_i) \geq \left\lfloor \frac{odd}{2} \right\rfloor$ implies that, for all $t \leq s$,

$$\sum_{i=1}^{t} \gamma(C_i) \geq \min\{t, \left\lfloor \frac{odd}{2} \right\rfloor\}. \tag{4}$$

Now we show that for every $k, 1 \leq k \leq m$, we have $\Delta(\mathcal{X},\pi,k) \leq \left\lfloor \frac{n+1}{2} \right\rfloor$. For $k < m$, there is a unique t such that $\sum_{i=1}^{t} e(C_i) \leq k < \sum_{i=1}^{t+1} e(C_i)$. By Lemma 4(a) we get

$$\Delta(\mathcal{X},\pi,k) = \Delta(\mathcal{X},\pi,\sum_{i=1}^{t} e(C_i)) + \Delta(\mathcal{X}_i,\pi_i,k - \sum_{i=1}^{t} e(C_i)),$$

where $\mathcal{X}_i$ is a subfamily of $\mathcal{X}$ comprising triples that are subsets of C_i, and π_i is the restriction of π to $\mathcal{X}_i$. From the case of a connected family $\mathcal{X}$ discussed above we have $\Delta(\mathcal{X}_i,\pi_i,k - \sum_{i=1}^{t} e(C_i)) \leq \left\lfloor \frac{|C_i|+1}{2} \right\rfloor$. Denote by odd_t the number of odd components among $C_1, ..., C_t$. Then, $\Delta(\mathcal{X},\pi,k) \leq \sum_{i=1}^{t} |C_i| - \sum_{i=1}^{t} e(C_i) + \left\lfloor \frac{|C_{t+1}|+1}{2} \right\rfloor \leq \left\lfloor \frac{|C_{t+1}|+1}{2} \right\rfloor + \sum_{i=1}^{t} (\left\lfloor \frac{|C_i|+1}{2} \right\rfloor - \gamma(C_i)) \leq$
$\left\lfloor \frac{|C_{t+1}|+1}{2} \right\rfloor + \sum_{|C_i| \text{even}} \frac{|C_i|}{2} + \sum_{|C_i| \text{odd}} \frac{|C_i|}{2} + \frac{odd_t}{2} - \sum_{i=1}^{t} \gamma(C_i) \leq \left\lfloor \frac{n+1}{2} \right\rfloor$ as,

by (4), $\frac{odd_t}{2} \leq \sum_{i=1}^{t} \gamma(C_i)$, and $\sum_{|C_i| \text{even}} \frac{|C_i|}{2} + \sum_{|C_i| \text{odd}} \frac{|C_i|}{2} + \left\lfloor \frac{|C_{t+1}|+1}{2} \right\rfloor \leq \left\lfloor \frac{n+1}{2} \right\rfloor$. ■

The next auxiliary bound deals with the case when $\mathcal{X}$ is disconnected.

Lemma 17 *If $\mathcal{X} \in \mathcal{S}_{n,m,3}$ is disconnected, and $m \geq n$, then $\Delta(\mathcal{X}) \leq \left\lfloor \frac{c+1}{2} \right\rfloor$, where c is the order of the largest component of $G_{\mathcal{X}}$.*

Proof. Let $C_1, ..., C_s$ be components of $G_{\mathcal{X}}$, and let $\mathcal{X}_i$ be the subfamily of $\mathcal{X}$ comprising triples that are subsists of C_i. As in the proof of Theorem 9, we set

$d(C) = |C_i| - e(C_i)$ and get $\sum_{i=1}^{s} d(C_i) = n - m \le 0$. Consider the standard ordering π of $\mathcal{X}$ such that the components are ordered in the increasing way with respect to the invariant d; without loss of generality. we assume that $C_1, ..., C_s$ is such ordering. Then $\sum_{i=1}^{t} d(C_i) \le 0$ for any $t \le s$. Let π_i be a restriction of π to $\mathcal{X}_i$. Then, by Lemma 16, for each component C_i, we have

$$\max_{1 \le k \le e(C_i)} \Delta(\mathcal{X}_i, \pi_i, k) \le \left\lfloor \frac{|C_i| + 1}{2} \right\rfloor$$

Further, $\Delta(\mathcal{X}_i, \pi_i, e(C_i)) = |C_i| - e(C_i) = d(C_i)$. Extending this conclusion to π we have: If a is the total number of triples in the first t components, then

$$\Delta(\mathcal{X}, \pi, a) = \sum_{i=1}^{t} d(C_i) \le 0.$$

Let $k < m$. Then there is a uniquely determined number t such that $\sum_{i=1}^{t-1} e(C_i) \le k < \sum_{i=1}^{t} e(C_i)$. Set $a = \sum_{i=1}^{t-1} e(C_i)$. By Lemma 4(a), $\Delta(\mathcal{X}, \pi, k) = \Delta(\mathcal{X}, \pi, a) + \Delta(\mathcal{X}_t, \pi_t, k - a) \le \Delta(\mathcal{X}_t, \pi_t, k - a) \le \left\lfloor \frac{|C_i|+1}{2} \right\rfloor$. The proof is complete. ∎

Before proving the upper bound we state one more lemma.

Lemma 18 *Let $\mathcal{X} \in \mathcal{S}_{n,n,3}$, and let $\Delta(\mathcal{X}, \pi, k) = \lceil \frac{n}{4} \rceil + 1$. Then there is an $\varepsilon \ge 0$ such that $k = \lceil \frac{n}{4} \rceil + \varepsilon$, and $\left| \bigcup_{i=1}^{k} X_{\pi(i)} \right| = \lceil \frac{n}{2} \rceil + \varepsilon$.*

Proof. First let $\mathcal{X}$ be connected, and τ be a standard ordering of $\mathcal{X}$. By (1), for $s \ge 1$ we have $\Delta(\mathcal{X}, \tau, s+1) - \Delta(\mathcal{X}, \tau, s) \le 1$. Therefore $\Delta(\mathcal{X}, \tau, s) \le 2 + (s-1) = s + 1$. Thus, if $\Delta(\mathcal{X}, \tau, s) = \lceil \frac{n}{4} \rceil + 1$ then $s \ge \lceil \frac{n}{4} \rceil$. Then, by the definition of π and k it is $k \ge \lceil \frac{n}{4} \rceil$, i.e., $k = \lceil \frac{n}{4} \rceil + \varepsilon$ for some $\varepsilon \ge 0$, which in turn implies $|M| = k + \Delta(\mathcal{X}, \pi, k) \ge \lceil \frac{n}{2} \rceil + \varepsilon$.

Now let $\mathcal{X}$ be disconnected. By Lemma 17, $\Delta(\mathcal{X}) \le \left\lfloor \frac{|C|+1}{2} \right\rfloor$, where C is the largest component of $G_{\mathcal{X}}$. As $\Delta(\mathcal{X}) > \lceil \frac{n}{4} \rceil + 2$, we get $|C| \ge \lceil \frac{n}{2} \rceil + 2$. Since the subfamily of $\mathcal{X}$ with its triples in C is connected, $e(C) \ge \left\lceil \frac{|C|-1}{2} \right\rceil \ge \lceil \frac{n}{4} \rceil$. Now it suffices to repeat the argument used in the case $\mathcal{X}$ is connected. ∎

Theorem 19 *For all $n \ge 4$, $f(n) \le \lceil \frac{n}{4} \rceil + 2$.*

Proof. Let $\mathcal{X} = \{X_1, ..., X_n\}$ be such that $\Delta(\mathcal{X}) = f(n)$ and, see Lemma 3, $|X_i| = 3$ for all $i \in [n]$. Assume by contradiction that $\Delta(\mathcal{X}) > \lceil \frac{n}{4} \rceil + 2$. We choose an ordering π of $\mathcal{X}$ and a number $k \in [n]$ so that k is the largest number with the property (a) $\Delta(\mathcal{X}, \pi, k) = \lceil \frac{n}{4} \rceil + 1$, and $\Delta(\mathcal{X}, \pi, s) \le \lceil \frac{n}{4} \rceil + 1$ for all $s \le k - 1$, (b) for all orderings τ of $\mathcal{X}$ there is $s_\tau \le k + 1$ such that

$\Delta(\mathcal{X}, \tau, s_\tau) > \lceil \frac{n}{4} \rceil + 1$. From (a) we have that $\Delta(\mathcal{X}, \pi, k+1) > \lceil \frac{n}{4} \rceil$ +1, and by Lemma 5,

$$\left| X_{\pi(t)} - \bigcup_{i=1}^{k} X_{\pi(i)} \right| \geq 2 \text{ for all } t > k.$$

Set $M = \bigcup_{i=1}^{k} X_{\pi(i)}$. Let $\overline{M}$ denote the complement to M with respect to the underlying set. By Lemma 9, $k \geq \lceil \frac{n}{4} \rceil + \varepsilon$, and $\left|\overline{M}\right| = n - (\lceil \frac{n}{2} \rceil + \varepsilon) = \lfloor \frac{n}{2} \rfloor - \varepsilon$, where $\varepsilon \geq 0$. Further, for $i = 2, 3$, let $\mathcal{A}_i$, $a_i = |\mathcal{A}_i|$, be the subfamily of $\mathcal{X}$, such that $X \in \mathcal{A}_i$ if $\left|X \cap \overline{M}\right| = i$. We note that $\mathcal{A}_2 \cup \mathcal{A}_3$ comprises $n - k$ sets of $\mathcal{X}$ that come in the ordering π after $X_{\pi(k)}$. We choose π so that all sets in $\mathcal{A}_2$ come in the ordering π before sets from $\mathcal{A}_3$. Two cases are considered.

First, let $a_2 = \left|\overline{M}\right| + \alpha$, where $\alpha \geq 0$. Let $\mathcal{B}_2$ be a family of 2-sets, $\mathcal{B}_2 = \{ B;\ B = \overline{M} \cap X$ for some $X \in \mathcal{A}_2\}$. Then, by Lemma 4(a), for each t, $k+1 \leq t \leq k + a_2$, we have $\Delta(\mathcal{X}, \pi, t) = \Delta(\mathcal{X}, \pi, k) + \Delta(\mathcal{B}_2, \pi', t-k) = \lceil \frac{n}{4} \rceil + 1 + \Delta(\mathcal{B}_2, \pi', t-k)$, where π' is the restriction of π to the set $\{k+1, ..., k+a_2\}$. By Theorem 9, for $m \geq n$, $f_2(n, m) = 1$. Hence, if we choose π' to be the same permutation as in the proof of Theorem 9, then $\Delta(\mathcal{B}_2, \pi', t-k) \leq 1$ for all $1 \leq t-k \leq a_2$. Thus, $\Delta(\mathcal{X}, \pi, s) \leq \lceil \frac{n}{4} \rceil + 2$ for all $k+1 \leq s \leq k+a_2$. Let $B_2 = \bigcup_{B \in \mathcal{B}_2} B$. By Lemma 4(a), $\Delta(\mathcal{X}, \pi, k+a_2) = \Delta(\mathcal{X}, \pi, k) + \Delta(\mathcal{B}_2, \pi', a_2) = \lceil \frac{n}{4} \rceil + 1 + (|B_2| - a_2)$. To finish the proof of this part we will show that $\Delta(X, \pi, s) \leq \lceil \frac{n}{4} \rceil + 2$ is true also for all s, $k + a_2 + 1 \leq s \leq n$. Again by Lemma 4(a),

$$\Delta(\mathcal{X}, \pi, s) = \Delta(\mathcal{X}, \pi, k+a_2) + \Delta(\mathcal{A}_3^*, \pi^*, s - (k+a_2))$$

where $\mathcal{A}_3^*$ consists of sets $X \cap (\overline{M} - B_2)$, $X \in \mathcal{A}_3$, and π^* is the restriction of π to $\mathcal{A}_3^*$. By Lemma 15, for any $s - (k+a_2) \leq a_3$ we have $\Delta(\mathcal{A}_3^*, \pi^*, s-(k+a_2)) \leq \frac{2}{3}(\left|\overline{M}\right| - |B_2|)$. Hence $\Delta(\mathcal{X}, \pi, s) = \lceil \frac{n}{4} \rceil + 1 + (|B_2| - \left|\overline{M}\right| - \alpha) + \frac{2}{3}(\left|\overline{M}\right| - |B_2|) \leq \lceil \frac{n}{4} \rceil + 1$ as $B_2 \subset \overline{M}$, i.e., $|B_2| \leq \left|\overline{M}\right|$ and $\alpha \geq 0$.

We are left with the case $a_2 = \left|\overline{M}\right| - \alpha$, where $\alpha > 0$. We consider an ordering τ, where the triples in $\mathcal{A}_3$ come at the very beginning of this ordering, followed by triples from $\mathcal{A}_2$. At the very end of the ordering come triples in $\mathcal{X} - (\mathcal{A}_2 \cup \mathcal{A}_3)$ in the same order as in the ordering π. As $\left|\overline{M}\right| = \lfloor \frac{n}{2} \rfloor - \varepsilon$, and $\alpha > 0$, we have $a_3 = n - k - a_2 = n - (\lceil \frac{n}{4} \rceil + \varepsilon) - (\left|\overline{M}\right| - \alpha) \geq \lfloor \frac{3n}{4} \rfloor - \varepsilon - \lfloor \frac{n}{2} \rfloor + \varepsilon - \alpha \geq \lfloor \frac{n}{4} \rfloor \geq \left\lceil \frac{|\overline{M}|-1}{2} \right\rceil$. Therefore, by Lemma 16, for all $s \leq a_3$, we have $\Delta(\mathcal{X}, \tau, s) \leq \left\lfloor \frac{|\overline{M}|+1}{2} \right\rfloor \leq \lceil \frac{n}{4} \rceil + 1$. Let $B_3 =: \bigcup_{X \in \mathcal{A}_3} X$. Then, $\Delta(\mathcal{X}, \tau, a_3) = |B_3| - a_3$. We get, by Lemma 4(b),

$$\max_{1 \leq s \leq a_2 + a_3} \Delta(\mathcal{X}, \tau, s) = \max\{\max_{1 \leq s \leq a_3} \Delta(\mathcal{X}, \tau, s), \max_{a_3+1 \leq s \leq a_2+a_3} \Delta(\mathcal{X}, \tau, s)\} \leq$$

$$\max\{\left\lceil \frac{n}{4} \right\rceil + 1, \Delta(\mathcal{X}, \tau, a_3) + \max_{1 \leq s \leq a_2} \Delta(\mathcal{A}_2^*, \tau', s)\},$$

where τ' is the restriction of τ to $\{a_3+1,...,a_3+a_2\}$ and $\mathcal{A}_2^*$ comprises sets $X-B_3, X\in\mathcal{A}_2$. As $|X\cap M|=1$ for all $X\in\mathcal{A}_2$ we further get

$$\begin{aligned}\max_{1\le s\le a_2+a_3}\Delta(\mathcal{X},\tau,s) &\le \max\{\left\lceil\frac{n}{4}\right\rceil+1, |B_3|-a_3+\max_{1\le s\le a_2}\left|\bigcup_{i=1}^{s}X_{\tau(i)}\cap(\overline{M}-B_3)\right|\}\\ &\le \max\{\left\lceil\frac{n}{4}\right\rceil+1, |B_3|-a_3+\left|\overline{M}-B_3\right|\}\le\max\{\left\lceil\frac{n}{4}\right\rceil+1,\left|\overline{M}\right|-a_3\},\end{aligned}$$

since $B_3\subset\overline{M}$. Finally, because $a_3\ge\left\lceil\frac{|\overline{M}|-1}{2}\right\rceil$ and $\left|\overline{M}\right|=\left\lfloor\frac{n}{2}\right\rfloor-\varepsilon$, we get

$$\max_{1\le s\le a_2+a_3}\Delta(\mathcal{X},\tau,s)\le\left\lceil\frac{n}{4}\right\rceil+1$$

Therefore, by the part (b) of definition of the value of k and the permutation π, we have $a_2+a_3\le k$. Since $k=n-a_2-a_3$, we get $k\ge\left\lceil\frac{n}{2}\right\rceil$. Hence $k=\left\lceil\frac{n}{2}\right\rceil+\varepsilon'$ for some $\varepsilon'\ge0$, and $|M|=\left|\bigcup_{i=1}^{k}X_{\pi(i)}\right|=k+\Delta(\mathcal{X},\pi,k)=\left\lceil\frac{n}{2}\right\rceil+\varepsilon'+\left\lceil\frac{n}{4}\right\rceil+1\ge\left\lceil\frac{3n}{4}\right\rceil+\varepsilon'$.

We have $a_2=\left|\overline{M}\right|-\alpha,\alpha>0$. Therefore,

$$a_3>\left|\overline{M}\right|.$$

Indeed, $a_3=n-k-a_2\ge n-\left\lceil\frac{n}{2}\right\rceil-\varepsilon'-\left|\overline{M}\right|+\alpha\ge\left\lfloor\frac{n}{2}\right\rfloor-\varepsilon'-\left\lfloor\frac{n}{4}\right\rfloor+\varepsilon'+\alpha\ge\left\lceil\frac{n}{4}\right\rceil>\left|\overline{M}\right|$.

Thus, as $a_3\ge\left|\overline{M}\right|\ge|B_3|$, by Lemma 16 , for each $s\le a_3,$ $\Delta(\mathcal{X},\tau,s)\le\left\lfloor\frac{|\overline{M}|+1}{2}\right\rfloor\le\left\lceil\frac{n}{4}\right\rceil+1$. We note that $\Delta(\mathcal{X},\tau,a_3)=|B_3|-a_3<0$. Further, by Lemma 4(b), we get

$$\max_{1\le s\le a_2+a_3}\Delta(\mathcal{X},\tau,s)=\max\{\max_{1\le s\le a_3}\Delta(\mathcal{X},\tau,s),\max_{a_3+1\le s\le a_2+a_3}\Delta(\mathcal{X},\tau,s)\}\le$$

$$\max\{\left\lceil\frac{n}{4}\right\rceil+1,\Delta(\mathcal{X},\tau,a_3)+\max_{1\le s\le a_2}\Delta(\mathcal{A}_2^*,\tau',s)\}\le$$

$$\max\{\left\lceil\frac{n}{4}\right\rceil+1,0+\max_{1\le s\le a_2}\left|\bigcup_{i=1}^{s}X_{\tau(i)}\cap(\overline{M}-B_3)\right|\}\le$$

$$\max\{\left\lceil\frac{n}{4}\right\rceil+1,\left|\overline{M}-B_3\right|\}\le\left\lceil\frac{n}{4}\right\rceil+1$$

because $a_3\ge\left|\overline{M}\right|$ and $B_3\subset\overline{M}$.

Clearly, $\Delta(\mathcal{X}, \tau, a_2+a_3) = |\overline{M}|+a_2-(a_2+a_3) < 0$. For $t = s+a_2+a_3, s \leq k$, we get by Lemma 4(a),

$$\begin{aligned}\Delta(\mathcal{X}, \tau, t) &= \Delta(\mathcal{X}, \tau, a_2+a_3) + \left|\bigcup_{i=a_2+a_3+1}^{t} X_{\tau(i)} - \bigcup_{i=1}^{a_2+a_3} X_{\tau(i)}\right| - s < \\ \left|\bigcup_{i=a_2+a_3+1}^{t} X_{\tau(i)} \cap M\right| - s &= \left|\bigcup_{i=1}^{s} X_{\pi(i)}\right| - s = \Delta(\mathcal{X}, \pi, s) \leq \left\lceil \frac{n}{4} \right\rceil + 1.\end{aligned}$$

We recall that triples not in $\mathcal{A}_2 \cup \mathcal{A}_3$ are in τ ordered the same way as in π and $X_{\pi(s)} \subset M$ for all $s \leq k$. We proved that $\Delta(\mathcal{X}, \pi, s) \leq \lceil \frac{n}{4} \rceil + 1$ for all $1 \leq s \leq n$, which contradicts that $\Delta(\mathcal{X}) > \lceil \frac{n}{4} \rceil + 2$. The proof is complete. ■

Acknowledgements. The authors are indebted to Noga Alon for discussions on expanders and on probabilistic methods, which lead to an improvement of the lower bound. This research of the first author was initiated thanks to the University of Washington – University of Bergen exchange program.

References

[1] M. Bardet, J-C. Faugère, B. Salvy, and P-J. Spaenlehauer, On the complexity of solving quadratic Boolean systems, J. Complexity 29 (2013), 53–75.

[2] N. Courtois, and J. Pierzyk, "Algebraic of block ciphers with over-defined systems of equations" in Advances of Cryptology Asiacrypt 2002, LNCS 2501, Springer, 2002, 267–287.

[3] S. Pissanetzky, Sparse Matrix Technology, electronic edition, 2007. website: http://www.SciControls.com

[4] Y. Saad, Iterative Methods for Sparse Linear Systems, SIAM, 2003, 2nd ed.

[5] I. Semaev, On solving sparse algebraic equations over finite fields, Designs, Codes and Cryptography 49(2008), 47–60.

[6] I. Semaev, New combinatorial problem and evaluation of sparse equations over finite fields, a manuscript.

[7] R. P. Tewarson, and Reginald P., Sparse Matrices (Part of the Mathematics in Science & Engineering series). Academic Press Inc. New York, London, 1973.