

# Asymptotic results for the number of Wagner's solutions to a generalised birthday problem

Alexey Lindo and Serik Sagitov

*Chalmers University of Technology and University of Gothenburg*

## Abstract

We study two functionals of a random matrix  $\mathbf{A}$  with independent elements uniformly distributed over the cyclic group of integers  $\{0, 1, \dots, M-1\}$  modulo  $M$ . One of them,  $V_0(\mathbf{A})$  with mean  $\mu$ , gives the total number of solutions for a generalised birthday problem, and the other,  $W(\mathbf{A})$  with mean  $\lambda$ , gives the number of solutions detected by Wagner's tree based algorithm.

We establish two limit theorems. Theorem 2.1 describes an asymptotical behaviour of the ratio  $\lambda/\mu$  as  $M \rightarrow \infty$ . Theorem 2.2 suggests Chen-Stein bounds for the total variation distance between Poisson distribution and distributions of  $V_0$  and  $W$ .

2010 Mathematics Subject Classification: 60B20, 60C05, 60F05

*Keywords:* Chen-Stein's method, Functionals of random matrices

## 1 Introduction

Let  $(N, M, L)$  be three natural numbers larger than or equal to 2. Assume that we have a random matrix

$$\mathbf{A} = (a_{ij}), \quad 1 \leq i \leq L, \quad 1 \leq j \leq N \quad (1)$$

with independent elements  $a_{ij}$  which are uniformly distributed on  $\{0, 1, \dots, M-1\}$ . Let  $\mathbf{J} = \{1, \dots, L\}^N$  be the set of matrix positions, so that  $|\mathbf{J}| = L^N$ . For each  $b \in \{0, 1, \dots, M-1\}$ , define  $V_b \equiv V_b(\mathbf{A})$  as the number of vectors  $\mathbf{i} = (i_1, \dots, i_N) \in \mathbf{J}$  with

$$a_{i_1,1} + \dots + a_{i_N,N} \stackrel{M}{=} b,$$

where the sign  $\stackrel{M}{=}$  means equality modulo  $M$ . Clearly,  $\sum_{b=0}^{M-1} V_b = L^N$ , so that by the assumption of uniform distribution,

$$\mu := E(V_0) = L^N M^{-1}.$$

The problem of finding all  $V_0$  zero-sum vectors

$$\mathbf{a}_{\mathbf{i}} = (a_{i_1,1}, \dots, a_{i_N,N}), \quad \mathbf{i} = (i_1, \dots, i_N) \in \mathbf{J} \quad (2)$$

for a given matrix  $\mathbf{A}$ , can be viewed as a generalised birthday problem. It arises naturally in a variety of situations including cryptography, see [7] and reference therein; ring linear codes [3]; abstract algebra, where in the theory of modules it is related to the notion of annihilator, see e.g. [4]. This problem can be solved only by exhaustive search and is *NP*-hard [6]. Wagner [7] proposed a subexponential algorithm giving hope to quickly detect at least some of the solutions to this kind of problems.

Assume that  $N = 2^n$ ,  $n \geq 1$  and  $M = 2^m + 1$ ,  $m \geq n$ . It will be convenient to use the symmetric form

$$D_m := \{-2^{m-1}, \dots, -1, 0, 1, \dots, 2^{m-1}\}$$

of  $\{0, 1, \dots, M-1\}$  as the set of possible values for  $a_{ij}$ . Wagner's algorithm has a binary tree structure, see Figure 1, starting from  $N$  leaves at level  $n$  and moving toward the top of the tree at level 0. For a given a vector  $\mathbf{x} = (x_1, \dots, x_{2^n})$  with  $x_j \in D_m$  the algorithm searches for the value

$$H_n(\mathbf{x}) := x_1^{(n)} \in D_{m-n} \cup \{\Delta\}, \quad (3)$$

obtained recursively in a way explained next (the special state  $\Delta$  indicates that the algorithm is terminated and a solution is not found). Put  $x_j^{(0)} \equiv x_j$ . For  $h = 1, \dots, n$

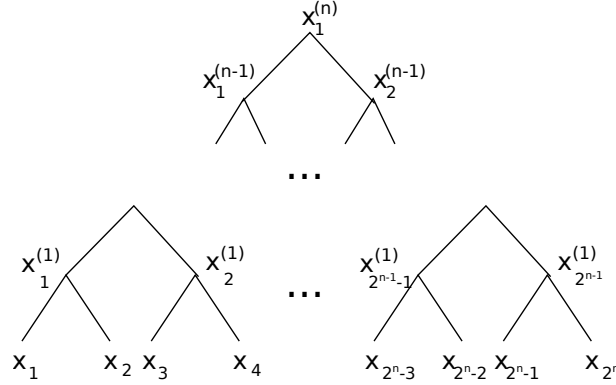


Figure 1: Wagner's algorithm

and  $j = 1, \dots, 2^{n-h}$ , let  $x_j^{(h)} = b$  if there exists such a  $b \in D_{m-h}$  that

$$x_{2j-1}^{(h-1)} + x_{2j}^{(h-1)} \stackrel{M}{=} b,$$

and put  $x_j^{(h)} = \Delta$  otherwise. In particular, if  $x_k^{(h-1)} = \Delta$  for at least one of the two indices  $k \in \{2j-1, 2j\}$ , then  $x_j^{(h)} = \Delta$ .

A vector  $\mathbf{x}$  will be called a Wagner's solution to the generalised birthday problem, if  $H_n(\mathbf{x}) = 0$ . The total number  $W \equiv W(\mathbf{A})$  of Wagner's solutions among the vectors (2) has mean

$$\lambda := E(W) = L^N p_{n,m},$$

where

$$p_{n,m} := P(H_n(\mathbf{a}_i) = 0), \quad \mathbf{i} \in \mathbf{J}.$$

The proportion of Wagner's solutions can be characterised by the ratio of the means

$$R_{n,m} := \lambda/\mu = (2^m + 1)p_{n,m}. \quad (4)$$

Clearly,  $R_{n,m}$  is the conditional probability of a given zero-sum random vector to be Wagner's solution.

There is a growing number of papers studying the properties of various tree based algorithms with some of them, in particular [5], suggesting further developments of Wagner's approach. The main results of this paper are stated in the next section. Theorem 2.1 gives an integral recursion for calculating the limit for the key ratio (4). Theorem 2.2 suggests Chen-Stein bounds for the total variation distance between Poisson distribution and distributions of  $V_0$  and  $W$ . (Among related results concerning speed of convergence for functional of random matrices over finite algebraical structures we can only name a recent paper [2].)

## 2 Main results

Define a sequence of polynomials  $\{\phi_n(x)\}_{n \geq 1}$  by

$$\phi_n(x) := \int_0^x \phi_{n-1}(u)\phi_{n-1}(x-u)du + 2 \int_x^{2^{-n}} \phi_{n-1}(u)\phi_{n-1}(u-x)du, \quad (5)$$

with  $\phi_1(x) \equiv 1$ .

**Theorem 2.1.** *For any fixed natural number  $n$ ,*

$$R_{n,m} \rightarrow \phi_n(0), \quad m \rightarrow \infty,$$

*where the limit is obtained from the integral recursion (5).*

To illustrate Theorem 2.1, take  $N = 16$ ,  $L = 1000$ , and  $M = 10^{45}$ . Then the expected number of zero-sum vectors is  $\mu = 1000$ . In practice, finding all zero-sum vectors out of  $L^N = 10^{48}$  candidates is a time consuming task. In this example we have  $n = 4$  and  $m$  is approximately 150. Judging from Figure 2 illustrating the typical values for the proportion factor  $R_{n,m}$  using numerical computations based on the recursions for (7) presented in the next section, out of a thousand solutions the Wagner algorithm will catch no more than one.

**Theorem 2.2.** *For a random matrix (1) consider the number  $V_0$  of vectors (2) such that  $a_{i_1,1} + \dots + a_{i_N,N} \stackrel{M}{=} 0$ . Then*

$$\sum_{k=0}^{\infty} \left| P(V_0 = k) - \frac{\mu^k e^{-\mu}}{k!} \right| \leq 4(1 - e^{-\mu})M^{-1},$$

*where  $\mu = L^N M^{-1}$ . Furthermore, if  $N = 2^n$  and  $M = 2^m + 1$ ,  $m > n$ , then with  $\lambda = L^N p_{n,m}$*

$$\sum_{k=0}^{\infty} \left| P(W = k) - \frac{\lambda^k e^{-\lambda}}{k!} \right| \leq 8(1 - e^{-\lambda})\mu N L^{-1}.$$

According to Theorem 2.2, Poisson approximation for  $V_0$  works well when  $L^N \ll M$ . For  $W$ , a sufficient condition for the Chen-Stein bound to be small is  $N L^{N-1} \ll M$ .

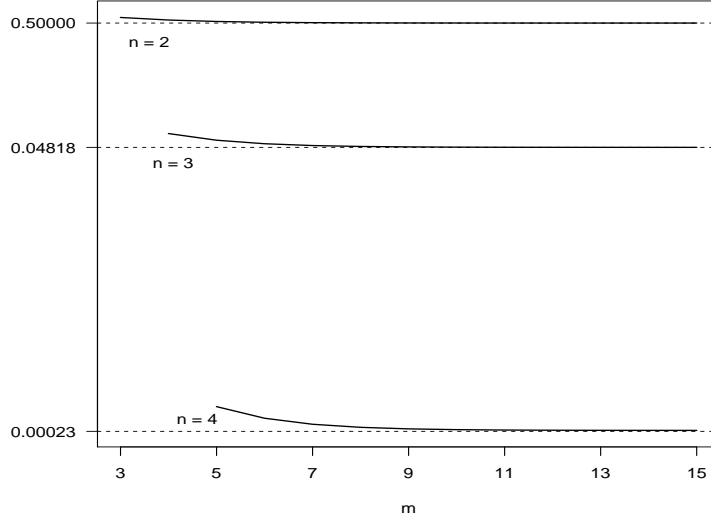


Figure 2: The ratio of the means (4) for  $n = 2, 3, 4$  are plotted as functions of  $m$ . The limits predicted by Theorem 2.1 are indicated by horizontal dotted lines.

### 3 Key recursion

Consider a backward recursion

$$v_i(j) = \sum_{k=0}^j v_{i+1}(k)v_{i+1}(j-k) + 2 \sum_{k=j+1}^{2^i} v_{i+1}(k)v_{i+1}(k-j) \quad (6)$$

involving a system of vectors  $(v_i(0), \dots, v_i(2^{i-1}))$  for  $i \geq 1$ . In particular, we have

$$v_i(0) = v_{i+1}^2(0) + 2 \sum_{k=1}^{2^i} v_{i+1}^2(k).$$

For  $1 \leq i \leq m-1$ , denote by  $v_i^{(m)}(j)$  the unique solution of (6) determined by the following frontier condition

$$v_{m-1}(0) = \dots = v_{m-1}(2^{m-2}) = (1 + 2^m)^{-1}.$$

By the forthcoming Corollary 3.2, we can write  $p_{n,m} = v_{m-n}^{(m)}(0)$  so that

$$R_{n,m} = (1 + 2^m)v_{m-n}^{(m)}(0), \quad n = 1, \dots, m-1. \quad (7)$$

**Lemma 3.1.** *Let  $1 \leq n \leq m-1$  and  $H_n(\mathbf{x})$  be defined by (3). Assuming that  $\mathbf{x}$  is a random vector with independent component uniformly distributed over  $D_m$ , put*

$$p_{i,m}(j) := P(H_i(\mathbf{x}) \stackrel{M}{=} j).$$

Then

$$p_{1,m}(-2^{m-2}) = \dots = p_{1,m}(2^{m-2}) = (2^m + 1)^{-1},$$

and for  $2 \leq i \leq m-1$  and  $0 \leq j \leq 2^{m-i-1}$ , we have  $p_{i,m}(-j) = p_{i,m}(j)$  with  $p_{i,m}(j)$  satisfying the recursion

$$p_{i,m}(j) = \sum_{k=0}^j p_{i-1,m}(k)p_{i-1,m}(j-k) + 2 \sum_{k=j+1}^{2^{m-i}} p_{i-1,m}(k)p_{i-1,m}(k-j).$$

*Proof.* There are exactly  $M = 2^m + 1$  different ordered pairs of numbers from the set  $D_m$  that add modulo  $M$  up to a given  $j \in D_{m-1}$ . These pairs have the form: for  $j = 0$ ,

$$(-2^{m-1} + k, 2^{m-1} - k), k = 0, \dots, 2^m,$$

for  $j = 1, \dots, 2^{m-2}$ ,

$$\begin{aligned} &(-2^{m-1} + k, -2^{m-1} + j - k - 1), \quad k = 0, \dots, j-1, \\ &(-2^{m-1} + k, 2^{m-1} + j - k), \quad k = j, \dots, 2^m, \end{aligned}$$

and for  $j = -2^{m-2}, \dots, -1$ ,

$$\begin{aligned} &(2^{m-1} - k, 2^{m-1} + j + k + 1), \quad k = 0, \dots, |j| - 1, \\ &(2^{m-1} - k, -2^{m-1} + j + k), \quad k = |j|, \dots, 2^m. \end{aligned}$$

Since these pairs appear with equal probability  $M^{-2}$ , the first claim follows.

On the other hand, for a given  $j \in D_{m-i}$  with  $i \geq 2$ , there are only  $M - |j|$  different ordered pairs of numbers from the set  $D_{m-i+1}$  that add modulo  $M$  up to  $j$ . These pairs have the form:

$$\begin{aligned} &(-2^{m-i} + k, 2^{m-i} + j - k), \quad k = j, \dots, 2^{m-i+1}, \quad j = 0, \dots, 2^{m-i-1}, \\ &(2^{m-i} - k, -2^{m-i} + j + k), \quad k = |j|, \dots, 2^{m-i+1}, \quad j = -2^{m-i-1}, \dots, -1. \end{aligned}$$

This yields for  $j = 1, \dots, 2^{m-i-1}$ ,

$$\begin{aligned} p_{i,m}(j) &= \sum_{k=j}^{2^{m-i+1}} p_{i-1,m}(-2^{m-i} + k)p_{i-1,m}(2^{m-i} - k + j), \\ p_{i,m}(-j) &= \sum_{k=j}^{2^{m-i+1}} p_{i-1,m}(2^{m-i} - k)p_{i-1,m}(-2^{m-i} + k - j). \end{aligned}$$

The stated symmetry property  $p_{i,m}(-j) = p_{i,m}(j)$  now follows recursively from the assumption of uniform distribution. To finish the proof of the lemma, it remains to observe that after replacing  $k - 2^{m-i}$  by  $l$  in the last relation for  $p_{i,m}(j)$  we get

$$p_{i,m}(j) = \sum_{l=j-2^{m-i}}^{2^{m-i}} p_{i-1,m}(l)p_{i-1,m}(j-l),$$

which in turn equals to

$$\begin{aligned} &\sum_{l=0}^j p_{i-1,m}(l)p_{i-1,m}(j-l) + \sum_{l=j+1}^{2^{m-i}} p_{i-1,m}(l)p_{i-1,m}(l-j) + \sum_{l=j-2^{m-i}}^{-1} p_{i-1,m}(-l)p_{i-1,m}(j-l) \\ &= \sum_{k=0}^j p_{i-1,m}(k)p_{i-1,m}(j-k) + 2 \sum_{k=j+1}^{2^{m-i}} p_{i-1,m}(k)p_{i-1,m}(k-j). \end{aligned}$$

□

**Corollary 3.2.** *Comparison of the key recursion in Lemma 3.1 with the recursion (6) yields*

$$p_{m-i,m}(j) = v_i^{(m)}(j).$$

## 4 Proof of Theorem 2.1

Recall (7) and put

$$R_{n,m}(j) = 2^m v_{m-n}^{(m)}(j), \quad \phi_{n,m}(x) := \phi_n(x2^{-m}).$$

We prove Theorem 2.1 by verifying a more general convergence result

$$\alpha_{n,m} := \max_{0 \leq j \leq 2^{m-n-1}} |R_{n,m}(j) - \phi_{n,m}(j)| \rightarrow 0, \quad m \rightarrow \infty. \quad (8)$$

To this end we use induction over  $n$ . The base case  $n = 1$  is trivial. To prove the inductive step observe first that by (6)

$$R_{n,m}(j) = 2^{-m} \sum_{k=0}^j R_{n-1,m}(k) R_{n-1,m}(j-k) + 2^{1-m} \sum_{k=j+1}^{2^{m-n}} R_{n-1,m}(k) R_{n-1,m}(k-j). \quad (9)$$

It is easy to see recursively that the constant

$$C_n := \sup_{m > n} \max_{0 \leq j \leq 2^{m-n-1}} R_{n,m}(j)$$

is finite.

On the other hand, by (5),

$$\phi_{n,m}(j) = 2^{-m} \int_0^j \phi_{n-1,m}(u) \phi_{n-1,m}(j-u) du + 2^{1-m} \int_j^{2^{m-n}} \phi_{n-1,m}(u) \phi_{n-1,m}(u-j) du,$$

so that

$$\begin{aligned} \phi_{n,m}(j) &= 2^{-m} \sum_{k=0}^j \phi_{n-1,m}(k) \phi_{n-1,m}(j-k) \\ &\quad + 2^{1-m} \sum_{k=j+1}^{2^{m-n}} \phi_{n-1,m}(k) \phi_{n-1,m}(k-j) + \epsilon_{n,m}(j), \end{aligned} \quad (10)$$

with accordingly defined remainder term  $\epsilon_{n,m}(j)$ . Uniform continuity of  $\phi_n(x)$  yields uniform convergence  $\epsilon_{n,m}(j) \rightarrow 0$  as  $m \rightarrow \infty$ , and (8) follows from (9) and (10), since

$$\alpha_{n,m} \leq 2 \left[ C_{n-1} + \max_{0 \leq x \leq 2^{-n}} \phi_n(x) \right] \alpha_{n-1,m} + \max_{0 \leq j \leq 2^{m-n}} |\epsilon_{n,m}(j)|.$$

## 5 Proof of Theorem 2.2

The following result is a straightforward corollary of Theorem 1 from [1] and is a key tool for our proof here.

**Lemma 5.1.** *Let  $Z = \sum_{\mathbf{i} \in \mathbf{J}} \chi_{\mathbf{i}}$  be a sum of possibly dependent indicator random variables with  $E(Z) = \zeta$ . Suppose there is a family of subsets  $\mathbf{J}_{\mathbf{i}} \subset \mathbf{J}$  such that for any  $\mathbf{i} \in \mathbf{J}$  and  $\mathbf{k} \notin \mathbf{J}_{\mathbf{i}}$ , indicators  $\chi_{\mathbf{i}}$  and  $\chi_{\mathbf{k}}$  are independent. Then*

$$\frac{\zeta}{4(1 - e^{-\zeta})} \sum_{k=0}^{\infty} \left| P(Z = k) - \frac{\zeta^k e^{-\zeta}}{k!} \right| \leq \sum_{\mathbf{i} \in \mathbf{J}} \sum_{\mathbf{k} \in \mathbf{J}_{\mathbf{i}}} E(\chi_{\mathbf{i}}) E(\chi_{\mathbf{k}}) + \sum_{\mathbf{i} \in \mathbf{J}} \sum_{\mathbf{k} \in \mathbf{J} \setminus \{\mathbf{i}\}} E(\chi_{\mathbf{i}} \chi_{\mathbf{k}}).$$

We start the proof of Theorem 2.2 by observing that  $V_0 = \sum_{\mathbf{i} \in \mathbf{J}} \chi_{\mathbf{i}}$ , where the indicator random variables

$$\chi_{\mathbf{i}} = 1_{\{a_{i_1,1} + \dots + a_{i_N,N} \stackrel{M}{=} 0\}}, \quad \mathbf{i} = (i_1, \dots, i_N)$$

are identically distributed with  $E(\chi_{\mathbf{i}}) = M^{-1}$ , and mutually independent. Independence is due to the defining property of the matrix  $\mathbf{A}$ . Indeed, if  $\mathbf{k} \neq \mathbf{i}$  and (without loss of generality)  $1, \dots, j$  are the coordinates where these two vectors differ, then

$$\begin{aligned} P(a_{k_1,1} + \dots + a_{k_N,N} \stackrel{M}{=} a_{i_1,1} + \dots + a_{i_N,N} \stackrel{M}{=} 0) \\ &= P(a_{k_1,1} + \dots + a_{k_j,j} \stackrel{M}{=} a_{i_1,1} + \dots + a_{i_j,j} \stackrel{M}{=} -a_{i_{j+1},j+1} - \dots - a_{i_N,N}) \\ &= \sum_{b \in D_m} P(a_{k_1,1} + \dots + a_{k_j,j} \stackrel{M}{=} b; a_{i_1,1} + \dots + a_{i_j,j} \stackrel{M}{=} b; a_{i_{j+1},j+1} + \dots + a_{i_N,N} \stackrel{M}{=} -b) \\ &= M^{-1} \sum_{b \in D_m} P(a_{i_1,1} + \dots + a_{i_j,j} \stackrel{M}{=} b; a_{i_{j+1},j+1} + \dots + a_{i_N,N} \stackrel{M}{=} -b) = M^{-2}. \end{aligned}$$

Therefore, we can apply Lemma 5.1 with  $\mathbf{J}_{\mathbf{i}} = \{\mathbf{i}\}$ , and the Chen-Stein bound for  $V_0$  follows from  $E(V_0) = \mu$  and

$$\sum_{\mathbf{i} \in \mathbf{J}} \sum_{\mathbf{k} \in B_{\mathbf{i}}} E(\chi_{\mathbf{i}}) E(\chi_{\mathbf{k}}) = L^N M^{-2} = \mu M^{-1}.$$

To obtain the Chen-Stein bound for  $W$ , we define  $\mathbf{J}_{\mathbf{i}}$  as the set of  $\mathbf{k} \in L$  such that vectors  $\mathbf{i}$  and  $\mathbf{k}$  share at least one component. Observe that

$$|\mathbf{J}_{\mathbf{i}}| = L^N - (L - 1)^N.$$

By definition of  $W$ ,

$$W = \sum_{\mathbf{i} \in \mathbf{J}} \chi_{\mathbf{i}}, \quad \chi_{\mathbf{i}} = 1_{\{H_n(\mathbf{a}_{\mathbf{i}}) = 0\}},$$

so that  $E(\chi_{\mathbf{i}}) = p_{n,m}$  and therefore,

$$\sum_{\mathbf{i} \in \mathbf{J}} \sum_{\mathbf{k} \in \mathbf{J}_{\mathbf{i}}} E(\chi_{\mathbf{i}}) E(\chi_{\mathbf{k}}) = L^N (L^N - (L - 1)^N) p_{n,m}^2 \leq N L^{-1} \lambda^2.$$

Since a Wagner's solution is necessarily is a zero-sum vector, we have for  $\mathbf{i} \neq \mathbf{k}$ ,

$$E(\chi_{\mathbf{i}} \chi_{\mathbf{k}}) = P(H_n(\mathbf{a}_{\mathbf{i}}) = 0; H_n(\mathbf{a}_{\mathbf{k}}) = 0) \leq P(a_{k_1,1} + \dots + a_{k_N,N} \stackrel{M}{=} 0; H_n(\mathbf{a}_{\mathbf{i}}) = 0).$$

Let  $l_1, \dots, l_j$  are the coordinates where the vectors  $\mathbf{i}, \mathbf{k}$  differ. Then it follows that

$$\begin{aligned} \mathbb{E}(\chi_{\mathbf{i}}\chi_{\mathbf{k}}) &\leq \sum_{b \in D_m} \mathbb{P}(a_{k_{l_1}, l_1} + \dots + a_{k_{l_j}, l_j} \stackrel{M}{=} b; a_{i_{l_1}, l_1} + \dots + a_{i_{l_j}, l_j} \stackrel{M}{=} b; H_n(\mathbf{a}_{\mathbf{i}}) = 0) \\ &= M^{-1} \sum_{b \in D_m} \mathbb{P}(a_{i_{l_1}, l_1} + \dots + a_{i_{l_j}, l_j} \stackrel{M}{=} b; H_n(\mathbf{a}_{\mathbf{i}}) = 0) = M^{-1} p_{n,m}, \end{aligned}$$

and we get

$$\sum_{\mathbf{i} \in \mathbf{J}} \sum_{\mathbf{k} \in \mathbf{J}_{\mathbf{i}} \setminus \{\mathbf{i}\}} \mathbb{E}(\chi_{\mathbf{i}}\chi_{\mathbf{k}}) \leq L^N (L^N - (L-1)^N) p_{n,m} M^{-1} \leq NL^{-1} \lambda \mu.$$

The proof is finished by applying once again Lemma 5.1.

**Acknowledgements.** The first author is grateful to Vladimir Vatutin and Andrey Zubkov for formulating an initial problem setting that eventually lead to this research project.

## References

- [1] ARRATIA, R., GOLDSTEIN, L., GORDON, L. (1989). Two moments suffice for Poisson approximation: the Chen-Stein method. *Ann. Prob.* **17**, 9–25.
- [2] FULMAN, J., GOLDSTEIN, L. (2015). Stein’s method and the rank distribution of random matrices over finite fields. *Ann. Prob.* **43**, 1274–1314.
- [3] GREFERATH, M. (2009). An introduction to ring-linear coding theory. *Gröbner Bases, Coding, and Cryptography*. Springer, 219–238.
- [4] LANG, S. (2002). *Abstract algebra*, 3rd edn. Springer, New York.
- [5] MINDER, L., SINCLAIR, A. (2012). The extended k-tree algorithm. *J. Cryptol.* **25**, 349–382.
- [6] SCHROEPPEL, R., SHAMIR, A. (1981). A  $T = O(2n/2)$ ,  $S = O(2n/4)$  algorithm for certain NP-complete problems. *SIAM J. Comput.* **10**, 456–464.
- [7] WAGNER, D. (2002). A generalized birthday problem. *CRYPTO 2002*. Springer, 288–303.