Федеральный исследовательский центр Информатика и Управление Российской академии наук

RUDN university

РЭУ.РФ Российский экономический университет имени Г.В. Плеханова

# КОМПЬЮТЕРНАЯ АЛГЕБРА

Материалы 6-й международной конференции

*Москва, 23–25 июня 2025 г.*

# COMPUTER ALGEBRA

**6th International Conference Materials**

***Moscow, June 23–25, 2025***

Moscow
2025

UDC 519.6(063)
BBC 22.19;431

*Editors:* PhD Anna A. Ryabenko, DSc Dmitry S. Kulyabov

*Reviewers:* PhD A. V. Nesterov, PhD K. P.Lovetskiy

Printed in author's edition

**Computer algebra:** 6th International Conference Materials. Moscow, 23–25 June, 2025 / ed. A. A. Ryabenko, D. S. Kulyabov. Moscow: RUDN University.

DOI: 10.22363/12585-2025-6-000.
EDN: DZBCNS.

The international conference is organized by Federal Research Center "Computer Science and Control" of RAS, Peoples' Friendship University of Russia and Plekhanov Russian University of Economics. The talks presented at the conference discuss actual problems of computer algebra — the discipline whose algorithms are focused on the exact solution of mathematical and applied problems using a computer.

For scientists, graduate and undergraduate students in mathematics, physics and computer science.

ISBN 978-5-209-12585-3

Scientific publication

# The Generic-Case Complexity of Finding a Binary Solution to a System of Linear Equations

Alexandr V. Seliverstov

*Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), 19, Bolshoy Karetny per., Moscow, 127051, Russian Federation*

**Abstract**

Using the Schwartz–Zippel lemma, we estimate the generic-case complexity of finding a binary solution to a system of linear equations over an infinite field. Compared to previous works, the new algorithm is applicable to systems with a smaller number of linear equations, but it requires more running time. The algorithm works over a field of any characteristic. If the system has many binary solutions, then our algorithm is not applicable. So, in the worst case, there is no polynomial upper bound on the computational complexity.

**Keywords**

boolean programming, knapsack, matrix rank, polynomial, heuristics, computational complexity

## 1. Introduction

Let us consider the recognition problem whether there is a $\{0,1\}$-solution to a system of linear equations. The problem is NP-complete not only over the ring of integers, but also over the field of residues modulo any odd prime. Over the ring of integers, under a constraint on the size of the coefficients, a heuristic polynomial-time algorithm is known based on finding the shortest nonzero vector in an integer lattice [5]. The problem is also related to some optimization problems known as the multidimensional knapsack problem and Boolean programming, refer to the review [3] and recent article [1]. Constraints for dimensionality reducing of the problem by means of projection onto a coordinate hyperplane are known [2]. On the other hand, some systems of quadratic equations have been considered recently [4].

Over an arbitrary field $K$ of characteristic $\mathrm{char}(K) \neq 2$, for almost all systems having at least $n - \sqrt{2n - o(n)}$ linear equations in $n$ variables, a heuristic polynomial-time algorithm had been proposed several years ago [9]. In this work, the restriction on the number of equations is significantly relaxed, although the computational complexity increases. We consider systems over an arbitrary computable field. The computational complexity is estimated by the number of algebraic operations over the field.

For a recognition problem, let us assume three possible answers: the input may not only be accepted or rejected, but also an explicit notification of uncertainty of the choice is possible. In any case, the answer must be obtained in a finite time and without errors, and if an easily verifiable condition is met, then the notification of uncertainty can be issued only for a small fraction of inputs among all inputs of a given length. Such algorithms are called *generic* [6] or *errorless heuristics*. It is known that an NP-complete problem can be split into several subproblems that are also NP-complete [8]. Generic algorithms for an NP-complete problem can be considered for determining NP-complete subproblems for which the generic algorithm gives an uncertain answer.

$6^{th}$ *International Conference "Computer Algebra", Moscow, June 23–25, 2025*
10.22363/12585-2025-6-022
EDN: EZKBKJ
slvstv@iitp.ru (A. V. Seliverstov)
0000-0003-4746-6396 (A. V. Seliverstov)

To estimate the number of inputs of a given size on which the algorithm quickly makes the correct decision, we use the Schwartz–Zippel lemma [7].

**Lemma 1.** *Given a non-constant polynomial $f(x_1, \ldots, x_n)$ of degree $d$ over a field $K$. If random variables $\xi_1$, ..., $\xi_n$ are independent and uniformly distributed on a finite set $S \subseteq K$ of cardinality $|S|$, then the inequality*

$$\text{Prob}\left[f(\xi_1, \ldots, \xi_n) = 0\right] \leq \frac{d}{|S|}$$

*holds, where* Prob$[\cdot]$ *denotes the probability of the condition indicated in square brackets.*

## 2. Results

Let us consider a system of $m$ linear equations in $n > m$ variables:

$$\begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n + 1 = 0 \\ \cdots \\ \alpha_{m1}x_1 + \cdots + \alpha_{mn}x_n + 1 = 0 \end{cases}.$$

Multiplying each linear equation by each of the variables and taking into account the equalities $x_k^2 = x_k$, which are satisfied with $\{0,1\}$-solutions, we obtain $mn$ new equations of the second degree. Discarding the terms depending only on one variable, we obtain a set of $mn$ bilinear forms, the coefficients of which form a matrix $W$. The rows correspond to the bilinear forms, and the columns correspond to monomials of the form $x_jx_k$ for $j < k$.

**Lemma 2.** *Let the matrix $W$ be computed for $m$ linear equations in $n$ variables over a purely transcendental extension of the field $K$, where all coefficients $\alpha_{ij}$ are algebraically independent of each other. The rank of the matrix satisfies the inequality*

$$\text{rank}(W) \geq mn - \frac{m(m+1)}{2}.$$

**Example 1.** For $n = 3$ and $m = 1$, the $3 \times 3$ matrix

$$W = \begin{pmatrix} \alpha_{12} & \alpha_{13} & 0 \\ \alpha_{11} & 0 & \alpha_{13} \\ 0 & \alpha_{11} & \alpha_{12} \end{pmatrix}$$

is degenerate over a field of characteristic $\text{char}(K) = 2$ because $\det(W) = -2\alpha_{11}\alpha_{12}\alpha_{13}$. Next, for $n = 5$ and $m = 2$, the $10 \times 10$ matrix $W$ is degenerate over any field because $\text{rank}(W) \leq 9$. For $n = 7$ and $m = 3$, the $21 \times 21$ matrix $W$ is also degenerate over any field because $\text{rank}(W) \leq 18$. (The rank is computed with SymPy.)

Let the number of equations $m$ be such that $mn \geq \text{rank}(W) + n - m$. This inequality holds for $m > n/2$. But a smaller number $m$ is sufficient because the rank of $W$ is small. In the general case, $n$ linearly independent linear equations can be derived from resulting quadratic equations as well as the initial linear equations. In particular, there are new linear equations. Next, using these $n$ linearly independent linear equations, one can find a solution and check whether it consists of zeros and ones. Of course, the method is not applicable when the system has many $\{0,1\}$-solutions. Thus, we have a polynomial upper bound on the generic-case complexity, but we have nothing in the worst case.

**Example 2.** Let us consider a linear equation in two variables $\alpha x_1 + \beta x_2 + 1 = 0$. Multiplying this equation by each of the variables and taking into account the equalities $x_k^2 = x_k$, which are satisfied with $\{0,1\}$-solutions, we obtain two equations:

$$\begin{cases} \beta x_1x_2 + (1+\alpha)x_1 = 0 \\ \alpha x_1x_2 + (1+\beta)x_2 = 0 \end{cases}.$$

This yields the linear equation $\alpha(1+\alpha)x_1 = \beta(1+\beta)x_2$. For $\alpha = \beta = -1$, this equation turns into the identity. But in the general case, it is a new linear equation, which is linearly independent of original one. So, one can either find the $\{0,1\}$-solution or prove its absence.

The probability of success is equal to the probability that the determinant of an $n \times n$ matrix does not vanish. A bound can be obtained using the Schwartz–Zippel lemma, i. e., Lemma 1. Let $K$ denote a field and $\varepsilon$ denote a positive real parameter.

**Theorem.** *For our algorithm, there is an univariate function $f(n)$ so that if $n$ is even, $m \geq n/2$, and the coefficients $\alpha_{ij}$ are uniformly and independently distributed on the set $S \subset K$ of cardinality $\lceil f(n)/\varepsilon \rceil$, then the upper bound on the probability of the uncertain answer equals $\varepsilon$. The generic-case complexity is equal to the complexity of finding the rank of $W$.*

*Proof.* Let us consider the special system of exactly $n/2$ linear equations, where the $k$-th equation depends on two variables $x_{2k-1}$ and $x_{2k}$:

$$\begin{cases} \beta_1 x_1 & +\beta_2 x_2 + 1 & = & 0 \\ \cdots & \cdots & & \\ \beta_{2k-1} x_{2k-1} & +\beta_{2k} x_{2k} + 1 & = & 0 \\ \cdots & & & \\ \beta_{n-1} x_{n-1} & +\beta_n x_n + 1 & = & 0 \end{cases}.$$

As in Example 2, new linear equations are

$$\begin{cases} \beta_1(1+\beta_1)x_1 & -\beta_2(1+\beta_2)x_2 & = & 0 \\ \cdots & \cdots & & \\ \beta_{2k-1}(1+\beta_{2k-1})x_{2k-1} & -\beta_{2k}(1+\beta_{2k})x_{2k} & = & 0 \\ \cdots & \cdots & & \\ \beta_{n-1}(1+\beta_{n-1})x_{n-1} & -\beta_n(1+\beta_n)x_n & = & 0 \end{cases}.$$

They together compose a system of $n$ linear equations in $n$ variables. Let us denote by $M$ the $n \times n$ matrix of linear term coefficients. The determinant is a polynomial in coefficients $\beta_1$, ..., $\beta_n$ of the initial system:

$$\det(M) = \pm \prod_{k=1}^{n/2} \beta_{2k-1}\beta_{2k}(2 + \beta_{2k-1} + \beta_{2k}),$$

where the sign depends on the order of equations. If all coefficients of the initial system are nonzero and inequalities $\beta_{2k-1} + \beta_{2k} \neq -2$ holds, then $\det(M)$ does not vanish.

In fact, the matrix $M$ is not unique. Its entries are rational functions in $\beta_1$, ..., $\beta_n$. But one can compute $M$ so that all entries are polynomials.

In the general case, using the first $n/2$ linear equations, our algorithm produces a nondegenerate system of $n$ linear equations in $n$ variables. It has unique solution. Let us consider the $n \times n$ matrix $M$ of linear term coefficients. Its entries are rational functions in coefficients $\alpha_{ij}$ of the initial system. Without loss of generality, let the entries be polynomials. Thus, $\det(M)$ is a polynomial too. Moreover, the polynomial does not vanish identically. Let the degree upper bound be the desired function $f(n)$. In accordance with Lemma 1, the determinant does not vanish for almost all coefficients, i. e., our algorithm fails with probability at most $\varepsilon$.

## 3. Discussion

In accordance with Lemma 2, in the general case, such bounds based on the Schwartz–Zippel lemma cannot be significantly improved without increasing runtime. However, such an improvement is possible for sparse systems of equations with a fixed arrangement of nonzero coefficients.

The algorithm can be useful over a finite field too, although the Schwartz–Zippel lemma requires sufficiently many elements depending on the number of variables. Of course, if $K$ is infinite, then a sufficiently large set $S \subset K$ exists for all $n$ and $\varepsilon$.

Of course, if the number of variables is sufficiently large, then the worst-case computational complexity remains high. Nevertheless, in accordance with our result, the Merkle–Hellman cryptosystem based on the subset sum problem can be broken in almost all cases by means of a broadcast attack against it, refer to [5]. Many related problems can also be reduced to the problem under consideration. Our algorithm can also be considered as method to compute the Gröbner basis of some zero-dimensional ideal in the ring of multivariate polynomial.

**Author Contributions:** Conceptualization and writing, Alexandr V. Seliverstov. The author has read and agreed to the published version of the manuscript.

**Funding:** The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation for IITP RAS.

**Data Availability Statement:** No new data were created or analyzed during this study. Data sharing is not applicable.

**Conflicts of Interest:** The author declares no conflict of interest.

**Acknowledgments:** We thank the conference organizers.

## References

[1] G. Alonistiotis, A. Antonopoulos, N. Melissinos, A. Pagourtzis, S. Petsalakis, M. Vasilakis, Approximating subset sum ratio via partition computations, Acta Informatica 61 (2024) 101–113. doi:10.1007/s00236-023-00451-7.

[2] A. A. Boykov, A. V. Seliverstov, On a cube and subspace projections, Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki 33 (2023) 402–415. doi:10.35634/vm230302.

[3] V. Cacchiani, M. Iori, A. Locatelli, S. Martello, Knapsack problems — an overview of recent advances. part II: Multiple, multidimensional, and quadratic knapsack problems, Computers and Operations Research 143 (2022) 1–14. doi:10.1016/j.cor.2021.105693.

[4] Y. G. Evtushenko, A. A. Tret'yakov, Exact formula for solving a degenerate system of quadratic equations, Computational Mathematics and Mathematical Physics 64 (2024) 365–369. doi:10.1134/S0965542524030072.

[5] Y. Pan, F. Zhang, Solving low-density multiple subset sum problems with SVP oracle, Journal of Systems Science and Complexity 29 (2016) 228–242. doi:10.1007/s11424-015-3324-9.

[6] A. N. Rybalov, Generic polynomial algorithms for the knapsack problem in some matrix semigroups, Siberian Electronic Mathematical Reports 20 (2023) 100–109. URL: http://semr.math.nsc.ru/v20/n1/p100-109.pdf. doi:10.33048/semi.2023.20.009.

[7] J. T. Schwartz, Fast probabilistic algorithms for verification of polynomial identities, J. ACM 27 (1980) 701–717. doi:10.1145/322217.322225.

[8] L. Zhang, M. Quweider, F. Khan, H. Lei, Splitting NP-complete sets infinitely, Information Processing Letters 186 (2024) 1–7. doi:doi.org/10.1016/j.ipl.2024.106472.

[9] O. A. Zverkov, A. V. Seliverstov, Effective lower bounds on the matrix rank and their applications, Programming and Computer Software 49 (2023) 441–447. doi:10.1134/S0361768823020160.