Eleventh International Workshop on Algebraic and Combinatorial Coding Theory
June 16-22, 2008, Pamporovo, Bulgaria *pp. 248-254*

On solving sparse algebraic equations over finite fields II. Extended abstract.

Igor Semaev Igor.Semaev@ii.uib.no
Department of Informatics, University of Bergen, NORWAY

1 Introduction

Let F_q be a finite field with q elements and X is a set of variables from F_q of size n. By X_i, $1 \le i \le m$ we denote subsets of X of size $l_i \le l$. Equations

$$f_1(X_1) = 0, \dots, f_m(X_m) = 0 \tag{1}$$

are considered, where f_i are polynomials over F_q and they only depend on variables X_i(l-sparse). We look for all solutions in F_q to (1). So we only consider polynomials of degree at most $q-1$ in each variable. They define mappings from all l_i-tuples over F_q to F_q and any such mapping is represented by a polynomial of degree at most $q-1$ in each variable. The equation $f_i(X_i) = 0$ is determined by (X_i, V_i), where V_i is the set of F_q-vectors in variables X_i, also called X_i-vectors, where f_i is zero. We call (X_i, V_i) a symbol. For $q = 2$ the polynomial f_i is uniquely defined by V_i. Given f_i, the set V_i is computed with q^{l_i} trials.

Deterministic Agreeing-Gluing Algorithm [6] and its average behavior are studied. Assume equiprobable distribution on (1). Given natural numbers m and $l_1, \dots, l_m \le l$, equations in (1) are independent. Each $f_i(X_i) = 0$ is determined by the subset X_i of size l_i taken uniformly at random, that is with the probability $\binom{n}{l_i}^{-1}$, and the mapping f_i taken, independently of X_i, with the probability $q^{-q^{l_i}}$. The running time of the Agreeing-Gluing Algorithm is a random variable.

For fixed q, l and $c \ge 1$ let $\beta = \beta(\alpha)$, where $0 \le \alpha \le l$, be the only root to

$$q^{\beta - \frac{\alpha}{l}} = q e^{g(\alpha)} (1 - \sum_{t=0}^{l} \binom{l}{t} \beta^{l-t} (1-\beta)^t (1 - \frac{1}{q})^{q^t})^{c - \frac{\alpha}{l}},$$

or $\beta(\alpha) = 0$ if there is not any root for some α. Here $g(\alpha) = f(z_\alpha) - \alpha + \alpha \ln \alpha - \frac{\alpha \ln q}{l}$ and $f(z) = \ln(e^z + q^{-1} - 1) - \alpha \ln(z)$, where by z_α we denote the only positive root of the equation $\frac{\partial f}{\partial z}(z) = 0$. We prove

Theorem 1 *Let* $\frac{l_1 + l_2 + \ldots + l_m}{ln}$ *tend to a constant* $c \ge 1$ *as* n *tends to* ∞ *while* $q \ge 2$ *and* $l \ge 3$ *are fixed. Let* $r(q, l, c)$ *be the maximal of* $\max_{0 \le \alpha \le l} q^{\beta(\alpha) - \frac{\alpha}{l}}$

Table 1: Algorithms' running time.

l	3	4	5	6
the worst case	1.324^n	1.474^n	1.569^n	1.637^n
Gluing1, expectation	1.262^n	1.355^n	1.425^n	1.479^n
Gluing2, expectation	1.238^n	1.326^n	1.393^n	1.446^n
Agreeing-Gluing1, expectation	1.113^n	1.205^n	1.276^n	1.334^n

and 1. *Then the expected complexity of the Agreeing-Gluing Algorithm is* $O((r(q,l,c)+\varepsilon)^n)$ *bit operations for any positive real* ε.

For any triple $q, l, c \geq 1$ the Theorem enables estimating the expected running time of the Agreeing-Gluing Algorithm with some mathematical software like Maple. To this end we realize that the equation $\frac{\partial f}{\partial z}(z) = 0$ is equivalent to $\frac{ze^z}{e^z+q^{-1}-1} = \alpha$. So $\alpha = \alpha(z)$ and $\beta = \beta(z)$ are functions in z and $z_\alpha = z$.

For some of $2, l, 1$(e.g. n Boolean equations in n variables each equation depends on l variables) we show the data obtained in Table 1 with the expected complexities of the Gluing1 and Gluing2 Algorithms from our previous work [7]. Agreeing-Gluing1 Algorithm is a variant of the Agreeing-Gluing Algorithm with the same asymptotical running time and polynomial in n memory requirement. In case $q = 2$ each instance of (1) may be encoded with a CNF formula in the same set of variables and of clause length at most l [7]. So l-SAT solving algorithms provide with the worst case complexity estimates, see [2], in the first line. We remark an exciting difference in the worst case complexity and expected complexity of the Agreeing-Gluing Algorithm. It is quite obvious that average instances of the l-SAT problem and that of (1) are different. That gives insight into why the expected complexity is so low in comparison with the worst case. The Agreeing-Gluing family algorithms seem better on sparse equation systems (1) than Gröbner Basis related algorithms, see conjectured estimates in [9].

This article was motivated by applications in cryptanalysis. Mappings implemented by modern ciphers are compositions of functions in small number of variables. Intermediate variables are introduced to simplify equations, describing the cipher, and get a system of sparse equations. We are studying an approach which exploits the sparsity of equations and doesn't depend on their algebraic degree. This approach was independently discovered in [10] and [5], where the Agreeing procedure(called local reduction in [10]) was described for the first time. The term Agreeing itself comes from [6]. No asymptotical estimates for that type of algorithms were given in [10, 5, 6]. We recommend to look also through our previous work [7], where some necessary basic facts were proved.

This is the extended abstract of [8]. The author is grateful to H.Raddum for careful reading the work and numerous remarks.

2 Gluing procedure and Gluing Algorithm

For symbols (X_i, V_i) for $i = 1, 2$, one defines $Z = X_1 \cup X_2$ and $Y = X_1 \cap X_2$ and the set of Z-vectors $U = \{(a_1, b, a_2) : (a_1, b) \in V_1, (b, a_2) \in V_2\}$. Here a_i is an $(X_i \backslash Y)$-vector and b is a Y-vector. We denote $(a_1, b, a_2) = (a_1, b)\circ(b, a_2)$ and say that (a_1, b, a_2) is the gluing of (a_1, b) and (b, a_2). To glue (X_1, V_1) and (X_2, V_2) one can sort V_1 or V_2 by Y-subvectors and only glues vectors with the same Y-subvector. So the complexity of the gluing is $O(|U| + (|V_1| + |V_2|)\log(|V_i|))$ operations. We use a simpler bound $O(|V_1||V_2| + |V_1| + |V_2|)$ in what follows. Denote $(Z, U) = (X_1, V_1) \circ (X_2, V_2)$.

Gluing Algorithm

input: the system (1) represented by symbols (X_i, V_i), where $1 \le i \le m$.

output: the set U of all solutions to (1) in variables $X(m) = X_1 \cup \ldots \cup X_m$.

put $(Z, U) \leftarrow (X_1, V_1)$ **and** $k \leftarrow 2$,

while $k \le m$ **do** $(Z, U) \leftarrow (Z, U) \circ (X_k, V_k)$ **and** $k \leftarrow k + 1$,

return (Z, U).

The set U is all solutions to (1) in variables $X(m)$. The Gluing Algorithm takes $O(\sum_{k=1}^{m-1} |U_k| + m)$ operations with F_q-vectors of length at most n, where q and l are fixed, and n or m may grow. The memory requirement is of the same magnitude. Here $(X(k), U_k) = (X_1, V_1) \circ \ldots \circ (X_k, V_k)$. The set U_k consists of all solutions to the first k equations in variables $X(k) = X_1 \cup \ldots \cup X_k$. The sequence of $|U_k|$ fully characterizes the running time of the algorithm. The asymptotical analysis of $|U_k|$ is done in [7] using Random Allocations Theory results found in [4, 3, 1]. Two technical statements from [7] are formulated here.

Lemma 1 *(Lemma 4 in [7]) Let the subsets of variables $X_1, \ldots, X_k$ be fixed while $f_1, \ldots, f_k$ are randomly chosen according to our model. Then the expected number of solutions to the first k equations in (1) is $E_{f_1,\ldots,f_k}|U_k| = q^{|X(k)|-k}$.*

Lemma 2 *(Lemma 5 in [7]) Let $L_k = l_1 + \ldots + l_k$ and $\alpha = L_k/n$, and $k \le n$. Let $0 < \delta < 1$ be fixed as n tends to ∞. Then $E|U_k|$, the expected number of solutions to the first k equations, is $< q^{n^\delta}$, if $L_k < n^\delta$, and $O((qe^{g(\alpha)} + \epsilon)^n)$ otherwise for any positive real number ϵ. Here $g(\alpha) = f(z_\alpha) - \alpha + \alpha \ln \alpha - \frac{\alpha \ln q}{l}$ and $f(z) = \ln(e^z + q^{-1} - 1) - \alpha \ln(z)$, where by z_α we denote the only positive root of the equation $\frac{\partial f}{\partial z}(z) = 0$.*

3 Agreeing procedure and Agreeing-Gluing Algorithm

For symbols (X_i, V_i) for $i = 1, 2$, one defines $Y = X_1 \cap X_2$. Let $V_{1,2}(V_{2,1})$ be the set of Y-subvectors of $V_1(V_2)$. We say the symbols (X_1, V_1) and (X_2, V_2) agree if $V_{1,2} = V_{2,1}$. Otherwise, we apply the procedure called agreeing. We delete from V_i all vectors whose Y-subvectors are not in $V_{2,1} \cap V_{1,2}$. So new symbols (X_i, V_i') are determined, where $V_i' \subseteq V_i$ consist of the vectors in V_i survived after agreeing. To agree (X_1, V_1) and (X_2, V_2) one sorts V_1 or V_2 by Y-subvectors and do agreeing by table look ups. So the complexity of the agreeing is at most$O((|V_1| + |V_2|) \log(|V_i|))$ operations. The following Agreeing-Gluing Algorithm combines the Agreeing and Gluing procedures to solve (1).
Agreeing-Gluing Algorithm

input: the system (1) represented by symbols (X_i, V_i), where $1 \le i \le m$.
output: the set U of all solutions to (1) in variables $X(m) = X_1 \cup \ldots \cup X_m$.
put $(Z, U) \leftarrow (X_1, V_1)$ **and** $k \leftarrow 2$,
while $k \le m$ **do** $s \leftarrow k$,
  while $s \le m$ **agree** (Z, U) **and** (X_s, V_s)**, put** $s \leftarrow s + 1$,
put $(Z, U) \leftarrow (Z, U) \circ (X_k, V_k)$ **and** $k \leftarrow k + 1$**,**
return (Z, U).

Assume $(X(0), U_0')$ trivial. For any $0 \le k < m$ let $(X(k+1), U_{k+1}')$ denote the symbol $(X(k), U_k') \circ (X_{k+1}, V_{k+1})$ after agreeing with $(m - k - 1)$ symbols (X_i, V_i), where $k + 1 < i \le m$. The Agreeing-Gluing Algorithm produces the sequence of $(X(k), U_k')$ and takes

$$O(m(\sum_{k=1}^{m-1} |U_k'| + 1)) \tag{2}$$

operations with F_q-vectors of length at most n, where q and l are fixed, and n or m may grow. (2) incorporates the cost of the gluing $(X(k), U_k') \circ (X_{k+1}, V_{k+1})$, which is $O(|U_k'|)$ operations, and the agreeing the resulting set of $X(k+1)$-vectors, of size at most $O(|U_k'|)$, with the rest $m-k-1$ symbols. In our setting $|U_k'|$ is a random variable. We estimate the expectation of $|U_k'|$ in Section 4, see Theorem 2. That will imply Theorem 1. From the definition of Gluing and Agreeing procedures we get:

Lemma 3 $(X(k), U_k')$ *is the symbol* $(X(k), U_k) = (X_1, V_1) \circ \ldots \circ (X_k, V_k)$ *after agreeing with* $(m - k)$ *symbols* (X_i, V_i)*, where* $k < i \le m$.

The space requirement of the Algorithm is as its running time. The Agreeing-Gluing1 Algorithm, similar to the Gluing1 Algorithm of [7], requires polynomial memory with the same running time. We do not go into detail here.

4 Complexity analysis of the Agreeing-Gluing Algorithm

We prove Theorem 1. Let $Z, X_1, \ldots, X_k$ be fixed subsets of variables and U be a fixed set of Z-vectors, so that (Z, U) is defined by an equation $f(Z) = 0$. Let V_i be the set of X_i-vectors, solutions to independent equations $f_i(X_i) = 0$ generated uniformly at random.

Lemma 4 *Let (Z, U') be produced from (Z, U) by agreeing with all (X_i, V_i). Then the expectation of $|U'|$ is given by $E_{f_1,\ldots,f_k}|U'| = |U|\prod_{i=1}^{k}(1-(1-\frac{1}{q})^{q^{|X_i\setminus Z|}})$, where $|X_i \setminus Z|$ stands for the number of variables X_i not occurring in Z.*

Proof. Assume $k = 1$. Let $Y_1 = Z \cap X_1$ and $|U| = \sum_a |U_a|$, where U_a is the subset of U-vectors whose projection to variables Y_1 is a. Similarly, $V_{1,a}$ is the subset of V_1-vectors whose projection to variables Y_1 is a. Then $|U'| = \sum_a |U_a| I_a$, where $I_a = 1$ for $V_{1,a} \neq \emptyset$ and $I_a = 0$ for $V_{1,a} = \emptyset$. Let W_a be the subset of all vectors in variables X_1 whose projection to variables Y_1 is a. We see that $|W_a| = q^{|X_1\setminus Y_1|}$. One computes $Pr(V_{1,a} = \emptyset) = Pr(f_1 \neq 0 \,\mathrm{on}\, W_a) = (1-\frac{1}{q})^{q^{|X_1\setminus Y_1|}}$. So $E_{f_1}(I_a) = 1-(1-\frac{1}{q})^{q^{|X_1\setminus Y_1|}} = 1-(1-\frac{1}{q})^{q^{|X_1\setminus Z|}}$. Then $E_{f_1}|U'| = \sum_a |U_a| E_{f_1}(I_a) = |U|(1-(1-\frac{1}{q})^{q^{|X_1\setminus Z|}})$. This proves the statement for $k = 1$. The Lemma is now shown true by induction.

Corollary 1 *Let f be generated independently to f_i. Then $E_{f,f_1,\ldots,f_k}|U'| = E_f|U|\prod_{i=1}^{k}(1-(1-\frac{1}{q})^{q^{|X_i\setminus Z|}})$.*

We will use the Corollary in order to estimate the expectation of $|U'_k|$.

Lemma 5 *Let $0 \leq \beta \leq 1$ be any number. Then*

$$E|U'_k| \leq q^{\beta n-k} + \sum_{|Z|>\beta n} Pr(X(k) = Z)\, q^{|Z|-k} \prod_{i=k+1}^{m} E_{X_i}(1-(1-\frac{1}{q})^{q^{|X_i\setminus Z|}}), \quad (3)$$

where Z runs over all subsets of X of size $> \beta n$.

Proof. For fixed X_i and random f_i, and by Lemma 3 and Corollary 1 we have

$$E_{f_1,\ldots,f_m}|U'_k| = q^{|X(k)|-k} \prod_{i=k+1}^{m} (1-(1-\frac{1}{q})^{q^{|X_i\setminus X(k)|}}), \quad (4)$$

as $E_{f_1,\ldots,f_k}|U_k| = q^{|X(k)|-k}$ by Lemma 2. Let We study the expectation of $|U'_k|$ when X_i are random too. So

$$E|U'_k| = \sum_{Z\subseteq X} Pr(X(k) = Z)\, q^{|Z|-k} \prod_{i=k+1}^{m} E_{X_i}(1-(1-\frac{1}{q})^{q^{|X_i\setminus Z|}})$$

We partition the last sum for $|Z| \le \beta n$ and $|Z| > \beta n$, and get the statement. In next three Lemmas(without proof here) we estimate the expectation

$$E_{X_i}(1-(1-\frac{1}{q})^{q^{|X_i \setminus Z|}}). \tag{5}$$

Lemma 6 *Let $Z \subseteq X$ be a fixed subset of variables. Then (5) only depends on the size of Z and doesn't depend on the set itself. The expectation is not decreasing as $|Z|$ is decreasing or $|X_i|$ is increasing.*

Lemma 7 *Let Z be a fixed u-subset of X and X_i be an l_i-subset of X taken uniformly at random. Then $Pr(|X_i \setminus Z| = t) = \frac{\binom{u}{l_i-t}\binom{n-u}{t}}{\binom{n}{l_i}}$.*

Lemma 8

1. *Let $|Z| > \beta n$, where $0 \le \beta \le 1$ is fixed as n tends to ∞, then (5) is bounded by $F(\beta) + O(\frac{1}{n})$, where $O(\frac{1}{n})$ doesn't depend on i.*
2. *The function $F(\beta) = 1 - \sum_{t=0}^{l}\binom{l}{t}\beta^{l-t}(1-\beta)^t(1-\frac{1}{q})^{q^t}$ is not increasing in $0 \le \beta \le 1$ and $\frac{1}{q} \le F(\beta) \le 1-(1-\frac{1}{q})^{q^l} < 1$.*

The inequality (3) then implies

$$E|U'_k| \le q^{\beta n-k} + E_{X_1,\dots,X_k}(q^{|X(k)|-k})\,(F(\beta)+\varepsilon)^{m-k}. \tag{6}$$

for any positive real ε as n tends to ∞. For $0 \le \alpha \le l$ we define the function $0 \le \beta(\alpha) \le 1$ by the rule: $\beta = \beta(\alpha)$ is the solution of the equation

$$q^{\beta-\frac{\alpha}{l}} = qe^{g(\alpha)}F(\beta)^{c-\frac{\alpha}{l}} \tag{7}$$

if such a solution exists and $\beta(\alpha) = 0$ otherwise. We know that $c_n = \frac{l_1+l_2+\ldots+l_m}{ln}$ tends to a constant $c \ge 1$ as n tends to ∞ while q and l are fixed.

Theorem 2

1. *The equation (7) has at most one solution for any $0 \le \alpha \le l$.*
2. *Let $L_k = l_1+\ldots+l_k$ and $\alpha = L_k/n$, and $k \le n$. Let $0 < \delta < 1$ be fixed as n tends to ∞. Then*

$$E|U'_k| = \begin{cases} < q^{n^\delta}, & \text{if } \; L_k < n^\delta; \\ O((q^{\beta(\alpha)-\frac{\alpha}{l}}+\varepsilon)^n), & \text{if } \; ln > L_k \ge n^\delta; \\ < 1, & \text{if } \; L_k \ge ln, \end{cases}$$

for any positive real ε.

Proof. We prove the second statement here. It is true for $L_k < n^\delta$ and $L_k \geq ln$. Let $ln > L_k \geq n^\delta$. Then by Lemma 2 we get from (6) that

$$E|U'_k| \leq (q^{\beta - \frac{\alpha}{l}})^n + O((qe^{g(\alpha)} + \varepsilon)^n (F(\beta) + \varepsilon)^{\frac{m-k}{n} n}),$$

as $\frac{\alpha}{l} \leq \frac{k}{n}$ and for any positive ε. We realize that $\frac{m-k}{n} \geq c_n - \frac{\alpha}{l}$, so

$$E|U'_k| \leq (q^{\beta - \frac{\alpha}{l}})^n + O((qe^{g(\alpha)} F(\beta)^{c - \frac{\alpha}{l}} + \varepsilon)^n) \quad (8)$$

for any real positive ε as n tends to ∞. If (7) has one solution, then the inequality $E|U'_k| = O((q^{\beta(\alpha) - \frac{\alpha}{l}} + \varepsilon)^n)$ follows from (8) and (7). When (7) has no solutions, the statement is easy. The Theorem is proved.

The main Theorem 1 now follows from Theorem 2 and formula (2).

References

[1] V. P. Chistyakov, Discrete limit distributions in the problem of shots with arbitrary probabilities of occupancy of boxes, *Mat. Zametki* 1, 1967, 9-16.

[2] K. Iwama, Worst-case upper bounds for kSAT, *The Bull. EATCS* 82, 2004, 61-71.

[3] V. Kolchin, The rate of convergence to limit distributions in the classical problem of shots, *Teoriya veroyatn. i yeye primenen.*, 11, 1966, 144-156.

[4] V. Kolchin, A. Sevast'yanov, V. Chistyakov, *Random allocations,* John Wiley & Sons, 1978.

[5] H. Raddum, Solving non-linear sparse equation systems over $GF(2)$ using graphs, Univ. Bergen, preprint, 2004.

[6] H. Raddum, I. Semaev, New technique for solving sparse equation systems, Cryptology ePrint Archive, 2006/475.

[7] I. Semaev, On solving sparse algebraic equations over finite fields, to appear in *Des., Codes Crypt.*, extended abstract in Proc. WCC'07, Versailles, France, INRIA, 361-370.

[8] I. Semaev, On solving sparse algebraic equations over finite fields II, Cryptology ePrint Archive: Report 2007/280.

[9] B.-Y. Yang, J-M. Chen, N. Courtois, On asymptotic security estimates in XL and Gröbner bases-related algebraic cryptanalysis, ICICS 2004, *Lect. Notes Comp. Sci.* 3269, Springer-Verlag, 2004, 401-413.

[10] A. Zakrevskij, I. Vasilkova, Reducing large systems of Boolean equations, *4th Intern. Workshop Bool. Probl.*, Freiberg Univ., 2000.