Singular 0/1-matrices, and the hyperplanes spanned by random 0/1-vectors

Thomas Voigt*        Günter M. Ziegler**
Inst. Mathematics, MA 6-2
TU Berlin, D-10623 Berlin, Germany
{tvoigt,ziegler}@math.tu-berlin.de

Abstract

Let $P_s(d)$ be the probability that a random 0/1-matrix of size $d \times d$ is singular, and let $E(d)$ be the expected number of 0/1-vectors in the linear subspace spanned by $d-1$ random independent 0/1-vectors. (So $E(d)$ is the expected number of cube vertices on a random affine hyperplane spanned by vertices of the cube.)

We prove that bounds on $P_s(d)$ are equivalent to bounds on $E(d)$: $P_s(d) = \left(2^{-d}E(d) + \frac{d^2}{2^{d+1}}\right)(1 + o(1))$.

We also report about computational experiments pertaining to these numbers.

1 Introduction

0/1-polytopes arise naturally in a great variety of interesting contexts, including a prominent role in combinatorial optimization, yet some basic characteristics of “typical” (that is, random) 0/1-polytopes are unknown. (For a survey of a variety of aspects of 0/1-polytopes see [10].)

One of the key open questions in this context is rather notorious:

- Pick $d+1$ random vertices of the $d$-cube independently (with respect to the uniform distribution). What is the probability that these vectors do not form a $d$-simplex?

If we assume without loss of generality that one of these points is the origin $0$ the question can be rephrased: Let $C^d = [0,1]^d$ be the $d$-dimensional unit hypercube, and let

\[ \mathcal{M}_d := \{0,1\}^{d \times d} \]

be the set of all 0/1-matrices of size $d \times d$.

- What is the asymptotic behaviour of the probability

\[ P_s(d) := \text{Prob} [\det(M) = 0 \mid M \in \mathcal{M}_d] \]

that a random square $d$-dimensional 0/1-matrix is singular?

*Supported by Deutsche Forschungs-Gemeinschaft (DFG) ZI 475/3
**Partially supported by Deutsche Forschungs-Gemeinschaft (DFG), FZT86 and ZI 475/4
This central but difficult question has received careful attention; see Komlós [6], Bollobás [2], Kahn, Komlós & Szemerédi [5]. It has been conjectured that

\[ P_s(d) = \frac{d^2}{2^d} (1 + o(1)), \]

which is essentially the probability that two rows or two columns of a random matrix are equal. However, the known upper bounds are far off this mark; currently the best upper bound is \( P_s(d) < (1 - \varepsilon)^d \), for some rather small \( \varepsilon > 0 \). (This was proved by Kahn, Komlós and Szemerédi in [5] with \( \varepsilon = 0.001 \).)

A closely related problem is as follows:

- Given \( r \) random vertices \( v_1, \ldots, v_r \) of \( C^d \), what is the expected number of 0/1-vectors in the affine subspace spanned by these vectors?

Improving a result by Odlyzko [7], Kahn, Komlós & Szemerédi derived in [5] that there exists a constant \( C \) independent from \( d \) such that the probability that such an affine subspace contains any 0/1-vector other than \( v_1, \ldots, v_r \) is \( 4\binom{r}{3} (\frac{3}{4})^d (1 + o(1)) \), provided that \( r < d - C \). However, so far no results were known for the case \( r = d \).

In this paper we will show that determining the expected number of vertices of \( C^d \) in the affine subspace spanned by \( d \) random vertices of \( C^d \) is just as hard as determining \( P_s(d) \). More precisely, let \( \mathcal{G} \) denote the set of all linearly independent \((d-1)\)-sets of 0/1-vectors of length \( d \) and for a set \( S \) of arbitrary vectors let \( v(S) \) be the number of 0/1-vectors in the linear subspace spanned by \( S \). Then the following theorem holds.

**Theorem 1.1.** Let

\[ E(d) := \frac{1}{|\mathcal{G}|} \sum_{G \in \mathcal{G}} v(G) \]

be the expected number of 0/1-points on the hyperplane spanned by a random linearly independent set of \( d-1 \) 0/1-vectors. Then

\[ P_s(d) = \left( \frac{1}{2^d} E(d) + \frac{d^2}{2^{d+1}} \right) (1 + o(1)). \]

We can give a (trivial) lower bound for \( E(d) \) by just considering the \( \binom{d}{2} + d \) “fat” hyperplanes (faces \( x_i = 0 \) and hyperplanes \( x_i - x_j = 0 \)) containing \( 2^{d-1} \) vertices each. Since \( d-1 \) points chosen randomly from such a hyperplane span the hyperplane with probability \( 1 - (1 - \varepsilon)^{d-1} \) (according to [5]) it is easy to verify that \( E(d) \geq \frac{d^2}{2} (1 + o(1)) \).

In fact the conjectured upper bounds on \( P_s(d) \) and \( E(d) \) are strictly equivalent:

**Corollary 1.2.** As \( d \to \infty \),

\[ P_s(d) = \frac{d^2}{2^d} (1 + o(1)) \]

if and only if

\[ E(d) = \frac{d^2}{2} (1 + o(1)). \]
Using symmetry we could switch to an affine version, replacing $G$ by the set of affinely independent $d$-sets of 0/1-vectors and checking the expected value of 0/1-vectors in a hyperplane spanned by such a set. However, for the purpose of this paper the linear version will be more convenient to handle; so we will consider only hyperplanes containing the origin $0$.

To our knowledge the problem of determining the expected number of 0/1-vectors on a hyperplane $h$ spanned by random vertices of $C^d$ has not been studied independently yet. Some basic results were derived in [2] and [5] by examining the structure of the defining equations $a$ for planes $h = \{x \in \mathbb{R}^d \mid a'x = 0\}$ (which is perhaps the most natural approach). The lemma of Littlewood-Offord (see Section 2) is a classical tool: It states that if all $a_j$ are nonzero then the number of 0/1-points in this plane is at most $\binom{d}{\lfloor d/2 \rfloor}$. If the coefficients satisfy additional conditions, this number can be reduced considerably (see Halász [3] [4]). In order to obtain such conditions it would be of considerable interest to learn more about the distribution of determinants of 0/1-matrices: If $d - 1$ vectors span a hyperplane and we write these vectors into a $d \times (d - 1)$ matrix $M$, then a defining equation $a'x = 0$ is given by $a_j = (-1)^j \det(r_j(M))$, where $r_j(M)$ is the matrix obtained from $M$ by deleting the $j$-th row.

The rest of this paper is organized as follows: In Section 2 we state some consequences of the Littlewood-Offord lemma. The proof of Theorem 1.1 is given in Section 3. In Section 4 we present some experimental estimates of $P_s(d)$ for $d \leq 30$.

Some definitions.
We use standard vector notation $a = (a_1, \ldots, a_d)^t$, where $d$ denotes the dimension. The expected value of a random variable $X$ is denoted by $E[X]$; the probability of an event $Y$ is $\text{Prob}[Y]$. Define $r(F)$ as the (linear) rank of a family or set of vectors $F$.

The next definition is useful for partitioning sets of matrices into subsets with “nice” properties and was frequently used in the analysis of 0/1-matrices (see [2] or [5]). Given a $d \times d$ matrix $M$ we define the strong rank $\overline{r}(M)$ as the largest $k \leq d$ such that all $k$-subsets of columns from $M$ are independent. (Equivalently, it is the largest $k$ such that the truncation to rank $k$ of the matroid given by the columns of the matrix $m$ is uniform of rank $k$.) We also consider the strong rank of sets and of families of $d$-dimensional vectors.

2 The Littlewood-Offord lemma

The “Littlewood-Offord lemma” is a classical tool [2] [7] for obtaining upper bounds on $P_s(d)$.

Lemma 2.1 (Littlewood-Offord). Let $s \in \mathbb{R}$, $n \in \mathbb{N}$ and let $a_i \in \mathbb{R}$ with $|a_i| \geq 1$ for $1 \leq i \leq n$. Then at most $\binom{n}{\lfloor n/2 \rfloor}$ of the $2^n$ sums $\sum_{i=1}^{n} \varepsilon_i a_i$, $\varepsilon_i = \pm 1$ fall in the open interval $(s - 1, s + 1)$.

Corollary 2.2. Let $a_i \in \mathbb{R}, i = 1, \ldots, n$ with at least $t$ of the $a_i$ nonzero. Then at most $\binom{t}{\lfloor t/2 \rfloor}2^{n-t} \approx \frac{2^n}{\sqrt{\pi t}}$ of the $2^n$ sums $\sum_{i=1}^{n} \varepsilon_i a_i$, $\varepsilon_i \in \{0,1\}$ can have the same value.
As observed in [5], this lemma suffices to show that with very high probability the strong rank of a random 0/1-matrix is either close to \(d\) or at most 1.

**Lemma 2.3.** Let \(M \in \mathcal{M}_d\) be a random matrix. Let \(E\) be the event that \(M\) has a \(d \times (k+1)\) submatrix of strong rank \(k\) for some \(k \in \{2, \ldots, d - \frac{d}{\ln(d)}\}\). Then for large \(d\),

\[
\text{Prob}[E] \leq 2^{-d}.
\]

**Proof.** The proof follows [2, Chapter 14.2] (see also [5, Section 3.1]) and is sketched here for the reader’s convenience.

Let \(M\) be a random 0/1-matrix and \(k < d\). If \(M\) contains \(k + 1\) columns \(c_1, \ldots, c_{k+1}\) of strong rank \(k\) then clearly we can find a \(k \times (k+1)\) submatrix of \(M\) of strong rank \(k\) by deleting \(d - k\) linearly dependent rows from \((c_1, \ldots, c_{k+1})\).

If we want to upper bound the probability that arbitrarily chosen columns \(c_1, \ldots, c_{k+1}\) have strong rank \(k\), then it suffices to give an upper bound on the probability that \(c_1, \ldots, c_{k+1}\) have rank \(k\) conditioned on the event that an arbitrary \(k \times (k+1)\) submatrix \(\tilde{M}\) of \((c_1, \ldots, c_{k+1})\) has strong rank \(k\):

\(\tilde{M}\) has strong rank \(k\) if and only if the last column of \(\tilde{M}\) is a unique linear combination of the first \(k\) columns and all coefficients in this combination are non-zero. Under this condition the probability that any of the remaining \(d - k\) rows of \(c_1, \ldots, c_{k+1}\) satisfy the linear dependency equation defined by \(\tilde{M}\) is at most \(2^{-k\left(\frac{k}{\lceil k/2 \rceil}\right)}\) by Lemma 2.2, so the probability that \(c_1, \ldots, c_{k+1}\) have rank \(k\) is at most \((2^{-k\left(\frac{k}{\lceil k/2 \rceil}\right)})^{d-k}\). Since there are at most \(\binom{d}{k}\binom{d}{k+1}\) such submatrices \(\tilde{M}\) we find

\[
\text{Prob}[\tilde{\Phi}(M) = k] \leq \binom{d}{k}\binom{d}{k+1}\left(2^{-k\left(\frac{k}{\lceil k/2 \rceil}\right)}\right)^{d-k}.
\]

We derive

\[
\sum_{k=3}^{\lfloor d/\ln(d) \rfloor} \binom{d}{k}\binom{d}{k+1}\left(2^{-k\left(\frac{k}{\lceil k/2 \rceil}\right)}\right)^{d-k} \leq 2^{-d}
\]

by checking that each summand in (1) is at most \(\frac{1}{d^{2d}}\) if \(d\) is large (using Stirling’s formula and elementary, but somewhat tedious calculations).

To complete the proof of Lemma 2.3, we observe that the event \(\tilde{\Phi}(M) = 2\) depends on the existence of three columns \(m_i, m_j, m_k\) such that \(m_i + m_j = m_k\), which happens with probability \(\Theta(d^3(\frac{3}{\ln(d)})^d)\).

\[\square\]

**Corollary 2.4.** Let \(M \in \mathcal{M}_d\) be a random matrix. Then

\[
\text{Prob}\left[\tilde{\Phi}(M) \leq d - 3\frac{d}{\ln(d)}\right] \leq \frac{d^2}{2^{d+1}}(1 + o(1)).
\]
3 Proof of Theorem 1.1

Let $S \subset \mathcal{M}_d$ be the set of singular matrices and $R = \mathcal{M}_d \setminus S$. We will partition $S$ into subsets $S_j \subset S, j \in \{1, \ldots, 4\}$ and derive precise bounds on the sizes of two of these sets in terms of $|\mathcal{G}|$ and $E(d)$. The other two sets are small. This allows us to estimate the value $P_s(d) = \frac{|S|}{|S| + |R|}$.

Let $N_d := \left\lfloor d - \frac{3d}{\ln(d)} \right\rfloor$ and partition $S$ into the disjoint sets

- $S_1 := \{M \in \mathcal{M}_d \mid r(M) = d - 1, \bar{r}(M) = 1\}$
- $S_2 := \{M \in \mathcal{M}_d \mid r(M) = d - 1, \bar{r}(M) > N_d\}$
- $S_3 := \{M \in \mathcal{M}_d \mid \bar{r}(M) \in \{0, 2, \ldots, N_d\}\}$
- $S_4 := \{M \in \mathcal{M}_d \mid r(M) < d - 1, \bar{r}(M) = 1 \text{ or } \bar{r}(M) > N_d\}$

We will give precise estimates for the sizes of the sets $R, S_1, S_2$, and check that the sets $S_3$ and $S_4$ are small enough. More precisely, we will show that

$$|R| = |\mathcal{G}| \frac{2d - E(d)}{d^2}$$
$$|S_1| = |\mathcal{G}| \frac{d - 1}{2d}$$
$$|S_2| = |\mathcal{G}| \frac{E(d)}{d} (1 + o(1))$$
$$|S_1| \leq |S_2| (1 + o(1))$$
$$|S_3| \leq \frac{c_1}{d} |S_1|$$
$$|S_4| \leq \frac{c_2}{\sqrt{d}} (|S_1| + |S_2|)$$

for some constants $c_1, c_2 > 0$.

- While most matrices from $\mathcal{M}_d$ with two equal columns are in $S_1$, most matrices with two equal rows lie in $S_2$. To see this, pick a random $(d - 1) \times d$ matrix $N = (n_1, \ldots, n_d)$. Using the result of Kahn, Komlós and Szemerédi [5] that $P_s(d) \leq (1 - \varepsilon)^d$ for some $\varepsilon \geq 0.001$, we obtain $d(1 - \varepsilon)^{d-1}$ as an upper bound on the probability that at least one of the $(d - 1) \times (d - 1)$ submatrices $c_j(N)$ is singular, where $c_j(N)$ is the matrix obtained from $N$ by deleting the $j$-th column $n_j$. Cramer’s rule gives $\sum_{j=1}^{d} (-1)^j d_j n_j = 0$ for the determinants $d_j = \det(c_j(N))$. Thus, $N$ has strong rank $d - 1$ if all determinants are nonzero, which establishes (3):

$$|S_1| \leq |S_2| (1 + o(1))$$

- By Lemma 2.3 a random matrix $M \in \mathcal{M}_d$ lies in $S_3$ with probability at most $(d + 1) 2^{-d}$. The probability that two columns are equal is $d^2 2^{-d-1} (1 + o(1))$. Again almost all matrices with two identical columns have strong rank $d - 1$ and are in $S_1$ (up to an exponentially small subset), which implies (6):

$$|S_3| = O\left(\frac{1}{d} |S_1|\right)$$
For each matrix $M \in \mathcal{R} \cup \mathcal{S}_1 \cup \mathcal{S}_2$ there is at least one $G \in \mathcal{G}$ that is a subset of the column set of $M$. The estimates (2), (3) and (4) are obtained by examining this in detail:

- For each $G \in \mathcal{G}$ we have exactly $\frac{d!}{2}(d-1)$ matrices from $\mathcal{S}_1$ containing only columns from $G$ (since we have $d-1$ choices for a duplicate column and $\frac{d!}{2}$ permutations). This gives (3):

$$|\mathcal{S}_1| = |\mathcal{G}| \frac{d!}{2}(d-1)$$

- For any fixed $G \in \mathcal{G}$ we can construct $d!(v(G) - d)$ different matrices $S \in \mathcal{S}_2 \cup \mathcal{S}_3$ (using columns from $G$ and an additional nonzero column in the span of $G$ that is not in $G$). Summing over $G \in \mathcal{G}$ we obtain $d!E(d)(1 + o(1))|\mathcal{G}|$ matrices in $\mathcal{S}_2$, since (5) and (6) imply that $|\mathcal{S}_3|$ is small compared to $|\mathcal{S}_2|$. On the other hand each matrix $M \in \mathcal{S}_2$ is constructed $d$ times. This gives (4):

$$|\mathcal{S}_2| = \frac{1}{d - o(d)}d!E(d)(1 + o(1))|\mathcal{G}|$$
$$= d! \frac{E(d)}{d}(1 + o(1))|\mathcal{G}|.$$  

- Similarly, we get $d!(2^d - E(d))|\mathcal{G}|$ matrices in $\mathcal{R}$ and each matrix $M \in \mathcal{R}$ is constructed $d$ times. This gives (2):

$$|\mathcal{R}| = \frac{d!}{d} (2^d - E(d)) |\mathcal{G}|.$$  

A little more work is required for the upper bound (7) on $|\mathcal{S}_4|$. So far we established an upper bound on the number of matrices of rank $d - 1$ in terms of the number of regular matrices. A similar argument will be used to show that for any $k \leq d - 2$ there are significantly fewer matrices of rank $k$ than matrices of rank $k + 1$, which gives the desired result:

(i) First consider the matrices $\hat{\mathcal{S}}$ with the property that the rows or the columns admit more than one trivial dependency (i.e. zero-vectors or pairs of identical vectors). This probability is dominated by the probability that a matrix has two pairs of identical rows or columns, which happens with probability $O\left(\frac{d}{4}2^{-2d}\right)$, so clearly $|\mathcal{S}_4 \cap \hat{\mathcal{S}}|$ is exponentially smaller than $\frac{1}{\sqrt{d}}(|\mathcal{S}_1| + |\mathcal{S}_2|)$.

(ii) Let $\hat{\mathcal{S}}$ be the set of matrices whose columns or rows have a subset with strong rank in $\{2, \ldots, N_d\}$. Lemma 2.3 gives that this happens with probability of at most $2^{-d}$, while the probability that two columns are equal is $d^2 2^{-d-1}(1 + o(1))$. This implies $|\hat{\mathcal{S}}| \leq O\left(\frac{1}{d^2} |\mathcal{S}_1|\right)$.

(iii) To estimate the number of the remaining matrices in $\mathcal{S}_4$, we use similar techniques as in [2, Chapter 14.2]:

\[\]
We can use the Littlewood-Offord lemma to give an upper bound on the number of 0/1-vectors in the span of a set of vectors \( C \): Let \( a \) be in the orthogonal space of \( C \), i.e. \( a^T c = 0 \) for all \( c \in C \). Clearly all vectors \( v \) in the span of \( C \) satisfy \( a^T v = 0 \). If \( s \) is the number of nonzero entries in \( a \) then Lemma 2.2 assures us that the span of \( C \) contains at most \( \left( \binom{s}{\lfloor s/2 \rfloor} 2^{d-s} \right) 0/1 \)-vectors.

Let \( S_4(k) \) be the matrices in \( S_4 \setminus (\bar{S} \cup \bar{S}) \) of rank \( k \). For a fixed \( k \leq d - 2 \) and \( m \in S_4(k) \) we know that the columns and rows of \( m \) admit at most one trivial dependency (by excluding \( \bar{S} \)) and that neither rows nor columns have a submatrix of strong rank between 2 and \( N_d \) (by excluding \( \bar{S} \)). Thus both \( \ker(m) \) and \( \ker(m^T) \) contain vectors with more than \( N_d \) nonzero entries, since they are at least 2-dimensional. Choose any such vectors \( a \in \ker(m) \) and \( b \in \ker(m^T) \).

If \( m \) is chosen uniformly at random from \( S_4(k) \), then the probability that \( a_d \neq 0 \) is at least \( \frac{N_d}{d} = 1 - \frac{3}{\log d} \). If we condition on this event (that the last column of \( m \) is a nontrivial linear combination of the remaining columns) and consider all 0/1-matrices having the same first \( d-1 \) columns as \( m \), then (by the observation above) at most \( 2^{d-N_d} \binom{N_d}{\lfloor N_d/2 \rfloor} \) of these matrices have rank \( k \), since the last column \( v \) has to satisfy \( b^T v = 0 \). Stirling’s formula implies that \( 2^{d-N_d} \binom{N_d}{\lfloor N_d/2 \rfloor} \approx 2^d \sqrt{\frac{2}{\pi N_d}} = O\left( \frac{1}{\sqrt{d}} 2^d \right) \).

Removing the condition \( a_d \neq 0 \) changes the number of matrices only by a factor of \( 1 + \frac{3}{\log d} \), so we find that

\[
|S_4(k)| = \begin{cases} 
O\left( \frac{1}{\sqrt{N_d}} |S_4(k+1)| \right) & \text{if } k < d - 2, \\
O\left( \frac{1}{\sqrt{N_d}} |S_1| + |S_2| \right) & \text{if } k = d - 2.
\end{cases}
\]

This establishes (7):

\[
|S_4| \leq \frac{c_2}{\sqrt{d}} (|S_1| + |S_2|)
\]

for some constant \( c_2 > 0 \).

Thus we have

\[
P_s(d) = \frac{|S|}{|R| + |S|} = \frac{(|S_1| + |S_2|)(1 + o(1))}{|R| + (|S_1| + |S_2|)(1 + o(1))} = \frac{\left( d-1 \right)^{\frac{d-1}{2} + \frac{E(d)}{d}}}{\left( \frac{d-1}{2} + \frac{E(d)}{d} + \frac{2^d - E(d)}{d} \right)} (1 + o(1)) = \left( \frac{1}{2^d} E(d) + \frac{d^2}{2^{d+1}} \right) (1 + o(1))
\]

This concludes the proof of Theorem 1.1.

\[\square\]
4 Experiments in small dimensions

Complete enumeration of the 0/1-matrices of size $d \times d$ is feasible up to dimension 7 (see [110]), while hyperplanes were enumerated up to dimension 8 (see Aichholzer & Aurenhammer [1]). For some higher dimensions we generated 25,000,000 random matrices and determined an experimental probability $P_x(d)$ that a random matrix is singular. The significance of these numbers is limited for high dimensions (we found very few singular matrices and 25 million is tiny compared to the number of 0/1-matrices), but since the number of singular matrices is sharply concentrated around the expected value the results should still be close to the real values. Up to dimension 17 $P_x(d)$ decreases at a slower rate than the natural lower bound $d^2 2^{-d}$ while in higher dimensions $P_x(d)$ seems to approach this bound.

| $d$ | matrices | singular | $P_x(d)$ | $\frac{d^2}{2^d}$ | $P_x(d) 2^d d^{-2}$ |
|-----|----------|----------|----------|------------------|---------------------|
| 1   | $2^1$    | 1        | 0.500000 | 0.500000         | 1.000              |
| 2   | $2^2$    | 10       | 0.625000 | 1.000000         | 0.625              |
| 3   | $2^3$    | 338      | 0.660156 | 1.125000         | 0.587              |
| 4   | $2^{16}$ | 42976    | 0.655761 | 1.000000         | 0.666              |
| 5   | $2^{25}$ | 21040112 | 0.627044 | 0.781250         | 0.803              |
| 6   | $2^{36}$ | 3.98 $\cdot$ 10$^{10}$ | 0.580372 | 0.562500         | 1.032              |
| 7   | $2^{49}$ | 2.92 $\cdot$ 10$^{14}$ | 0.519766 | 0.382812         | 1.358              |
| 8   | 25000000 | 11230864 | 0.449234 | 0.250000         | 1.797              |
| 9   | 25000000 | 9331895  | 0.373275 | 0.158203         | 2.359              |
| 10  | 25000000 | 7430305  | 0.297212 | 0.097656         | 3.043              |
| 11  | 25000000 | 5657196  | 0.226287 | 0.059082         | 3.830              |
| 12  | 25000000 | 4108304  | 0.164332 | 0.035156         | 4.674              |
| 13  | 25000000 | 2837245  | 0.113489 | 0.020629         | 5.501              |
| 14  | 25000000 | 1868850  | 0.074754 | 0.011962         | 6.249              |
| 15  | 25000000 | 1175425  | 0.047017 | 0.006865         | 6.847              |
| 16  | 25000000 | 707571   | 0.028303 | 0.003906         | 7.246              |
| 17  | 25000000 | 407077   | 0.016283 | 0.002204         | 7.385              |
| 18  | 25000000 | 225820   | 0.009032 | 0.001236         | 7.308              |
| 19  | 25000000 | 121157   | 0.004846 | 0.000686         | 7.038              |
| 20  | 25000000 | 62500    | 0.002500 | 0.000381         | 6.554              |
| 21  | 25000000 | 31779    | 0.001271 | 0.000210         | 6.045              |
| 22  | 25000000 | 15393    | 0.000615 | 0.000115         | 5.336              |
| 23  | 25000000 | 7383     | 0.000295 | 0.000063         | 4.683              |
| 24  | 25000000 | 3515     | 0.000140 | 0.000034         | 4.095              |
| 25  | 25000000 | 1722     | 0.000069 | 0.000018         | 3.698              |
| 26  | 25000000 | 736      | 0.000029 | 0.000010         | 2.923              |
| 27  | 25000000 | 345      | 0.000013 | 0.000005         | 2.541              |
| 28  | 25000000 | 164      | 0.000006 | 0.000003         | 2.246              |
| 29  | 25000000 | 81       | 0.000003 | 0.000002         | 2.068              |
| 30  | 25000000 | 37       | 0.000001 | 0.000000         | 1.766              |
Note added in proof:
Recently, T. Tao and V. H. Vu [8] have significantly improved the upper bound on $P_s(d)$, by proving that $P_s(d) = \left(\frac{3}{4} + o(d)\right)^d$.
Furthermore, M. Živković has recently computed the number of singular 0/1-matrices of size $8 \times 8$ exactly [9]. From this we get that $P_s(8) = 0.4492003726$, so our estimate $P_s(8) = 0.4492346$ wasn’t bad.

References

[1] O. Aichholzer and F. Aurenhammer, Classifying hyperplanes in hypercubes, SIAM J. Discrete Math., 9 (1996), pp. 225–232.

[2] B. Bollobás, Random Graphs, Academic Press, London, 1985.

[3] G. Halász, On the distribution of additive arithmetic functions, Acta Arith., 27 (1975), pp. 143–152.

[4] G. Halász, Estimates for the concentration function of combinatorial number theory and probability, Period. Math. Hung., 8 (1977), pp. 197–211.

[5] J. Kahn, J. Komlós, and E. Szemerédi, On the probability that a random ±1-matrix is singular, J. Amer. Math. Soc., 8 (1995), pp. 223–240.

[6] J. Komlós, On the determinant of (0, 1) matrices, Stud. Sci. Math. Hung., 2 (1967), pp. 7–21.

[7] A. Odlyzko, On subspaces spanned by random selections of ±1 vectors, J. Comb. Theory, Ser. A, 47 (1988), pp. 124–133.

[8] T. Tao and V. H. Vu, On the singularity probability of random Bernoulli matrices, J. Amer. Math. Soc., 20 (2007), pp. 603–628. arXiv:math.CO/0411095.

[9] M. Živković, Classification of small (0, 1) matrices, 2005. arXiv.org:math/0511636.

[10] G. M. Ziegler, Lectures on 0/1-polytopes, in Polytopes — Combinatorics and Computation, G. Kalai and G. M. Ziegler, eds., vol. 29 of DMV Seminar, Birkhäuser-Verlag, Basel, 2000, pp. 1–41.