SERIES EXPANSION FOR THE PROBABILITY THAT A RANDOM BOOLEAN MATRIX IS OF MAXIMAL RANK

UDC 519.21

V. V. MASOL

ABSTRACT. We consider a random $(N \times n)$ matrix in the field $GF(2)$ and establish relations that allow one to find the coefficients of the expansion of the probability that a given matrix is of maximal rank into a series in powers of a small parameter. We give explicit formulas for the cases of $n = 1$ and $n = 2$, $N \geq n$.

1. SETTING OF THE PROBLEM

Let $\mathbf{A} = (a_{ij})_{i \in I, j \in J}$ be a matrix with N rows and n columns, where $I = \{1, \dots, N\}$ and $J = \{1, \dots, n\}$. The entries of the matrix $\mathbf{A}$ are independent random variables that assume values in the field $GF(2)$ and have distribution

$$(1) \quad \mathbf{P}\{a_{ij} = 0\} = 1 - \mathbf{P}\{a_{ij} = 1\} = 2^{-1} (1 + \varepsilon x_{ij})$$

where ε is a fixed small number, $\varepsilon \geq 0$, and $x_{ij} \in (-\infty, \infty)$. Denote by $\chi(\mathbf{A})$ the following indicator:

$$\chi(\mathbf{A}) = \begin{cases} 1 & \text{if the matrix } \mathbf{A} \text{ contains } n \text{ linearly independent (in the field } GF(2)) \\ & N\text{-dimensional columns;} \\ 0, & \text{otherwise.} \end{cases}$$

Using relation (1), the probability of the event $\{\chi(\mathbf{A}) = 1\}$ can be represented in the following form:

$$(2) \quad \mathbf{P}\{\chi(\mathbf{A}) = 1\} = \sum_{s=0}^{nN} \varepsilon^s f^{(s)}(x_{ij}, i \in I, j \in J)$$

where the coefficients $f^{(s)}(x_{ij}, i \in I, j \in J)$, $s \geq 0$, are real numbers that do not depend on ε .

Let $m = N - n$. In the case of $m = 0$, a recurrence relation with respect to n is found in [1] to evaluate $f^{(s)}(x_{ij}, i \in I, j \in J)$, $s \geq 0$; for the case of

$$(3) \quad m \geq 0$$

the coefficients $f^{(s)}(x_{ij}, i \in I, j \in J)$, $s \in \{0, 1, 2\}$, are found in [2] in an explicit form by applying different approaches depending on $s \in \{0, 1, 2\}$.

The aim of this paper is to find a relation that allows one to evaluate the coefficients

$$f^{(s)}(x_{ij}, i \in I, j \in J), \quad s \geq 1,$$

of the expansion of the probability that a random $(N \times n)$ matrix in the field $GF(2)$ is of the maximal rank n into a series in terms of powers of a small parameter ε . Our

methods are based on the results obtained in [3] and on the explicit expansion for the cases of $n = 1$ and $n = 2$.

2. MAIN RESULTS

In what follows we need some notation. Let $R(s)$ be a set of s distinct elements,

$$R(s) = \{(i_1, j_1), \dots, (i_s, j_s)\},$$

and let $t_{R(s)}$ be the coefficient of $x_{i_1 j_1} \cdots x_{i_s j_s}$ in the representation of

$$f^{(s)}(x_{ij}, i \in I, j \in J),$$

$t_{R(s)} = \text{coef}_{x_{i_1 j_1} \cdots x_{i_s j_s}} f^{(s)}(x_{ij}, i \in I, j \in J)$ (here and in what follows the parameters i and j with or without superscripts are elements of the sets I and J , respectively, namely $i \in I$ and $j \in J$). Then

$$(4) \quad f^{(s)}(x_{ij}, i \in I, j \in J) = \sum t_{R(s)} x_{i_1 j_1} \cdots x_{i_s j_s}$$

where the sum is taken over all different sets $R(s)$.

Remark 1. In what follows we assume that the equality $R_1(s) = R_2(s)$ holds if and only if the set $R_1(s)$ can be obtained from $R_2(s)$ by permuting the elements $(i_\nu^{(2)}, j_\nu^{(2)})$, $\nu = 1, 2, \dots, s$, and vice versa, where

$$R_t(s) = \{(i_1^{(t)}, j_1^{(t)}), \dots, (i_s^{(t)}, j_s^{(t)})\}, \quad t = 1, 2.$$

Put

$$(5) \quad \zeta(j) = \{i : (i, j) \in R(s)\}, \quad j \in J.$$

Theorem 1. *Let a collection $\{j_1, \dots, j_s\}$ contain k elements of J , that is,*

$$\{j_1, \dots, j_s\} = \{\mu_1, \dots, \mu_k\}, \quad 1 \leq \mu_1 < \dots < \mu_k \leq n.$$

If conditions (1) and (3) hold, then

$$(6) \quad t_{R(s)} = -2^{-(N-1)k-1} \frac{P(N-k)}{P(m)} \sum_{c \in \gamma_0} (-1)^\tau$$

where

$$P(N) = \prod_{\nu=1}^N (1 - 2^{-\nu}), \quad P(0) = 1;$$

γ_0 is the set of matrices c , $c = (c_{ij})_{i \in I, j \in \{1, \dots, k-1\}}$, in the field $GF(2)$ such that the rank of c is $k-1$ and c satisfy the following condition:

$$(7) \quad \bigoplus_{i \in \zeta(\mu_k)} c_{ij} = 0, \quad j \in \{1, \dots, k-1\}.$$

Here $\tau = \bigoplus_{\omega=1}^{k-1} \bigoplus_{i \in \zeta(\mu_\omega)} c_{i\omega}$ and the symbol $\oplus$ stands for the operation of summation in the field $GF(2)$.

Remark 2. In what follows we assume that $\sum_{c \in \gamma_0} (-1)^\tau \equiv 1$ if $k = 1$.

Let $R(s) = \{(i_1, j_1), \dots, (i_s, j_s)\}$ and

$$(8) \quad \{j_1, \dots, j_s\} = \{\mu_1, \mu_2\},$$

$\mu_1, \mu_2 \in J$, $\mu_1 \neq \mu_2$, and $s \geq 2$. Put $\zeta_{12} = \zeta(\mu_1) \cap \zeta(\mu_2)$, $s_q = |\zeta(\mu_q) \setminus \zeta_{12}|$, $q = 1, 2$, and $s_{12} = |\zeta_{12}|$.

Theorem 2. (i) If $N \geq 1$, $n = 1$, and condition (1) holds, then

$$P\{\chi(\mathbf{A}) = 1\} = 1 - 2^{-N} - 2^{-N} \sum_{s=1}^N \varepsilon^s \sum_{1 \leq i_1 < \dots < i_s \leq N} \prod_{q=1}^s x_{i_q 1};$$

(ii) if $N \geq 2$, $n = 2$, and condition (1) holds, then

$$\begin{aligned} P\{\chi(\mathbf{A}) = 1\} &= (1 - 2^{-N}) (1 - 2^{-N+1}) \\ &\quad - 2^{-N} (1 - 2^{-N+1}) \sum_{s=1}^N \varepsilon^s \sum_{1 \leq i_1 < \dots < i_s \leq N} \left[\sum_{j=1}^2 \prod_{q=1}^s x_{i_q j} + \varepsilon^s \prod_{j=1}^2 \prod_{q=1}^s x_{i_q j} \right] \\ &\quad + 2^{-2N+1} \sum_{s=2}^{2N-1} \varepsilon^s \sum_{1 \leq i_1 < \dots < i_{s_1} \leq N} 1 \\ &\quad \times \sum_{\substack{1 \leq i'_1 < \dots < i'_{s_{12}} \leq N \\ i'_l \notin \{i_1, \dots, i_{s_1}\}}} \sum_{\substack{1 \leq i''_1 < \dots < i''_{s_2} \leq N \\ i''_l \notin \{i_1, \dots, i_{s_1}, i'_1, \dots, i'_{s_{12}}\}}} \left(\prod_{q=1}^{s_1} x_{i_q 1} \right) \\ &\quad \times \left(\prod_{j=1}^2 \prod_{q=1}^{s_{12}} x_{i'_q j} \right) \left(\prod_{q=1}^{s_2} x_{i''_q 2} \right) \end{aligned}$$

where the sum is taken over all nonnegative integers s_1 , s_{12} , and s_2 such that $s_1 + 2s_{12} + s_2 = s$ and either $s_{12} = 0$, $s_1 \geq 1$, $s_2 \geq 1$ or $s_{12} \geq 1$, $s_1 + s_2 \geq 1$; we also put $\prod_{q=1}^0 \equiv 1$.

3. PROOF OF THEOREM 1

It is proved in [3] that

$$(9) \quad t_{R(s)} = 2^{-nN} \sum_{\alpha'_0, \alpha'_0 \subseteq \alpha_0} \sum_{c \in \alpha'_0} (-1)^\sigma$$

where α_0 is the collection of $(N \times n)$ matrices c of rank n in the field $GF(2)$ such that $c = (c_{ij})_{i \in I, j \in J}$ and $\bigoplus_{i \in \zeta(\mu_k)} c_{ij} = 0$, $j \in \{\mu_1, \dots, \mu_{k-1}\}$; $\sigma = \bigoplus_{q=1}^s c_{i_q j_q}$; α'_0 is the subset of α_0 , $\alpha'_0 \subseteq \alpha_0$, consisting of matrices $c \in \alpha_0$ such that $c^{(l)}$ and $c^{(t)}$ belong to α'_0 if and only if $c_j^{(l)} = c_j^{(t)}$ for $j \in \{\mu_1, \dots, \mu_k\}$, $l \neq t$; $c_j^{(\xi)}$ is the column j of the matrix $c^{(\xi)}$ for $c^{(\xi)} \in \alpha_0$, $j \in J$, $\xi = 1, 2, \dots$.

It is easy to see that

$$(10) \quad \sum_{c \in \alpha'_0} (-1)^\sigma = (-1)^\sigma B_k$$

for an arbitrary collection α'_0 , $\alpha'_0 \subseteq \alpha_0$, where B_k is the cardinality of the set α'_0 , $B_k = |\alpha'_0|$. Relations (9) and (10) imply that

$$(11) \quad t_{R(s)} = 2^{-nN} B_k \sum_{c \in \beta_0} (-1)^\sigma$$

where β_0 is the set of $(N \times k)$ matrices c , $c = (c_{ij})_{i \in I, j \in \{1, \dots, k\}}$, of rank k in the field $GF(2)$ satisfying condition (7).

Now we show that

$$(12) \quad \sum_{c \in \beta_0} (-1)^\sigma = -2^{k-1} \sum_{c \in \gamma_0} (-1)^\tau.$$

Indeed, the sum on the left-hand side of (12) can be rewritten as follows:

$$(13) \quad \sum_{c \in \beta_0} (-1)^\sigma = \sum_{\beta'_0, \beta'_0 \subseteq \beta_0} \sum_{c \in \beta'_0} (-1)^\sigma$$

where β'_0 is a subset of β_0 , $\beta'_0 \subseteq \beta_0$, consisting of matrices $c \in \beta_0$ such that both $c^{(l)}$ and $c^{(t)}$ belong to β'_0 if and only if $c_j^{(l)} = c_j^{(t)}$ for $j = 1, 2, \dots, k-1$, $l \neq t$, $c^{(\xi)} \in \beta_0$, $\xi = 1, 2, \dots$. Fix a set β'_0 . Let the sum $\sum_{c \in \beta'_0} (-1)^\sigma$ contain μ_0 terms $(-1)^\tau$ and let μ_0^- be the number of changes of the sign $(-1)^\tau$ in this sum. It is clear that $\mu_0 = \gamma_1 \gamma_2 - \nu$ and $\mu_0^- = \gamma_1^- \gamma_2 - \nu^-$ where γ_1 (γ_1^-) is the total number of ways to place an even (odd) number of nonzero elements of the field $GF(2)$ to those positions of the column k in the matrix c , $c \in \beta'_0$, whose indices belong to the set $\zeta(\mu_k)$. The numbers ν (ν^-) are defined similarly to the numbers γ_1 (γ_1^-) under the additional condition that the elements of the column k are linear combinations of the corresponding elements in the first $k-1$ columns of the matrix c ; γ_2 is the total number of ways to place elements of the field $GF(2)$ to the positions $I \setminus \zeta(\mu_k)$ of the N -dimensional column k in the matrix c , $c \in \beta'_0$. It is proved in [3] that $\gamma_1 = \gamma_1^- = 2^{|\zeta(\mu_k)|-1}$, $\gamma_2 = 2^{N-|\zeta(\mu_k)|}$, $\nu = 2^{k-1}$, and $\nu^- = 0$. Thus $\mu_0 - \mu_0^- = -2^{k-1}$, whence

$$(14) \quad \sum_{\beta'_0, \beta'_0 \subseteq \beta_0} \sum_{c \in \beta'_0} (-1)^\sigma = -2^{k-1} \sum_{c \in \gamma_0} (-1)^\tau.$$

Relations (13) and (14) prove (12).

Now we show that

$$(15) \quad B_k = (2^N - 2^k) \dots (2^N - 2^{n-1}), \quad k \geq 1.$$

Indeed, according to the definition of B_k

$$(16) \quad B_k = \prod_{l=1}^{n-k} b_{\delta_l}$$

where $1 \leq \delta_1 < \dots < \delta_{n-k} \leq n$, $\delta_1, \dots, \delta_{n-k} \notin \{\mu_1, \dots, \mu_k\}$, and b_{δ_l} is the total number of ways to place elements of the field $GF(2)$ to an N -dimensional column such that this column is linearly independent of the columns with indices $\mu_1, \dots, \mu_k, \delta_1, \dots, \delta_{l-1}$. It is clear that

$$b_{\delta_l} = 2^N - 2^{k+l-1}, \quad l = 1, 2, \dots, n-k.$$

Taking into account (16) we get (15). Using relations (11), (12), and (15) we prove (6) by an obvious calculation. Theorem 1 is proved. $\square$

4. APPLICATIONS OF THEOREM 1

Example 1. If $s = 1$, then

$$(17) \quad f^{(s)}(x_{ij}, i \in I, j \in J) = -2^{-N} \frac{P(N-1)}{P(m)} \sum_{i=1}^N \sum_{j=1}^n x_{ij}.$$

Indeed, if $s = 1$, then

$$(18) \quad f^{(s)}(x_{ij}, i \in I, j \in J) = \sum_{i=1}^N \sum_{j=1}^n t_{R(s)} x_{ij}$$

where $R(s) = \{(i, j)\}$. The parameter k defined in Theorem 1 is equal to $k = 1$, thus we find from (6) and Remark 2 that

$$(19) \quad t_{R(s)} = -2^{-N} \frac{P(N-1)}{P(m)}.$$

Hence (18) and (19) imply (17).

Example 2. If $s = 2$, then

$$\begin{aligned}
 f^{(s)}(x_{ij}, i \in I, j \in J) \\
 (20) \quad &= -2^{-N} \frac{P(N-1)}{P(m)} \left(\sum_{j=1}^n \sum_{1 \leq i_1 < i_2 \leq N} x_{i_1 j} x_{i_2 j} + \sum_{i=1}^N \sum_{1 \leq j_1 < j_2 \leq n} x_{i j_1} x_{i j_2} \right) \\
 &+ 2^{-2N+1} \frac{P(N-2)}{P(m)} \sum_{1 \leq j_1 < j_2 \leq n} \sum_{i_1 \neq i_2} x_{i_1 j_1} x_{i_2 j_2}.
 \end{aligned}$$

Indeed, it follows from (4) that

$$\begin{aligned}
 f^{(s)}(x_{ij}, i \in I, j \in J) &= \sum_{j=1}^n \sum_{1 \leq i_1 < i_2 \leq N} t_{R_1(s)} x_{i_1 j} x_{i_2 j} + \sum_{i=1}^N \sum_{1 \leq j_1 < j_2 \leq n} t_{R_2(s)} x_{i j_1} x_{i j_2} \\
 (21) \quad &+ \sum_{1 \leq j_1 < j_2 \leq n} \sum_{i_1 \neq i_2} t_{R_3(s)} x_{i_1 j_1} x_{i_2 j_2}
 \end{aligned}$$

where $R_1(s) = \{(i_1, j), (i_2, j)\}$, $R_2(s) = \{(i, j_1), (i, j_2)\}$, and $R_3(s) = \{(i_1, j_1), (i_2, j_2)\}$. Using (6) for $k = 1$ and Remark 2 we get

$$(22) \quad t_{R_1(s)} = -2^{-N} \frac{P(N-1)}{P(m)}.$$

Now we check the relations

$$(23) \quad t_{R_2(s)} = -2^{-N} \frac{P(N-1)}{P(m)},$$

$$(24) \quad t_{R_3(s)} = 2^{-2N+1} \frac{P(N-2)}{P(m)}.$$

It follows from (6) for $k = 2$ that

$$(25) \quad t_{R_2(s)} = -2^{-2N+1} \frac{P(N-2)}{P(m)} \sum_{c \in \gamma_0} (-1)^\tau$$

where γ_0 is the set of all N -dimensional columns c , $c = (c_{\nu 1})_{\nu \in I}$, of rank 1 in the field $GF(2)$ such that $c_{i1} = 0$. Since $\tau = \bigoplus_{\nu \in \zeta(j_1)} c_{\nu 1} = c_{i1} = 0$, we have $\tau = 0$. Thus

$$(26) \quad \sum_{c \in \gamma_0} (-1)^\tau = |\gamma_0| = 2^{N-1} - 1.$$

Using (25) and (26) we get (23).

Further, relation (6) for $k = 2$ implies

$$(27) \quad t_{R_3(s)} = -2^{-2N+1} \frac{P(N-2)}{P(m)} \sum_{c \in \gamma_0} (-1)^\tau$$

where γ_0 is the set of all N -dimensional columns c , $c = (c_{\nu 1})_{\nu \in I}$, of rank 1 in the field $GF(2)$ such that $c_{i21} = 0$. Note that $\tau = c_{i11}$. Hence

$$\sum_{c \in \gamma_0} (-1)^\tau = \sum_{c \in \gamma_0^+} 1 - \sum_{c \in \gamma_0^-} 1$$

where $\gamma_0^+ \subseteq \gamma_0$ ($\gamma_0^- \subseteq \gamma_0$) and $c_{i11} = 0$ ($c_{i11} = 1$) for any column $c \in \gamma_0^+$ ($c \in \gamma_0^-$).

It is clear that $|\gamma_0^+| = 2^{N-2} - 1$ and $|\gamma_0^-| = 2^{N-2}$. Thus

$$(28) \quad \sum_{c \in \gamma_0} (-1)^\tau = -1.$$

Substituting (28) into (27) we obtain (24). Relations (21)–(24) prove (20).

Example 3. If $s = nN$, then

$$(29) \quad f^{(s)}(x_{ij}, i \in I, j \in J) = -2^{-N} \frac{P(N-1)}{P(m)} \prod_{i=1}^N \prod_{j=1}^n x_{ij}.$$

Indeed, (4) implies

$$(30) \quad f^{(s)}(x_{ij}, i \in I, j \in J) = t_{R(s)} \prod_{i=1}^N \prod_{j=1}^n x_{ij}$$

where $R(s) = \{(i, j), i \in I, j \in J\}$. Now we show that

$$(31) \quad t_{R(s)} = -2^{-N} \frac{P(N-1)}{P(m)}.$$

Using (6) for $k = n$ we obtain

$$(32) \quad t_{R(s)} = -2^{(N-1)n-1} \sum_{c \in \gamma_0} (-1)^\tau$$

where γ_0 is the set of all $(N \times (n-1))$ matrices c , $c = (c_{ij})_{i \in I, j \in \{1, \dots, n-1\}}$, of rank $n-1$ in the field $GF(2)$ such that

$$\bigoplus_{i=1}^N c_{ij} = 0, \quad j \in \{1, 2, \dots, n-1\}.$$

This implies that $\tau = 0$, since

$$\tau = \bigoplus_{\omega=1}^{n-1} \bigoplus_{i \in \zeta(\omega)} c_{i\omega} = \bigoplus_{\omega=1}^{n-1} \bigoplus_{i=1}^N c_{i\omega} = 0.$$

Therefore

$$(33) \quad \sum_{c \in \gamma_0} (-1)^\tau = |\gamma_0|.$$

Now we prove that

$$(34) \quad |\gamma_0| = (2^{N-1} - 1) \cdots (2^{N-1} - 2^{n-2}).$$

Indeed, $|\gamma_0| = b_1 \cdots b_{n-1}$ where b_q ($q = 1, 2, \dots, n-1$) is the total number of ways to place elements of the field $GF(2)$ to an N -dimensional column such that the number of unit elements in the column is even and the column does not linearly depend on the columns with indices $1, 2, \dots, q-1$. It is clear that

$$b_q = 2^{N-1} - 2^{q-1}, \quad q = 1, 2, \dots, n-1.$$

This implies (34). Relations (33) and (34) allow one to represent (32) in the form of (31). Substituting (31) into (30) we get (29).

Example 4. If $s = 3$, then

$$\begin{aligned}
 f^{(s)}(x_{ij}, i \in I, j \in J) &= -2^{-N} \frac{P(N-1)}{P(m)} \left(\sum_{1 \leq i_1 < i_2 < i_3 \leq N} \sum_{j=1}^n \prod_{l=1}^3 x_{i_l j} + \sum_{1 \leq j_1 < j_2 < j_3 \leq n} \sum_{i=1}^N \prod_{l=1}^3 x_{i j_l} \right) \\
 &+ 2^{-2N+1} \frac{P(N-2)}{P(m)} \left\{ \sum_{1 \leq j_1 < j_2 \leq n} 1 \right. \\
 &\quad \times \left[\sum_{1 \leq i_1 < i_2 \leq N} (x_{i_1 j_1} x_{i_1 j_2} x_{i_2 j_1} + x_{i_1 j_1} x_{i_1 j_2} x_{i_2 j_2} \right. \\
 &\quad \left. \left. + x_{i_1 j_1} x_{i_2 j_1} x_{i_2 j_2} + x_{i_1 j_2} x_{i_2 j_1} x_{i_2 j_2} \right) \right. \\
 &\quad \left. + \sum_{1 \leq i_1 < i_2 < i_3 \leq N} \sum_{(\nu_1, \nu_2, \nu_3) \in \pi(j_1, j_2)} x_{i_1 \nu_1} x_{i_2 \nu_2} x_{i_3 \nu_3} \right] \\
 &\quad + \sum_{1 \leq j_1 < j_2 < j_3 \leq n} \sum_{1 \leq i_1 < i_2 \leq N} 1 \\
 &\quad \times \sum_{(\lambda_1, \lambda_2, \lambda_3) \in \pi(i_1, i_2)} x_{\lambda_1 j_1} x_{\lambda_2 j_2} x_{\lambda_3 j_3} \Big\} \\
 &- 2^{-3N+3} \frac{P(N-3)}{P(m)} \sum_{1 \leq j_1 < j_2 < j_3 \leq n} \sum_{1 \leq i_1 < i_2 < i_3 \leq N} \\
 &\quad \times \sum_{(\gamma_1, \gamma_2, \gamma_3) \in \pi(j_1, j_2, j_3)} x_{i_1 \gamma_1} x_{i_2 \gamma_2} x_{i_3 \gamma_3}
 \end{aligned} \tag{35}$$

where $\pi(j_1, j_2)$ ($\pi(j_1, j_2, j_3)$) is the set of all permutations of the sets $\{j_1, j_1, j_2\}$ and $\{j_2, j_2, j_1\}$ ($\{j_1, j_2, j_3\}$).

Indeed, relation (4) implies for $s = 3$ that

$$\begin{aligned}
 f^{(s)}(x_{ij}, i \in I, j \in J) &= \sum_{1 \leq i_1 < i_2 < i_3 \leq N} \sum_{j=1}^n t_{R_{11}(s)} \prod_{l=1}^3 x_{i_l j} + \sum_{1 \leq j_1 < j_2 < j_3 \leq n} \sum_{i=1}^N t_{R_{12}(s)} \prod_{l=1}^3 x_{i j_l} \\
 &+ \sum_{1 \leq j_1 < j_2 \leq n} \sum_{1 \leq i_1 < i_2 \leq N} (t_{R_{211}(s)} x_{i_1 j_1} x_{i_1 j_2} x_{i_2 j_1} + t_{R_{212}(s)} x_{i_1 j_1} x_{i_1 j_2} x_{i_2 j_2} \\
 &\quad + t_{R_{213}(s)} x_{i_1 j_1} x_{i_2 j_1} x_{i_2 j_2} + t_{R_{214}(s)} x_{i_1 j_2} x_{i_2 j_1} x_{i_2 j_2}) \\
 &+ \sum_{1 \leq j_1 < j_2 \leq n} \sum_{1 \leq i_1 < i_2 < i_3 \leq N} \sum_{q=1}^6 t_{R_{22q}(s)} x_{i_1 \nu_1^{(q)}} x_{i_2 \nu_2^{(q)}} x_{i_3 \nu_3^{(q)}} \\
 &+ \sum_{1 \leq j_1 < j_2 < j_3 \leq n} \sum_{1 \leq i_1 < i_2 \leq N} \sum_{q=1}^6 t_{R_{23q}(s)} x_{\lambda_1^{(q)} j_1} x_{\lambda_2^{(q)} j_2} x_{\lambda_3^{(q)} j_3} \\
 &+ \sum_{1 \leq j_1 < j_2 < j_3 \leq n} \sum_{1 \leq i_1 < i_2 < i_3 \leq N} \sum_{q=1}^6 t_{R_{3q}(s)} x_{i_1 \gamma_1^{(q)}} x_{i_2 \gamma_2^{(q)}} x_{i_3 \gamma_3^{(q)}}
 \end{aligned} \tag{36}$$

where

$$R_{11}(s) = \{(i_1, j), (i_2, j), (i_3, j)\}, \quad R_{12}(s) = \{(i, j_1), (i, j_2), (i, j_3)\},$$

$$\begin{aligned}
R_{211}(s) &= \{(i_1, j_1), (i_1, j_2), (i_2, j_1)\}, & R_{212}(s) &= \{(i_1, j_1), (i_1, j_2), (i_2, j_2)\}, \\
R_{213}(s) &= \{(i_1, j_1), (i_2, j_1), (i_2, j_2)\}, & R_{214}(s) &= \{(i_1, j_2), (i_2, j_1), (i_2, j_2)\}, \\
R_{22q}(s) &= \{(i_1, \nu_1^{(q)}), (i_2, \nu_2^{(q)}), (i_3, \nu_3^{(q)})\}, & (\nu_1^{(q)}, \nu_2^{(q)}, \nu_3^{(q)}) &\in \pi(j_1, j_2), \\
R_{23q}(s) &= \{(\lambda_1^{(q)}, j_1), (\lambda_2^{(q)}, j_2), (\lambda_3^{(q)}, j_3)\}, & (\lambda_1^{(q)}, \lambda_2^{(q)}, \lambda_3^{(q)}) &\in \pi(i_1, i_2), \\
R_{3q}(s) &= \{(i_1, \gamma_1^{(q)}), (i_2, \gamma_2^{(q)}), (i_3, \gamma_3^{(q)})\}, & (\gamma_1^{(q)}, \gamma_2^{(q)}, \gamma_3^{(q)}) &\in \pi(j_1, j_2, j_3), \\
&& q &= 1, \dots, 6.
\end{aligned}$$

To check the relations

$$(37) \quad t_{R_{1q}(s)} = -2^{-N} \frac{P(N-1)}{P(m)}, \quad q = 1, 2,$$

$$(38) \quad t_{R_{2lq}(s)} = 2^{-2N+1} \frac{P(N-2)}{P(m)}$$

for $l = 1$ and $q = 1, \dots, 4$ or $l \in \{2, 3\}$ and $q = 1, \dots, 6$ we apply Theorem 1 and proceed in the same way as in the proof of (22)–(24).

Let us prove that

$$(39) \quad t_{R_{3q}(s)} = -2^{-3N+3} \frac{P(N-3)}{P(m)}, \quad q = 1, \dots, 6.$$

Let $R_{31}(s) = \{(i_1, j_1), (i_2, j_2), (i_3, j_3)\}$. Then relation (6) implies for $k = 3$ that

$$(40) \quad t_{R_{31}(s)} = -2^{-3N+2} \frac{P(N-3)}{P(m)} \sum_{c \in \gamma_0} (-1)^\tau$$

where γ_0 is the set of all $(N \times 2)$ matrices c , $c = (c_{ij})_{i \in I, j \in \{1, 2\}}$, of rank 2 in the field $GF(2)$ such that $c_{i_3 j_1} = c_{i_3 j_2} = 0$. Note that $\tau = c_{i_1 j_1} \oplus c_{i_2 j_2}$. We represent the set γ_0 as the union

$$\gamma_0 = \bigcup_{\mu=1}^{16} \gamma_{0,\mu}$$

of disjoint subsets $\gamma_{0,\mu} \subseteq \gamma_0$, $\mu = 1, \dots, 16$, such that for any matrix $c^{(\mu)} \in \gamma_{0,\mu}$,

$$c^{(\mu)} = \left(c_{ij}^{(\mu)} \right)_{i \in I, j \in \{1, 2\}},$$

the elements $c_{i_1 j_1}^{(\mu)}$ and $c_{i_1 j_2}^{(\mu)}, c_{i_2 j_1}^{(\mu)}, c_{i_2 j_2}^{(\mu)}$ are fixed, $\mu = 1, \dots, 16$, and moreover

$$\left\{ c_{i_1 j_1}^{(l)}, c_{i_1 j_2}^{(l)}, c_{i_2 j_1}^{(l)}, c_{i_2 j_2}^{(l)} \right\} \neq \left\{ c_{i_1 j_1}^{(t)}, c_{i_1 j_2}^{(t)}, c_{i_2 j_1}^{(t)}, c_{i_2 j_2}^{(t)} \right\}$$

for $l \neq t$.

Putting, for example,

$$c_{i_1 j_1}^{(1)} = c_{i_1 j_2}^{(1)} = c_{i_2 j_1}^{(1)} = c_{i_2 j_2}^{(1)} = 0,$$

we get $\tau = 0$ and $|\gamma_{0,1}| = (2^{N-3} - 1)(2^{N-3} - 2)$, since the total number of ways to place nonzero elements of the field $GF(2)$ to the first column of the matrix $c^{(1)} \in \gamma_{0,1}$ is $2^{N-3} - 1$ in the case of $c_{i_1 j_1}^{(1)} = c_{i_2 j_1}^{(1)} = c_{i_3 j_1}^{(1)} = 0$, while the same number is $2^{N-3} - 2$ for the second column linearly independent of the first. Similarly, putting $c_{i_1 j_1}^{(2)} = c_{i_2 j_1}^{(2)} = c_{i_1 j_2}^{(2)} = 0$ and $c_{i_2 j_2}^{(2)} = 1$, we get $\tau = 1$ and $|\gamma_{0,2}| = (2^{N-3} - 1)2^{N-3}$. Now we evaluate the sum

$$\sum_{c \in \gamma_0} (-1)^\tau = \sum_{\mu=1}^{16} \sum_{c \in \gamma_{0,\mu}} (-1)^\tau = 2.$$

The latter two equalities together with (40) prove (39). Substituting (37)–(39) into (36) we get (35).

5. AUXILIARY RESULTS FOR THE PROOF OF THEOREM 2

Lemma 1. *Let $R(s) = \{(i_1, j_1), \dots, (i_s, j_s)\}$ and $j_1 = \dots = j_s$, $s \geq 1$. Then*

$$t_{R(s)} = -2^{-N} \frac{P(N-1)}{P(m)}.$$

Proof. It follows from the hypothesis of Lemma 1 that $\{j_1, \dots, j_s\} = \{\mu\}$ for some $\mu \in J$. Thus the parameter k defined in Theorem 1 is equal to 1. Taking (6) and Remark 2 into account we complete the proof of Lemma 1. $\square$

Lemma 2. *If the set $R(s)$ satisfies (8), then*

1) *for $s_1 = s_2 = 0$ and $s_{12} \geq 1$*

$$(41) \quad t_{R(s)} = -2^{-N} \frac{P(N-1)}{P(m)};$$

2) *for $s_{12} = 0$, $s_1 \geq 1$, and $s_2 \geq 1$*

$$(42) \quad t_{R(s)} = 2^{-2N+1} \frac{P(N-2)}{P(m)};$$

3) *for $s_{12} \geq 1$ and $s_1 + s_2 \geq 1$ relation (42) holds for $t_{R(s)}$.*

Proof. Let $s_1 = s_2 = 0$ and $s_{12} \geq 1$. Then we apply (6) for $k = 2$ and obtain

$$(43) \quad t_{R(s)} = -2^{-2N+1} \frac{P(N-2)}{P(m)} \sum_{c \in \gamma_0} (-1)^\tau$$

where γ_0 is the set of all nonzero N -dimensional columns c , $c = (c_{i1})_{i \in I}$, in the field $GF(2)$ such that $\bigoplus_{i \in \zeta_{12}} c_{i1} = \tau = 0$. Hence

$$(44) \quad \sum_{c \in \gamma_0} (-1)^\tau = |\gamma_0|.$$

Further we show that

$$(45) \quad |\gamma_0| = 2^{N-1} - 1.$$

Indeed, since $\bigoplus_{i \in \zeta_{12}} c_{i1} = 0$, the positions $i \in \zeta_{12}$ of the vector c contain an even number of unit elements of the field $GF(2)$; the positions $i \in I \setminus \zeta_{12}$ may contain arbitrary elements of the field $GF(2)$ such that the N -dimensional column is nonzero. Thus

$$|\gamma_0| = 2^{s_{12}-1} 2^{N-s_{12}} - 1 = 2^{N-1} - 1$$

and relation (45) is proved. Relation (41) follows from (43)–(45).

Now let $s_{12} = 0$, $s_1 \geq 1$, and $s_2 \geq 1$. Using relation (6) for $k = 2$ we prove equality (43) where γ_0 is the set of all nonzero N -dimensional columns c , $c = (c_{i1})_{i \in I}$, in the field $GF(2)$ such that

$$(46) \quad \bigoplus_{i \in \zeta(\mu_2)} c_{i1} = 0;$$

$\tau = \bigoplus_{i \in \zeta(\mu_1)} c_{i1}$. Since equality (46) holds for 2^{s_2-1} families of elements of the field $GF(2)$, the number of cases where the parameter τ is equal to 0 is the same as that where τ is equal to 1, namely 2^{s_1-1} . The positions $i \in I \setminus \zeta_{12}$ of the column $c \in \gamma_0$ can

be filled in an arbitrary way except for the case where the N -dimensional column is zero. Therefore

$$(47) \quad \sum_{c \in \gamma_0} (-1)^\tau = \sum_{c \in \gamma_0^+} 1 - \sum_{c \in \gamma_0^-} 1 = -1$$

where $\gamma_0^+, \gamma_0^- \subseteq \gamma_0$ ($\gamma_0^-, \gamma_0^- \subseteq \gamma_0$) is the collection of all columns of the set γ_0 such that $\tau = 0$ ($\tau = 1$). To get (47) we used the equalities

$$\sum_{c \in \gamma_0^+} 1 = 2^{s_2-1} 2^{s_1-1} 2^{N-(s_1+s_2)} - 1 = 2^{N-2} - 1$$

and $\sum_{c \in \gamma_0^-} 1 = 2^{N-2}$.

Substituting (47) into (43) we prove (42) for $s_{12} = 0$, $s_1 \geq 1$, and $s_2 \geq 1$.

Finally we prove the last statement of Lemma 2. Let $s_{12} \geq 1$ and $s_1 + s_2 \geq 1$. Then relation (43) holds with γ_0 the collection of all nonzero N -dimensional columns c , $c = (c_{i1})_{i \in I}$, in the field $GF(2)$ such that

$$(48) \quad \left(\bigoplus_{i \in \zeta_{12}} c_{i1} \right) \oplus \left(\bigoplus_{i \in \zeta(\mu_2) \setminus \zeta_{12}} c_{i1} \right) = 0;$$

$\tau = (\bigoplus_{i \in \zeta_{12}} c_{i1}) \oplus (\bigoplus_{i \in \zeta(\mu_1) \setminus \zeta_{12}} c_{i1})$. It follows from (48) that the number of unit elements among the terms of the sum $\bigoplus_{i \in \zeta_{12}} c_{i1}$ is even if and only if the number of unit elements among the terms of the sum $\bigoplus_{i \in \zeta(\mu_2) \setminus \zeta_{12}} c_{i1}$ is even. This easily implies relation (47). Indeed,

$$(49) \quad \sum_{c \in \gamma_0^+} 1 = b_1 + b_2$$

if $s_{12} \geq 1$, $s_1 \geq 1$, and $s_2 \geq 1$ where b_1 (b_2) is the total number of ways to place elements of the field $GF(2)$ to a nonzero N -dimensional column such that the number of unit elements in positions $i \in \zeta_{12}$, $i \in \zeta(\mu_1) \setminus \zeta_{12}$, $i \in \zeta(\mu_2) \setminus \zeta_{12}$ is even (odd). Obviously

$$b_1 = 2^{s_1-1} 2^{s_{12}-1} 2^{s_2-1} 2^{N-(s_1+s_2+s_{12})} - 1 = 2^{N-3} - 1, \quad b_2 = 2^{N-3}.$$

Therefore

$$(50) \quad \sum_{c \in \gamma_0^+} 1 = 2^{N-2} - 1.$$

In a similar way we obtain

$$(51) \quad \sum_{c \in \gamma_0^-} 1 = 2^{N-2}.$$

Relations (50) and (51) imply (47) for $s_{12} \geq 1$, $s_1 \geq 1$, and $s_2 \geq 1$.

If $s_{12} \geq 1$, $s_1 = 0$, and $s_2 \geq 1$, then

$$b_1 = 2^{s_2-1} 2^{s_{12}-1} 2^{N-(s_2+s_{12})} - 1 = 2^{N-2} - 1$$

and $b_2 = 0$ in equality (49), whence

$$(52) \quad \sum_{c \in \gamma_0^+} 1 = 2^{N-2} - 1.$$

Similarly we obtain

$$(53) \quad \sum_{c \in \gamma_0^-} 1 = 2^{N-2}.$$

Relations (52) and (53) prove equality (47) for $s_{12} \geq 1$, $s_1 = 0$, and $s_2 \geq 1$.

Finally if $s_{12} \geq 1$, $s_1 \geq 1$, and $s_2 = 0$, then

$$b_1 = 2^{s_1-1} 2^{s_{12}-1} 2^{N-(s_1+s_{12})} - 1 = 2^{N-2} - 1, \quad b_2 = 0$$

in equality (49) and thus $\sum_{c \in \gamma_0^+} 1 = 2^{N-2} - 1$. The equality $\sum_{c \in \gamma_0^-} 1 = 2^{N-2}$ is easy to prove. Therefore (47) is proved for $s_{12} \geq 1$ and $s_1 + s_2 \geq 1$. It follows from (47) and (43) that (42) holds for $s_{12} \geq 1$ and $s_1 + s_2 \geq 1$. Lemma 2 is proved. $\square$

Lemma 3. *If condition (1) holds, then*

$$f^{(0)}(x_{ij}, i \in I, j \in J) = \frac{P(N)}{P(m)}$$

for $N \geq n \geq 1$.

Lemma 3 is proved in [2].

6. PROOF OF THEOREM 2

Statement (i) of Theorem 2 can easily be proved by equality (2) for $n = 1$ and by Lemmas 1 and 3.

We prove statement (ii). Using representation (2) and (4) we find for $N \geq 2$ and $n = 2$ that

$$\begin{aligned}
 P\{\chi(\mathbf{A}) = 1\} &= f^{(0)}(x_{ij}, i \in I, j \in \{1, 2\}) \\
 &+ \sum_{s=1}^N \varepsilon^s \sum_{1 \leq i_1 < \dots < i_s \leq N} \left[\sum_{j=1}^2 t_{R_j(s)} \prod_{q=1}^s x_{i_q j} + \varepsilon^s t_{R_3(s)} \prod_{j=1}^2 \prod_{q=1}^s x_{i_q j} \right] \\
 &+ \sum_{s=2}^{2N} \varepsilon^s \sum_{1 \leq i_1 < \dots < i_s \leq N} 1 \\
 &\times \sum_{\substack{1 \leq i'_1 < \dots < i'_{s_{12}} \leq N \\ i' \notin \{i_1, \dots, i_{s_1}\}}} \sum_{\substack{1 \leq i''_1 < \dots < i''_{s_2} \leq N \\ i'' \notin \{i_1, \dots, i_{s_1}, i'_1, \dots, i'_{s_{12}}\}}} t_{R_4(s)} \left(\prod_{q=1}^{s_1} x_{i_q 1} \right) \\
 &\times \left(\prod_{j=1}^2 \prod_{q=1}^{s_{12}} x_{i'_q j} \right) \left(\prod_{q=1}^{s_2} x_{i''_q 2} \right)
 \end{aligned} \tag{54}$$

in view of condition (1) where

$$R_j(s) = \{(i_1, j), \dots, (i_s, j)\}, \quad j \in \{1, 2\},$$

$$R_3(s) = \{(i_1, 1), \dots, (i_s, 1), (i_1, 2), \dots, (i_s, 2)\},$$

$$R_4(s) = \{(i_1, 1), \dots, (i_{s_1}, 1), (i'_1, 1), \dots, (i'_{s_{12}}, 1), (i'_1, 2), \dots, (i'_{s_{12}}, 2), (i''_1, 2), \dots, (i''_{s_2}, 2)\}.$$

Taking Lemma 1 into account we obtain for $j \in \{1, 2\}$ that

$$(55) \quad t_{R_j(s)} = -2^{-N} (1 - 2^{-N+1}).$$

Lemma 2 implies that

$$(56) \quad t_{R_3(s)} = -2^{-N} (1 - 2^{-N+1})$$

and

$$(57) \quad t_{R_4(s)} = 2^{-2N+1}.$$

By Lemma 3

$$(58) \quad f^{(0)}(x_{ij}, i \in I, j \in \{1, 2\}) = (1 - 2^{-N}) (1 - 2^{-N+1}).$$

Substituting (55)–(58) into (54) we prove statement (ii). Theorem 2 is proved. $\square$

7. CONCLUDING REMARKS

Theorems 1 and 2 together with results of [1]–[3] allow one to find the distribution of the rank of an $(N \times n)$ matrix whose entries are independent nonidentically distributed random variables assuming values in the field $GF(2)$. Matrices with nonidentically distributed entries for which the difference between their distributions and the equiprobable distribution on $GF(2)$ is small appear not only in the theory ([4]–[6]) but also in some applied problems (say, when testing the quality of pseudorandom $(0, 1)$ -sequences). One of the results in [4]–[6] is that, under certain conditions, the limit distribution (as $n \rightarrow \infty$) of the rank of a random Boolean matrix is invariant and coincides with that in the case of the equiprobable distribution on $GF(2)$. At the same time, the use of the asymptotic results for finding the probability that a finite Boolean random matrix has maximal rank leads to a certain error, which can be “remedied” by using the results presented in this paper.

BIBLIOGRAPHY

1. V. V. Masol, *An expansion in a small parameter of the probability that a random determinant in the field $GF(2)$ is 1*, Teor. Imovirnost. Mat. Stat. **64** (2001), 102–105; English transl. in Theor. Probability Math. Statist. **64** (2002), 117–121. MR1922957 (2003g:60015)
2. V. V. Masol, *Explicit representation of some coefficients in the expansion of the random matrix rank distribution in the field $GF(2)$* , Theory Stoch. Process. **6(22)** (2000), no. 3–4, 122–126.
3. V. V. Masol, *Expansion in terms of powers of small parameter of the maximum rank distribution of a random Boolean matrix*, Kibernetika i Sistemnyi Analiz **38** (2002), no. 6, 176–180; English transl. in Cybernetics and Systems Analysis **38** (2003), no. 6, 938–942.
4. I. N. Kovalenko, *Invariance theorems for random Boolean matrices* Kibernetika **11** (1975), no. 5, 138–152; English transl. in Cybernetics **11** (1976), no. 5, 818–834. MR0458552 (56:16752)
5. A. A. Levitskaya, *Invariance theorems for a system of random linear equations over an arbitrary finite ring*, Dokl. AN SSSR **263** (1982), no. 2, 289–291; English transl. in Soviet Math. Dokl. **25** (1982), 340–342. MR0650154 (83g:60046)
6. C. Cooper, *On the rank of random matrices*, Random Structures and Algorithms **16** (2000), no. 2, 209–232. MR1742352 (2000k:15050)

DEPARTMENT OF PROBABILITY THEORY AND MATHEMATICAL STATISTICS, MECHANICS AND MATHEMATICS FACULTY, NATIONAL TARAS SHEVCHENKO UNIVERSITY, ACADEMICIAN GLUSHKOV AVENUE 6, KYIV 03127, UKRAINE

E-mail address: vicamasol@pochtamt.ru

Received 15/APR/2003

Translated by V. SEMENOV