Combinatorics, Probability and Computing (2001) 10, 137–157. Printed in the United Kingdom
© 2001 Cambridge University Press

Singularity Probabilities for Random Matrices over Finite Fields

JEFF KAHN[1]† and JÁNOS KOMLÓS[2]‡

[1] Department of Mathematics and RUTCOR,
Rutgers University, New Brunswick, NJ 08903, USA
(e-mail: jkahn@math.rutgers.edu)

[2] Department of Mathematics,
Rutgers University, New Brunswick, NJ 08903, USA
(e-mail: komlos@math.rutgers.edu)

Received 20 August 1999; revised 25 October 2000

Fix q and let M_n be an $n \times n$ matrix with entries drawn independently from the finite field $\mathbf{F}_q$ according to some distribution μ_n. It is shown that, except in certain pathological cases, the probability that M_n is nonsingular is asymptotically the same as for uniform entries; that is,

$$\Pr(M_n \text{ is nonsingular}) \to \prod_{i \geqslant 1}(1 - q^{-i}) \quad \text{as } n \to \infty.$$

1. Introduction

Let $q = p^e$ be a prime power and write $\mathbf{F}$ for the Galois field $\mathbf{F} = \mathbf{F}_q = GF(q)$. We suppose throughout that $\mathbf{F}$ is fixed and $M = M_n$ is a random $n \times n$ matrix with entries chosen independently from $\mathbf{F}$.

If the entries are chosen *uniformly* from $\mathbf{F}$, then it is well known (and easy to see) that

$$\Pr(M_n \text{ is nonsingular}) \to \prod_{i \geqslant 1}(1 - q^{-i}) \quad \text{as } n \to \infty. \tag{1.1}$$

A nice surprise, discovered by various authors, is that (1.1) is quite robust. Specifically, Charlap, Rees and Robbins ([3], see also [5]) proved the following theorem in 1990.

Theorem 1.1. *A sufficient condition for* (1.1) *is that the entries of* M_n *be drawn from*

† Research supported in part by NSF grants DMS-9622966 and DMS-9970433.
‡ Research supported in part by NSF grant DMS-9801396.

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

some fixed probability distribution μ which is not concentrated on any proper affine subspace of $\mathbf{F}$.

They also discuss the extent to which the restriction on the support of μ in Theorem 1.1 can be relaxed. Note that some restriction is required since we clearly cannot expect (1.1) if μ is concentrated on a single point or a proper subfield. On the other hand, as noted in [3], the condition of Theorem 1.1 seems stronger than necessary; for example (again as noted in [3]), uniform distribution on some 2-dimensional subspace of $\mathbf{F} = \mathbf{F}_8$ regarded as a vector space over $\mathbf{F}_2$ is a first case not covered by Theorem 1.1.

A main result of the present paper (Theorem 1.3) is that replacing 'proper affine *subspace*' with 'proper affine *subfield*' in Theorem 1.1 gives a necessary and sufficient condition in the case of a fixed probability distribution. (So, for instance, (1.1) does hold for the example at the end of the preceding paragraph.)

Another extension of the uniform limit is to random matrices with μ depending on n. One limitation here is that, if $\|\mu\|_\infty > 1 - (\log n + O(1))/n$, then there is an $\Omega(1)$ chance of having either a row or column of zeros (if $\|\mu\|_\infty = \mu(0)$) or two or more equal rows or columns, so it is no longer reasonable to expect (1.1). (In fact, if $\|\mu\|_\infty > 1-(\log n-\omega(1))/n$, then M_n is a.s. singular.)

This question was considered by Kovalenko, Levitskaya and Savchuk [6]. They prove the standard limit (1.1) under the condition that the entries m_{ij} of M are independent and $\Pr(m_{ij} = \alpha) > (\log n + \omega(1))/n$ for all $\alpha \in \mathbf{F}$. (They prove the conclusion even when the entries are not identically distributed, but then the above lower bound on $\Pr(m_{ij} = \alpha)$ should be uniform in i, j.)

The behaviour of the nullity of M_n for $1-\mu(0) =: p$ close to $\log n/n$ and $\mu(\alpha) = p/(q-1)$ for $\alpha \neq 0$ was considered by Blömer, Karp and Welzl [2] in 1997; this was the original motivation for the present work. When $q = 2$ very precise information for the model of [2] is given by Cooper in [4].

In the present paper we address both these issues, giving, in Theorem 1.2, sufficient (and probably necessary; see below) conditions for (1.1) when the entries of $M = M_n$ are drawn from a distribution $\mu = \mu_n$ which may depend on n. The special case of fixed μ is completely settled in Theorem 1.3.

We will say that $\mathbf{H}$ is an *affine subfield* of $\mathbf{F}$ if $\mathbf{H} = \alpha\mathbf{K} + \beta$ with $\mathbf{K}$ a subfield of $\mathbf{F}$, $\alpha, \beta \in \mathbf{F}$ and $\alpha \neq 0$, and that $\mathbf{H}$ is *proper* if it is not equal to $\mathbf{F}$.

We will say that a matrix or vector is chosen *according to* μ if its entries are chosen independently, each according to μ.

Theorem 1.2. *Let μ_n be a probability distribution on $\mathbf{F} = \mathbf{F}_q$ (q fixed), and let M_n be a random $n \times n$ $\mathbf{F}$-matrix chosen according to μ_n. Then* (1.1) *holds provided*

$$\|\mu_n\|_\infty < 1 - \frac{\log n + \omega(1)}{n} \tag{1.2}$$

and

$$\mu_n(\mathbf{F} \setminus \mathbf{H}) > \omega(n^{-2}) \quad \textit{for each proper affine subfield } \mathbf{H} \textit{ of } \mathbf{F}. \tag{1.3}$$

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

(Throughout the paper log denotes natural logarithm.)

For fixed μ we present the definitive result mentioned earlier.

Theorem 1.3. *Let M_n be a random $n \times n$ $\mathbf{F}$-matrix with entries chosen according to some fixed non-degenerate probability distribution μ on $\mathbf{F}$. Then* (1.1) *holds if and only if the support of μ is not contained in any proper affine subfield of $\mathbf{F}$.*

Sufficiency of the conditions on μ – the more interesting part of this – is given by Theorem 1.2. A proof of necessity is provided in Section 7. (Here Conjecture 7.2 seems particularly interesting, both in its own right – since it appears to require some new ideas – and because it would complete a determination of $\lim_{n\to\infty} \Pr(M_n \text{ is nonsingular})$ for an *arbitrary* fixed μ.)

The main idea in the proof of Theorem 1.2 is that there is very little chance of 'non-generic' dependencies among the columns of M_n. Once these have been ruled out, probabilities for more typical dependencies behave essentially as for uniform entries. This vague description is implemented in Section 3, where the proof of Theorem 1.2 is given modulo results (Lemmas 3.1 and 3.2) proved in later sections.

The most complicated part of the whole business is the proof of Lemma 3.2 given in Section 6. This is probably not optimal, though we were unable to find a satisfyingly simple way to do it. As mentioned in the remarks following (6.7), most of this work can be avoided if we are willing to strengthen (1.3) to '$\mu_n(\mathbf{F} \setminus \mathbf{H}) > \omega(1/n)$', so in particular if we just want Theorem 1.3.

As suggested above, we believe (1.2) and (1.3) are also *necessary* for (1.1); in fact the probability of nonsingularity should be less than the right side of (1.1) if either of these conditions fails.

Both restrictions are at least natural. The reason for (1.2) (avoidance of row or column dependencies of length 1 or 2) was noted earlier. Similarly, failure of (1.3) gives an $\Omega(1)$ probability that M_n has no entries outside some proper affine subfield $\mathbf{H}$, which should increase the probability of singularity. (This is less obvious when $\mathbf{H} \not\ni 0$, but still seems to be true.)

A proof of necessity should identify anomalous situations arising under failure of (1.2) and/or (1.3), and show that the probability of nonsingularity is less than the right side of (1.1) in such situations, and asymptotically at most the right side of (1.1) otherwise. (One cautionary example: if $\mathbf{H}$ is a proper subfield and we condition on M_n having exactly one entry in $\mathbf{F} \setminus \mathbf{H}$, then the limit of $\Pr(M_n \text{ is nonsingular})$ will generally be strictly less than the right side of (1.1).) We believe such a program could eventually be carried out mainly via the present methods, but this seems likely to involve considerable extra effort.

2. Notation and background

To repeat, we take $\mathbf{F} = \mathbf{F}_q$, where the prime p and the prime-power $q = p^e$ are fixed. When working in the m-dimensional vector space $\mathbf{F}^m$, we take $u \cdot v$ to be the usual inner product, and, as usual, set $U^\perp = \{v \in \mathbf{F}^m : u \cdot v = 0 \text{ for all } u \in U\}$ for any $U \subseteq \mathbf{F}^m$.

We will sometimes regard $\mathbf{F}$ itself as a vector space over one of its subfields $\mathbf{H}$; subspaces

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

of this vector space are then called **H**-*subspaces*. (In other words, an **H**-subspace of **F** is an additive subgroup of **F** closed under multiplication by **H**.) Of course, any vector space over **F** is also a vector space over any subfield of **F**, but here we will use the natural default: unless otherwise specified, 'subspace of $\mathbf{F}^m$' refers to $\mathbf{F}^m$ as a vector space over **F**.

Vectors are column vectors by default. For a real matrix M we write M^T for the transpose. Hence r^T is a row vector if r is a column vector. The *support* of a vector v (or that of a discrete measure) is the set $\mathrm{supp}(v) = \{i : v_i \neq 0\}$. For a vector v and a set $S \subseteq \mathbf{F}$ of scalars, we write $Sv = \{\alpha v : \alpha \in S\}$. In particular, $\mathbf{F}\underline{1}$ is the set of column vectors $\{(\alpha, \alpha, \ldots, \alpha)^T : \alpha \in \mathbf{F}\}$. For $v \in \mathbf{F}^m$ and $I \subseteq [m]$, v_I is the restriction of v to I: $v_I = (v_i : i \in I)$.

If M is a $k \times m$ **F**-matrix (matrix with entries in **F**), then $\mathrm{row}(M)$ is the row space of M (linear span of the rows of M).

We use $[n]$ for the set $\{1, 2, \ldots, n\}$. Given positive sequences $\{f(n)\}$ and $\{g(n)\}$, we use the standard notations $f(n) = O(g(n))$, $f(n) = o(g(n))$, $f(n) = \Omega(g(n))$, and $f(n) = \omega(g(n))$, to denote, as $n \to \infty$, the limit relations $\limsup f(n)/g(n) < \infty$, $\lim f(n)/g(n) = 0$, $\liminf f(n)/g(n) > 0$, and $\lim f(n)/g(n) = \infty$, respectively. We will also use $f(n) = o(g(n))$ (with the same meaning) in cases where f is not required to be positive.

In general we pretend large numbers are integers, nothing we do being affected by such rounding.

Throughout the paper – as in the above theorems – we work with a 'hidden parameter' n, so that a statement like '$f(u) = o(1)$' really refers to a sequence $\{u_n\}$ and asserts that $f(u_n) \to 0$ as $n \to \infty$. We assume throughout that n is large enough to support our assertions, but emphasize again that the underlying field **F**, and thus also the parameters p, e and q, are fixed.

As usual, we say that an event $Q = Q_n$ holds *almost surely* (a.s.) if its probability tends to 1 as $n \to \infty$. For μ a probability distribution on X and $K \subseteq X$ we write μ_K for the conditional distribution $\mu(\cdot|K)$.

3. Main ingredients and Proof of Theorem 1.2

Here we give just enough of the results of later sections to allow us to prove Theorem 1.2. We assume throughout that $\mu = \mu_n$ satisfies (1.2) and a concrete version of (1.3):

$$\mu_n(\mathbf{F} \setminus \mathbf{H}) > h(n)/n^2 \quad \text{for each proper affine subfield } \mathbf{H} \text{ of } \mathbf{F}, \tag{3.1}$$

with $h(n) = \omega(1)$.

Eventually our main task (Lemma 3.2) will be to show that the first $n - \omega(1)$ rows of M_n are a.s. 'generic', meaning that they are (i) linearly independent, and (ii) not simultaneously annihilated by any 'atypical' vectors, formally those not belonging to

$$\mathscr{C} = \mathscr{C}(n, s_0) := \{b \in \mathbf{F}^n : |b^{-1}(\gamma)| \geqslant s_0 \quad \text{for all } \gamma \in \mathbf{F}\},$$

where

$$s_0 = (1 - \log^{-1} n)n/q \tag{3.2}$$

(and, of course, $b^{-1}(\gamma) = \{i : b_i = \gamma\}$).

The next two lemmas are proved in Section 4 (see Lemma 4.4) and Section 6, respectively.

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

Lemma 3.1. *Let U be an i-dimensional subspace of $\mathbf{F}^n$ contained in $\mathscr{C} \cup \{\underline{0}\}$ and let r be chosen from $\mathbf{F}^n$ according to μ. Then $\Pr(r \in U^\perp) = q^{-i} + o(1/n)$ uniformly in i and U.*

For Lemma 3.2 and the following discussion, let $h'(n)$ be any function with $\omega(1) < h'(n) < o(n \log^{-2} n)$, and set $h''(n) = \min\{h(n)/(3e), h'(n)\}$ (where $q = p^e$).

Lemma 3.2. *Let N be the matrix consisting of the first ℓ rows of M.*

(a) *If $\ell < n - \omega(1)$, then the rows of N are linearly independent a.s.*
(b) *If $\ell \geqslant n - h''(n)$, then a.s.*

$$Nb = \underline{0} \Rightarrow b \in \mathscr{C} \cup \{\underline{0}\}.$$

Given these we finish easily, as follows. Let $\ell = n - h''(n)$. Let $M^{(i)}$ be the matrix consisting of the first i rows of M and define the events

$$R = \{M^{(\ell)} b \neq \underline{0} \text{ for all } b \in \mathbf{F}^n \setminus (\mathscr{C} \cup \{\underline{0}\})\},$$

$$Q_i = R \wedge \{\operatorname{rank}(M^{(i)}) = i\}.$$

Then $\Pr(Q_\ell) = 1 - o(1)$ (by Lemma 3.2), and

$$\Pr(M \text{ is nonsingular}) = \Pr(Q_n) = \Pr(Q_\ell) \Pr(Q_{\ell+1} | Q_\ell) \cdots \Pr(Q_n | Q_{n-1}).$$

So it is enough to show

$$\Pr(Q_{i+1} | Q_i) = (1 - o(1/n))(1 - q^{-(n-i)}),$$

or equivalently, with r^i the ith row of M,

$$\Pr(r^{i+1} \notin \operatorname{row}(M^{(i)}) | Q_i) = (1 - o(1/n))(1 - q^{-(n-i)}), \quad \ell \leqslant i \leqslant n - 1. \tag{3.3}$$

But on Q_i, $\operatorname{row}(M^{(i)})^\perp$ is an $(n-i)$-dimensional subspace of $\mathbf{F}^n$ contained in $\mathscr{C} \cup \{\underline{0}\}$; so (3.3) is an instance of Lemma 3.1. □

4. Fourier analysis

Background

Let $\zeta = e^{2\pi i/p}$. Recall that the trace function $Tr(\alpha) = \alpha + \alpha^p + \cdots + \alpha^{p^{e-1}}$ is a linear map from $GF(p^e)$ onto $GF(p)$ (that is, from $\mathbf{F}$ onto $\mathbf{Z}_p$), and hence, for any fixed $u \in \mathbf{F}^m$, the function $\chi_u(v) = \zeta^{Tr(u \cdot v)}$, $v \in \mathbf{F}^m$, is a character of the additive group $\mathbf{F}^m$.

For $f : \mathbf{F}^m \to \mathbf{C}$, the Fourier transform of f is given by

$$\hat{f}(u) = \sum_{v \in \mathbf{F}^m} \zeta^{Tr(u \cdot v)} f(v).$$

The following identity is easy to verify (see also Lemma 2 in [3]): for any ($\mathbf{F}$-)subspace V of $\mathbf{F}^m$,

$$\sum_{v \in V} f(v) = \frac{1}{|U|} \sum_{u \in U} \hat{f}(u) \quad \text{where} \quad U = V^\perp.$$

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

(This is the Poisson Summation Formula for the special case when the subgroup V is actually a subspace, in which case $Tr(u \cdot v) = 0$ for all $v \in V$ if and only if $u \in V^{\perp}$.)

Replacing $f(v)$ with $f(z+v)$ and V with $U^{\perp}$, we get the form we will need.

Lemma 4.1. *For any subspace U of $\mathbf{F}^m$ and any $z \in \mathbf{F}^m$,*

$$\sum_{v \in U^{\perp}+z} f(v) = \frac{1}{|U|} \sum_{u \in U} \zeta^{-Tr(z \cdot u)} \hat{f}(u).$$

Note that this contains the inversion formula

$$f(z) = q^{-m} \sum_{u \in \mathbf{F}^m} \zeta^{-Tr(z \cdot u)} \hat{f}(u).$$

Convolutions

The following easy (and undoubtedly known) fact will be needed here and again later on. The symbol $*$ here is used for convolution of probability measures.

Proposition 4.2. *Suppose $\sigma_1, \sigma_2, \ldots$ are probability distributions on an additive group $\mathbf{F}$, with $\|\sigma_i\|_{\infty} \leqslant 1 - \eta$ for each i, where $\eta \leqslant 1/2$. Then, for any k,*

$$\begin{aligned} \|\sigma_1 * \cdots * \sigma_k\|_{\infty} &\leqslant \frac{1 + (1-2\eta)^k}{2} \\ &\leqslant \begin{cases} 1 - k\eta + k^2\eta^2, & \text{for all } \eta, \\ 3/4, & \text{if } k\eta > 1/2. \end{cases} \end{aligned} \tag{4.1}$$

Note that the first inequality is exact when $1 + 1 = 0$ in $\mathbf{F}$ and $\sigma_i(0) = 1 - \eta = 1 - \sigma_i(1)$.

Proof. We proceed by induction:

$$\begin{aligned} \sigma_1 * \cdots * \sigma_k(\gamma) &= \sum_{\alpha} \sigma_1 * \cdots * \sigma_{k-1}(\alpha)\sigma_k(\gamma - \alpha) \\ &\leqslant \frac{1 + (1-2\eta)^{k-1}}{2}(1-\eta) + \frac{1 - (1-2\eta)^{k-1}}{2}\eta = \frac{1 + (1-2\eta)^k}{2}. \end{aligned}$$

For this last inequality we used the easy fact that, if the numbers $a_i, b_i \geqslant 0$ satisfy $\sum a_i = \sum b_i = 1$ and $a_i \leqslant 1 - \delta$, $b_i \leqslant 1 - \eta$ for all i, where $\eta, \delta \leqslant 1/2$, then $\sum a_i b_i \leqslant (1-\delta)(1-\eta) + \delta\eta$.

The bounds (4.1) are standard calculus. □

Main inequalities

We now turn to the main business of this section. As earlier, set

$$\mathscr{C}(m,s) = \mathscr{C}_{\mathbf{F}}(m,s) = \{b \in \mathbf{F}^m : |b^{-1}(\gamma)| \geqslant s \text{ for all } \gamma \in \mathbf{F}\}.$$

We want to say (roughly) that, for suitable s, inner products of vectors from $\mathscr{C}(m,s)$ with $r \in \mathbf{F}^m$ chosen according to an appropriate σ are nicely behaved (see the Main Lemma, Lemma 4.4). We will get at this via the Fourier transform on $\mathbf{F}^m$. Let f be a probability

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

distribution on $\mathbf{F}^m$ and let U be a subspace of $\mathbf{F}^m$ of dimension d. Applying Lemma 4.1 (and using $\hat{f}(0) = \sum f(v) = 1$) we get, for r chosen from $\mathbf{F}^m$ according to f,

$$\Pr(r \in U^\perp + z) = q^{-d}\left[1 + \sum_{u \in U \setminus \{\underline{0}\}} \zeta^{-Tr(z \cdot u)} \hat{f}(u)\right].$$

This proves the following lemma.

Lemma 4.3. *Let r be a random m-vector chosen according to the distribution f on $\mathbf{F}^m$. Then, for any $z \in \mathbf{F}^m$ and any subspace U of $\mathbf{F}^m$ of dimension d,*

$$|\Pr(r \in U^\perp + z) - q^{-d}| \leqslant q^{-d} \sum_{u \in U \setminus \{\underline{0}\}} |\hat{f}(u)|. \tag{4.2}$$

When f is a product distribution σ^m on $\mathbf{F}^m$ (*i.e.*, $f(v) = \prod_{i=1}^m \sigma(v_i)$, with σ some probability distribution on $\mathbf{F}$), we have

$$\hat{f}(u) = \prod \hat{\sigma}(u_i), \tag{4.3}$$

where $\hat{\sigma}(\gamma) = \sum_{\alpha \in \mathbf{F}} \zeta^{Tr(\alpha\gamma)}\sigma(\alpha)$.

Assume from now on that $\|\sigma\|_\infty \leqslant 1 - \eta$. We first observe that

$$\prod_{\alpha \in \mathbf{F}} \hat{\sigma}(\alpha) \leqslant 1 - q\eta + q^2\eta^2. \tag{4.4}$$

To see this, consider some $u \in \mathbf{F}^q$ having one entry equal to γ for each $\gamma \in \mathbf{F}$, and let $r \in \mathbf{F}^q$ be chosen according to $f = \sigma^q$. Applying Lemma 4.1 (with $z = \underline{0}$ and $U = \langle u \rangle$, the one-dimensional subspace generated by u) and (4.3) yields

$$\Pr(r^T u = 0) = q^{-1} \sum_{\alpha \in \mathbf{F}} \hat{f}(u\alpha) = q^{-1}\left[1 + \sum_{\alpha \in \mathbf{F}^\times} \prod_{i=1}^{q} \hat{\sigma}(u_i\alpha)\right] = q^{-1}\left[1 + (q-1)\prod_{\alpha \in \mathbf{F}} \hat{\sigma}(\alpha)\right]. \tag{4.5}$$

On the other hand we have, from Proposition 4.2,

$$\Pr(r^T u = 0) \leqslant \frac{1 + (1-2\eta)^{q-1}}{2} \leqslant 1 - (q-1)\eta + (q-1)^2\eta^2,$$

and this with (4.5) implies (4.4). □

This observation easily gives our basic inequality.

Lemma 4.4 (Main Lemma). *Let N be a $d \times n$ $\mathbf{F}$-matrix of rank d with $\mathrm{row}(N) \subseteq \mathscr{C}(n,s) \cup \{\underline{0}\}$, and let σ be a probability distribution with $\|\sigma\|_\infty \leqslant 1 - \eta$. Then, for $r \in \mathbf{F}^n$ chosen according to $f = \sigma^n$ and any $c \in \mathbf{F}^d$,*

$$|\Pr(Nr = c) - q^{-d}| \leqslant (1 - q\eta + q^2\eta^2)^s.$$

Note that this lemma includes Lemma 3.1: choose $d = i$, $\sigma = \mu$, $c = \underline{0}$, $\eta = (\log n + \omega(1))/n$, and $s = s_0$. The error term $(1 - q\eta + q^2\eta^2)^s$ is then $o(1/n)$ and is indeed independent of i and U.

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

Proof. Fix $z \in \mathbf{F}^n$ with $Nz = c$ and set $\mathrm{row}(N) = U$. Then $Nr = c$ is the same as $r \in U^\perp + z$. By Lemma 4.3,

$$|\Pr(Nr = c) - q^{-d}| \leqslant q^{-d} \sum_{u \in U \setminus \{0\}} |\hat{f}(u)|. \tag{4.6}$$

But, since $|\hat{\sigma}(\gamma)| \leqslant 1$ for all γ, (4.3) and (4.4) imply $|\hat{f}(u)| \leqslant (1 - q\eta + q^2\eta^2)^s$ whenever $u \in \mathscr{C}(n, s)$ (in particular for each u appearing on the right-hand side of (4.6)) and the lemma follows. □

This will be our main tool for dealing with possibilities for annihilation of M_n by fairly 'typical' vectors – those in which each $\gamma \in \mathbf{F}$ appears many times. The next section develops methods for dealing with more pathological vectors.

5. Elementary bounds

For a fixed finite set X and $\sigma = \sigma_n$ a probability distribution on X, we say that σ is *almost uniform* (on X), and that an element of X chosen according to σ is *drawn almost uniformly* from X, if

$$\max_{\gamma \in X} |\sigma(\gamma) - |X|^{-1}| = o(1)$$

(equivalently, the total variation distance between σ and the uniform distribution is $o(1)$).

Lemma 5.1. *Let $\sigma = \sigma_n$ be a probability distribution with support $K \subseteq \mathbf{F}$, assumed to contain 0, and let A be the additive subgroup of $\mathbf{F}$ generated by K. Let $\gamma_1, \ldots, \gamma_m$ be drawn independently from K, each according to σ, and set $\gamma^* = \sum \gamma_i$. Then γ^* is drawn almost uniformly from A provided $m \cdot \min\{\sigma(\gamma) : \gamma \in K\} \to \infty$.*

Proof. We may regard γ^* as the mth state of a Markov chain on A with 0th state 0 and transition probabilities $p(\gamma, \delta) = \sigma(\delta - \gamma)$. The eigenvalues of the transition matrix are $\lambda_\chi = \sum\{\sigma(\gamma)\chi(\gamma) : \gamma \in K\}$, with χ ranging over the characters of A. Now γ^* is almost uniform from A provided $|\lambda_\chi| < 1 - \omega(1/m)$ for each $\chi \neq 1$; but this is an easy consequence of the hypotheses of the proposition. (An elementary proof would be equally easy.) □

Remark. If we omit the assumption $0 \in K$ then we should take A to be generated by $K - \gamma_0$ for some (any) $\gamma_0 \in K$, and conclude that γ^* is almost uniform from $A + m\gamma_0$.

For a set $A \subseteq \mathbf{F}$ and a vector $B = (\beta_1, \ldots, \beta_k) \in \mathbf{F}^k$, let

$$A \circ B = \left\{ \sum_{i=1}^{k} \alpha_i \beta_i : \alpha_1, \ldots, \alpha_k \in A \right\}.$$

For a set $S \subseteq \mathbf{F}$ with k elements we define $A \circ S = A \circ B$, where $B = (\beta_1, \ldots, \beta_k)$ is a list of the elements of S.

Lemma 5.2. *Fix $B = (\beta_1, \ldots, \beta_k) \in \mathbf{F}^k$ and let $\alpha_1, \ldots, \alpha_k$ be chosen uniformly and independently from an additive subgroup A of $\mathbf{F}$. Then $\sum \alpha_i \beta_i$ is chosen uniformly from $A \circ B$.*

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

Proof. For $a = (\alpha_1, \ldots, \alpha_k) \in A^k$, set $a \circ B = \sum \alpha_i \beta_i$. That A is a subgroup implies that, for any fixed $a' \in A^k$ and a chosen uniformly from A^k, $(a + a') \circ B$ has the same distribution as $a \circ B$. So, in particular, for any $y \in A \circ B$, we may fix $a' \in A^k$ with $a' \circ B = y$, and then

$$\Pr(a \circ B = y) = \Pr((a + a') \circ B = y) = \Pr(a \circ B = 0). \qquad \square$$

Corollary 5.3. *Assume the conditions of Lemma 5.2, but with $\alpha_1, \ldots, \alpha_k$ only assumed to be* almost *uniform from A. Then $\sum \alpha_i \beta_i$ is chosen almost uniformly from $A \circ B$.* $\square$

Lemma 5.4. *Let $B = (\beta_1, \ldots, \beta_k) \in \mathbf{F}^k$, $K \subseteq \mathbf{F}$ with $0 \in K$, σ some probability distribution on K, and A the additive subgroup of $\mathbf{F}$ generated by K. Let $b \in \mathbf{F}^m$ and $I_1, \ldots, I_k$ disjoint subsets of $[m]$ with $b_i = \beta_j$ for all $i \in I_j$, and assume*

$$\min_j |I_j| \min_{\gamma \in K} \sigma(\gamma) = \omega(1).$$

Then, for $r \in K^m$ chosen according to σ^m, and any $\gamma \in \mathbf{F}$,

$$\Pr(r^T b = \gamma) < [|A \circ B| - o(1)]^{-1}.$$

Proof. Let us condition on any specification of $(r_i : i \notin \cup I_j)$, say with $\sum_{i \notin \cup I_j} r_i \beta_i = \gamma'$, and let $r_j^* = \sum_{i \in I_j} r_i$. By Lemma 5.1, r_j^* is chosen almost uniformly from A; so by Corollary 5.3, $r^T b = \gamma' + \sum_{j=1}^k r_j^* \beta_j$ is almost uniform from $\gamma' + A \circ B$. The lemma follows. $\square$

Lemma 5.5. *For any $A \subseteq \mathbf{F}$ with $|A| \geqslant 2$ and $\emptyset \neq B \subseteq \mathbf{F}$, $|A \circ B| \geqslant |B|$. Equality holds if and only if there exist a subfield $\mathbf{H}$ of $\mathbf{F}$ and $\alpha, \beta \in \mathbf{F}$, $\alpha \neq 0$, such that A is a subset of the affine subfield $\mathbf{H}' = \alpha \mathbf{H} + \beta$, and B is an $\mathbf{H}$-subspace of $\mathbf{F}$.*

Proof. Since the lemma is not affected by applying a (non-degenerate) affine transformation to A, we may assume $0, 1 \in A$. In this case we have $B \subseteq A \circ B$, which gives the inequality.

Now assume equality holds, so that $A \circ B = B$. If $|B| = 1$ then $|A \circ B| = |B| < |A|$ implies $B = \{0\}$. If $|B| \geqslant 2$, then $B + B \subseteq B$, so that B (being finite) is an additive subgroup of $\mathbf{F}$. Hence the set $\mathbf{H} := \{\gamma \in \mathbf{F} : \gamma B \subseteq B\}$ is a subfield of $\mathbf{F}$, and B is an $\mathbf{H}$-subspace of $\mathbf{F}$. The relation $AB \subseteq B$ shows that $A \subseteq \mathbf{H}$. $\square$

Finally, we will need the following form of the Chernoff bound (see, *e.g.*, [1, p. 238, Theorem A.13]).

Proposition 5.6. *If the random variable X has a binomial distribution with mean μ then for any $x \leqslant \mu$ we have* $\Pr(X \leqslant x) \leqslant \exp\{-(\mu - x)^2/(2\mu)\}$.

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

6. Proof of Lemma 3.2

We need a concrete version of (1.2):

$$\|\mu_n\|_\infty \leqslant 1 - \frac{\log n + g(n)}{n} =: 1 - \varepsilon, \tag{6.1}$$

with $g(n) = \omega(1)$. For the rest of the paper we write μ for μ_n.

We will generalize slightly and assume throughout that L is an $\ell \times m$ $\mathbf{F}$-matrix chosen according to μ with

$$(1 - o(\log^{-2} n))n < \ell, m \leqslant n. \tag{6.2}$$

(The assumption $l \leqslant n$ is unnecessary – Lemmas 6.1 and 6.2 trivially become 'more true' as l grows – but will spare us some distractions.)

Set $\mathscr{C} = \mathscr{C}_{\mathbf{F}}(m, s_0)$ with s_0 as in (3.2). All bs in this section are assumed to lie in $\mathbf{F}^m \setminus \{\underline{0}\}$.

Lemma 6.1. *If $\ell > m + \omega(1)$ then a.s. there is no $b \in \mathscr{C}$ with $Lb \in \mathbf{F}\underline{1}$.*

Lemma 6.2. *If* (*in addition to* (6.2))

$$m \leqslant \ell + h(n)/(3e), \tag{6.3}$$

where as usual $q = p^e$, then

$$\text{a.s. there is no } b \in \mathbf{F}^m \setminus (\mathscr{C} \cup \{\underline{0}\}) \text{ with } Lb \in \mathbf{F}\underline{1}. \tag{6.4}$$

(In fact we only need Lemma 6.2 with '$Lb = \underline{0}$', but the version given is convenient because it will allow us to assume that the affine subfield $\mathbf{H}$ in the proof is actually a subfield. Again in Lemma 6.1 '$Lb = \underline{0}$' would suffice for Theorem 1.2, but we will make some use of the stronger statement in Section 7.)

Combined (and transposed) these give Lemma 3.2(a), while Lemma 6.2 itself contains Lemma 3.2(b).

Proof of Lemma 6.1. For $b \in \mathscr{C}$, $\gamma \in \mathbf{F}$ and $r \in \mathbf{F}^m$ chosen according to μ, Lemma 4.4 gives

$$\Pr(r^T b = \gamma) \leqslant 1/q + (1 - q\varepsilon + q^2\varepsilon^2)^{s_0} = 1/q + o(1/n).$$

So the probability that $Lb \in \mathbf{F}\underline{1}$ for some $b \in \mathscr{C}$ is at most $q^{m+1}[1/q + o(1/n)]^\ell = o(1)$. □

Proof of Lemma 6.2. Let $f(n)$ be some function satisfying

$$\min\{\log\log n/(3q), g(n)/(2q)\} > f(n) = \omega(1) \tag{6.5}$$

($g(n)$ as in (6.1)).

Let $K = \{\alpha \in \mathbf{F} : \mu(\alpha) > f(n)/n\}$, and let $\mathbf{H}$ be the (unique) minimal affine subfield of $\mathbf{F}$ containing K. Since the truth of the lemma is not affected when we replace μ by μ' given by $\mu'(\gamma) = \mu(\alpha\gamma + \beta)$ ($\alpha, \beta \in \mathbf{F}$, $\alpha \neq 0$), we may assume $0, 1 \in K$, so that $\mathbf{H}$ is actually a subfield. (Note $|K| \geqslant 2$ by (6.1) – or (1.2) – and (6.5).) Let A be the additive subgroup of $\mathbf{F}$ generated by K.

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

Set $T = nf^{-1/2}(n)$ (we just need $T/n \to 0$ and $Tf(n)/n \to \infty$), and for $b \in \mathbf{F}^m \setminus \{\underline{0}\}$ write

$$\mathbf{F}'(b) = \{\gamma \in \mathbf{F} : |b^{-1}(\gamma)| \geqslant T\}.$$

For the proof of Lemma 6.2 we consider an (unfortunately large) number of possibilities for b. Set

$$\begin{aligned}
\mathscr{B}_0 &= \mathscr{C} \cup \{\underline{0}\},\\
\mathscr{B}_1 &= \mathscr{C}(m,s) \setminus \mathscr{C}, \quad \text{with } s = m/(3q),\\
\mathscr{B}_2 &= \{b : |\mathrm{supp}(b)| < qT\},\\
\mathscr{B}_3 &= \{b : \mathbf{F}'(b) \text{ is not an } \mathbf{H}\text{-subspace of } \mathbf{F}\} \setminus \cup_{i<3}\mathscr{B}_i,\\
\mathscr{B}_4 &= \{b : |b^{-1}(\gamma)| \leqslant m/(2|\mathbf{F}'(b)|) \text{ for some } \gamma \in \mathbf{F}'(b)\} \setminus \cup_{i<4}\mathscr{B}_i,\\
\mathscr{B}_5 &= \mathbf{F}^m \setminus (\cup_{i<5}\mathscr{B}_i).
\end{aligned}$$

Set $Q_i = \{Lb \in \mathbf{F}\underline{1} \text{ for some } b \in \mathscr{B}_i\}$. (We will follow this convention throughout with superscripts and other decorations: *any decorated Q is the event that $Lb \in \mathbf{F}\underline{1}$ for some b in the correspondingly decorated $\mathscr{B}$.*)

To prove Lemma 6.2 we need to show that $\Pr(Q_i) = o(1)$ for $1 \leqslant i \leqslant 5$, assuming (6.2) and (6.3); it will be convenient to show a bit more:

$$\text{if } \ell, m \text{ satisfy (6.2), then } \Pr(Q_i) = o(1) \quad \text{for} \quad 1 \leqslant i \leqslant 4; \tag{6.6}$$

$$\text{if } \ell, m \text{ satisfy (6.2) and (6.3), then } \Pr(Q_5) = o(1). \tag{6.7}$$

Remarks. If $\mathbf{H} = \mathbf{F}$ then $\mathscr{B}_5 = \emptyset$, since $\{b : \mathbf{F}'(b) \neq \mathbf{F}\} \subseteq \mathscr{B}_3 \cup \mathscr{B}_2$, and $\{b : \mathbf{F}'(b) = \mathbf{F}\} \subseteq \mathscr{B}_4 \cup \mathscr{C}(m,s)$. This greatly simplifies the following proof, most of which is devoted to (6.7). In particular, if $h(n) = \omega(n)$ – that is, if the lower bound in (3.1) is $\omega(1/n)$ – then we can be a little more careful with our choice of f, adding '$f(n) < h(n)/(qn)$' to the requirements (6.5), and then necessarily $\mathbf{H} = \mathbf{F}$. So, for instance, we get a reasonably simple proof of Theorem 1.3.

We should also point out here that something close to (6.3) really is needed for (6.7) (and for Lemma 6.2). For suppose $m - \ell > h(n)$ and $\mathbf{H}'$ is a subfield of $\mathbf{F}$ with $(h(n)/n^2 <) \mu(\mathbf{F} \setminus \mathbf{H}') \leqslant (m-l)/n^2$. Then with probability $\Omega(1)$ there are fewer than $m - l$ non-$\mathbf{H}'$ entries in L, implying that there is an $\ell \times m'$ submatrix L' of L with $m' > \ell$ and all entries of L' in $\mathbf{H}'$. But in this case $L'b' = \underline{0}$ for some (nonzero) $b' \in (\mathbf{H}')^{m'}$, and then $Lb = \underline{0}$ when b is b' suitably augmented by 0s (and clearly $b \notin \mathscr{C}$).

For the proofs of (6.6) and (6.7), we consider the $\mathscr{B}_i$s in order.

$\mathscr{B}_1$. We again apply Lemma 4.4, in this case exploiting the fact that the bs under consideration are rare. For $t = \alpha m/q \in [s, s_0)$, set $\mathscr{B}_1^t = \{b : \min_{\gamma \in \mathbf{F}} |b^{-1}(\gamma)| = t\}$. By Proposition 5.6,

$$|\mathscr{B}_1^t| \leqslant q^{m+1} \exp\left\{-\frac{(m/q - t)^2}{2m/q}\right\} = q^{m+1} \exp\left\{-\frac{(1-\alpha)^2 m}{2q}\right\}$$

(the extra q-factors are for the choice of γ), while for $b \in \mathscr{B}_1^t$, $\gamma \in \mathbf{F}$, and $r \in \mathbf{F}^m$ chosen

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

according to μ, Lemma 4.4 gives

$$\Pr(r^T b = \gamma) \leqslant q^{-1}[1 + q(1 - q\varepsilon + q^2\varepsilon^2)^t].$$

Thus, since

$$q(1 - q\varepsilon + q^2\varepsilon^2)^t < q\exp\{-\alpha(\log n + \omega(1))\} = o(n^{-\alpha})$$

(using $m\varepsilon = \log n + \omega(1)$, a consequence of (6.2)), we have

$$\begin{aligned}\Pr(Q_1^t) \leqslant |\mathscr{B}_1^t|\,(\max \Pr(r^T b = \gamma))^\ell &< q^{m+1-\ell}\exp\{-(1-\alpha)^2 m/(2q) + o(\ell n^{-\alpha})\}\\ &< q^{m+1-\ell}\exp\{o(n^{1-\alpha}) - \Omega((1-\alpha)^2 n)\},\end{aligned}$$

where the max is over $b \in \mathscr{B}_1^t$ and $\gamma \in \mathbf{F}$. Since $\alpha < 1 - \log^{-1} n$, the term in brackets is $-\Omega((1-\alpha)^2 n) = -\Omega(n\log^{-2} n)$, and this with (6.2) gives

$$\Pr(Q_1) \leqslant \sum_{t=s}^{s_0} \Pr(Q_1^t) = o(1).$$

$\mathscr{B}_2$. For $b \in \mathbf{F}^m$ with $|\mathrm{supp}(b)| = x$, $\gamma \in \mathbf{F}$ and $r \in \mathbf{F}^m$ chosen according to μ, Proposition 4.2 gives

$$\Pr(r^T b = \gamma) \leqslant \begin{cases} 1 - x\varepsilon + (x\varepsilon)^2, & \text{if } x\varepsilon \leqslant 1/2,\\ 3/4, & \text{if } x\varepsilon > 1/2.\end{cases} \tag{6.8}$$

Thus, setting $x_0 = \lfloor 1/(2\varepsilon)\rfloor$,

$$\Pr(Q_2) \leqslant q\sum_{x=1}^{x_0}\binom{m}{x}(q-1)^x(1 - x\varepsilon + (x\varepsilon)^2)^\ell + q\sum_{x=x_0+1}^{qT}\binom{m}{x}(q-1)^x(3/4)^\ell. \tag{6.9}$$

(The initial q-factors are for the choice of γ in (6.8).)

That the right-hand side of (6.9) is $o(1)$ is a routine calculation using (6.2) and $T = o(n)$.

$\mathscr{B}_3$. For $b \in \mathscr{B}_3$, say with $|\mathbf{F}'(b)| = k$, Lemmas 5.4 and 5.5 imply that, for $r \in \mathbf{F}^m$ chosen according to μ and any $\gamma \in \mathbf{F}$,

$$\Pr(r^T b = \gamma) < \frac{1}{k + 1 - o(1)}\,. \tag{6.10}$$

To see this, condition on $S := \{i : b_i \notin \mathbf{F}'(b) \text{ or } r_i \notin K\}$ setting $\overline{S} = [m] \setminus S$. Then $r_{\overline{S}}$ is chosen from $K^{\overline{S}}$ according to μ_K. So, since $r^T b = r_S^T b_S + r_{\overline{S}}^T b_{\overline{S}}$ (and since $Tf(n)/n \to \infty$), it follows from Lemma 5.4 (with σ, b and r replaced by μ_K, $b_{\overline{S}}$ and $r_{\overline{S}}$) and Lemma 5.5 that, for any $\gamma \in \mathbf{F}$, $\Pr(r^T b = \gamma | S) < 1/(k + 1 - o(1))$, provided (say)

$$|b_{\overline{S}}^{-1}(\gamma)| > T/2 \quad \text{for all} \quad \gamma \in \mathbf{F}'(b). \tag{6.11}$$

But (6.11) holds with probability $1 - o(1)$, so we have (6.10). (We have used the assumption $b \in \mathscr{B}_3$ to guarantee $\mathbf{F}'(b) \neq \emptyset$ (because $b \notin \mathscr{B}_2$) and to rule out equality in the application of Lemma 5.5.)

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

On the other hand, for the number of such bs we can use the crude bound

$$|\{b \in \mathbf{F}^m : |\mathbf{F}'(b)| = k\}| < \binom{q}{k}\binom{m}{qT} q^{qT} k^{m-qT} < e^{o(n)} k^m.$$

Combining these observations, we have

$$\Pr(Q_3) < q \sum_{k=1}^{q-1} e^{o(n)} k^m (k + 1 - o(1))^{-\ell} = o(1).$$

(The initial q corresponds to the choice of γ in (6.10).)

$\mathscr{B}_4$. Fix an $\mathbf{H}$-subspace $\mathbf{F}'$ of $\mathbf{F}$, say with $|\mathbf{F}'| = k$, and let $\mathscr{B}_4(\mathbf{F}') = \{b \in \mathscr{B}_4 : \mathbf{F}'(b) = \mathbf{F}'\}$. For any $b \in \mathscr{B}_4(\mathbf{F}')$ and $\gamma \in \mathbf{F}$, Lemmas 5.4 and 5.5 imply

$$\Pr(Lb = \gamma \underline{1}) < (k - o(1))^{-\ell}. \tag{6.12}$$

On the other hand, noting that $|b^{-1}(\mathbf{F}')| > m' := m - qT$, we have

$$|\mathscr{B}_4(\mathbf{F}')| \leqslant \binom{m}{qT} q^{qT} k^{m'} k \exp\{-\Omega(n)\}. \tag{6.13}$$

Namely, $|\mathscr{B}_4(\mathbf{F}')|$ is less than the number of ways to choose a set of qT coordinates, assign arbitrary elements of $\mathbf{F}$ to these coordinates, and for the remaining coordinates choose values from $\mathbf{F}'$ with some γ appearing at most $m/(2k)$ times. (The last factor – in which the exponent may be given more precisely as $((m'/k - m/(2k))^2/(2m'/k)$ – again comes from Proposition 5.6.) Combining (6.12) and (6.13) we have $\Pr(Q_4(\mathbf{F}')) \leqslant |\mathscr{B}_4(\mathbf{F}')| q (k - o(1))^{-\ell} = o(1)$, so also $\Pr(Q_4) \leqslant \sum_{\mathbf{F}'} \Pr(Q_4(\mathbf{F}')) = o(1)$.

$\mathscr{B}_5$. We again fix an $\mathbf{H}$-subspace $\mathbf{F}'$ of $\mathbf{F}$ of size k, and set $\mathscr{B} = \mathscr{B}_5(\mathbf{F}') = \{b \in \mathscr{B}_5 : \mathbf{F}'(b) = \mathbf{F}'\}$ (and $Q = \{Lb \in \mathbf{F}\underline{1} \text{ for some } b \in \mathscr{B}\}$).

Notice that $\mathscr{B}_5(\{0\}) = \emptyset$, since $\mathbf{F}'(b) = \{0\}$ implies $b \in \mathscr{B}_2$, and $\mathscr{B}_5(\mathbf{F}) = \emptyset$ since $\mathbf{F}'(b) = \mathbf{F}$ implies $b \in \mathscr{B}_0 \cup \mathscr{B}_1 \cup \mathscr{B}_4$; so we may assume $\mathbf{F}' \neq \{0\}, \mathbf{F}$.

Set $\mathbf{G} = \{\alpha \in \mathbf{F} : \alpha\mathbf{F}' = \mathbf{F}'\} \cup \{0\}$. Then $\mathbf{G} \supseteq \mathbf{H}$, $\mathbf{G}$ is easily seen to be a subfield of $\mathbf{F}$, and $\mathbf{G} \neq \mathbf{F}$ since $\mathbf{F}' \neq \{0\}, \mathbf{F}$. Moreover, crucially,

$$\text{if } \alpha, \beta \in \mathbf{F} \text{ and } \alpha\mathbf{F}' = \mathbf{F}' + \beta, \text{ then } \alpha \in \mathbf{G} \text{ (and } \beta \in \mathbf{F}') \tag{6.14}$$

(since $0 \in \alpha\mathbf{F}' = \mathbf{F}' + \beta \;\Rightarrow\; \beta \in \mathbf{F}' \;\Rightarrow\; \mathbf{F}' + \beta = \mathbf{F}' \;\Rightarrow\; \alpha \in \mathbf{G}$). So in fact,

$$\alpha \notin \mathbf{G} \Rightarrow |\mathbf{F}' \cap (\alpha\mathbf{F}' + \beta)| \leqslant |\mathbf{F}'|/p \tag{6.15}$$

(since $\mathbf{F}' \cap (\alpha\mathbf{F}' + \beta)$ is an affine $\mathbf{G}$-subspace properly contained in $\mathbf{F}'$).

Let $\delta = \mu(\mathbf{F} \setminus \mathbf{G})$. By (3.1) and the definition of K, we have

$$qf(n)/n > \delta > h(n)n^{-2}. \tag{6.16}$$

Let $\mathbf{F}''$ be a $\mathbf{G}$-subspace of $\mathbf{F}$ complementary to $\mathbf{F}'$. For $\alpha \in \mathbf{F}$ write $\alpha = \alpha' + \alpha''$ with $\alpha' \in \mathbf{F}'$, $\alpha'' \in \mathbf{F}''$, and similarly $b = b' + b''$ for $b \in \mathbf{F}^m$. Then, for $b \in \mathbf{F}^m$, $r \in \mathbf{G}^m$ and $\gamma \in \mathbf{F}$,

$$r^T b = \gamma \quad \text{if and only if} \quad r^T b' = \gamma' \quad \text{and} \quad r^T b'' = \gamma''. \tag{6.17}$$

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

Write r^i for the ith row of L. For the proof of (6.7) we fix and condition on the value of $R := \{i \in [\ell] : r^i \in \mathbf{G}^m\}$, setting $|R| = t$.

Note first that, for each $i \in [\ell]$, $\Pr(i \in R) = (1-\delta)^m =: \rho$. It then follows from (6.16) (and (6.5)) that a.s.

$$n/\sqrt{\log n} < \rho\ell/2 < t \quad \text{and} \quad \ell - t > (1-\rho)\ell/2 \tag{6.18}$$

(*e.g.*, via Proposition 5.6, noting that (6.16) implies $(1-\rho)\ell$ is $\omega(1)$). So we only need to consider t satisfying (6.18).

We need to break up the event Q a bit further. Let L_R consist of those rows of L indexed by R. Let $\mathscr{B}^x$ consist of those $b \in \mathscr{B}$ with $|\{i : b_i \notin \mathbf{F}'\}| = x$. By the definition of $\mathbf{F}'(b)$, $\mathscr{B}^x = \emptyset$ unless

$$x < qT, \tag{6.19}$$

so we assume from now on that x is in this range. To complete the proof of (6.7) it will suffice to show

$$\Pr(\cup_{x<qT} Q^x) = o(1). \tag{6.20}$$

(We will eventually need to give special treatment to $x = 0$.)

Now, in view of (6.17), Q^x requires that $L_R b'' \in \mathbf{F}\underline{1}$ for some $b'' \in (\mathbf{F}'')^m$ with $|\mathrm{supp}(b'')| = x$. So, again using Proposition 4.2, we have

$$\Pr(Q^x) \leqslant \binom{m}{x} |\mathbf{F}''|^{x+1} \left(\frac{1+(1-2\eta)^x}{2}\right)^t, \tag{6.21}$$

with $\eta = 1-(1-\varepsilon)/(1-\delta)$. (Recall we condition on R. Note that (6.16) and the second upper bound in (6.5) imply $\eta = (\log n + \omega(1))/n$. The extra factor of $|\mathbf{F}''|$ is for the choice of the common value of the products $(r^i)^T b''$, $i \in R$.)

For the last term in (6.21) we may again apply the bounds (4.1), and a calculation similar to that required in connection with $\mathscr{B}_2$ then shows that

$$\sum_{1 \leqslant x < qT} \Pr(Q^x) = o(1)$$

if $t > t_0 := (1 - \log^{-1} n)n$, and, with $n' = n^{1-\rho/4}$,

$$\sum_{n' \leqslant x < qT} \Pr(Q^x) = o(1)$$

if $t > n/2$. (When $x = o(1/\eta)$ our bound on $\Pr(Q^x)$ is essentially $\exp[x(\log(n/x) - \eta t)]$, which, in view of the lower bounds on t in (6.18), is acceptably small provided $x > n'$.)

So, to finish the proof of (6.20) it will suffice to show

$$\Pr(\cup_{0 \leqslant x < qT} Q^x) = o(1) \quad \text{if} \quad t \leqslant n/2, \tag{6.22}$$

$$\Pr(\cup_{0 \leqslant x < n'} Q^x) = o(1) \quad \text{if } t \leqslant t_0, \tag{6.23}$$

and

$$\Pr(Q^0) = o(1) \quad \text{if} \quad t > t_0. \tag{6.24}$$

For the proofs of (6.22) and (6.23) we fix (condition on) the values $s_i := |\{j : r^i_j \notin \mathbf{G}\}|$

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

(recall r^i is the ith row of L). Note that a.s. $s_i < \log n$ for all i (since $\delta < \log\log n/(3n)$ by (6.5) and (6.16)). We will show, with $b \in \mathscr{B}$, $r = r^i$ and $\gamma \in \mathbf{F}$,

$$\Pr(r^T b = \gamma | s_i = 0) < (1 + n^{-\Omega(1)})/k \tag{6.25}$$

(recall $k = |\mathbf{F}'|$) and

$$\Pr(r^T b = \gamma | s_i = s) < (1 - \Omega(1))/k \quad \text{if} \quad 1 \leqslant s < \log n. \tag{6.26}$$

Once we have these, it follows that, for any A,

$$\begin{aligned} \Pr(\cup_{x<A} Q^x) &< o(1) + \sum_{x<A} \binom{m}{x} q^x k^{m-x} [(1 + n^{-\Omega(1)})/k]^t [(1 - \Omega(1))/k]^{\ell - t} \\ &< o(1) + \sum_{x<A} \binom{m}{x} q^x k^{o(n \log^{-2} n)} \exp\{n^{1-\Omega(1)} - \Omega(\ell - t)\}. \end{aligned}$$

(The $o(1)$s are for the possibility that some s_i exceeds $\log n$.) This easily gives (6.22) and (6.23). (For the latter, note $\ell - t = \Omega(n/\log n)$ since $t < t_0$.)

For (6.25) set $I = \{i : b_i \in \mathbf{F}'\}$, with $\bar{I} := [m] \setminus I$, and condition on any specification of I and $r_{\bar{I}}$, letting $r_{\bar{I}}^T b_{\bar{I}} = \gamma'$. Note that we may then assume $\gamma - \gamma' \in \mathbf{F}'$, since otherwise we cannot have $r^T b = \gamma$. Treating elements of $\mathbf{F}'$ as vectors in $\mathbf{G}^d$ (where $|\mathbf{F}'| = |\mathbf{G}|^d$), we may represent b_I as a $d \times |I|$ $\mathbf{G}$-matrix N, so that, under our conditioning, the probability on the left-hand side of (6.25) is $\Pr(Nr_I = \gamma - \gamma')$ with $r_I \in \mathbf{G}^I$ chosen according to $\mu_{\mathbf{G}}$. So, noting that $b \notin \mathscr{B}_4$ implies that every vector in $\mathbf{G}^d$ appears at least $s := m/(2k)$ times as a column of N, whence $\operatorname{rank}(N) = d$ and $\operatorname{row}(N) \subseteq \mathscr{C}_{\mathbf{G}}(|I|, s)$, we may apply Lemma 4.4 with $\mathbf{F}$ replaced by $\mathbf{G}$ (and q by $|\mathbf{G}|$) and $\eta = (\varepsilon - \delta)/(1 - \delta)$ to obtain

$$|\Pr(r^T b = \gamma | s_i = 0) - k^{-1}| < (1 - |\mathbf{G}|\eta + O(\eta^2))^s < n^{-\Omega(1)}.$$

(Actually it is easy to see that $\operatorname{row}(N) \subseteq \mathscr{C}_{\mathbf{G}}(|I|, m/(2|\mathbf{G}|))$. We will need this stronger statement in the proof of (6.24).)

To verify (6.26) we regard r as given in three stages.

Stage 1. **(a)** Choose those r_js for which $b_j \notin \mathbf{F}'$.

(b) If fewer than s of these are in $\mathbf{F} \setminus \mathbf{G}$ then also choose all but one of those r_js which lie in $\mathbf{F} \setminus \mathbf{G}$. (Note that for these additional choices – and also for those in Stage 2 – both the indices j and the values r_j are random.)

Stage 2. Choose the last r_j belonging to $\mathbf{F} \setminus \mathbf{G}$. (If this was already chosen in Stage 1(a), then Stage 2 is vacuous.)

Stage 3. Choose the remaining r_js.

Let J be the (partly random) set of indices j for which r_j is chosen in the first two stages. Note that (since $j \notin J \Rightarrow b_j \in \mathbf{F}'$, $r_j \in \mathbf{G}$) we cannot have $r^T b = \gamma$ unless $r_J^T b_J \in \mathbf{F}' + \gamma$. So (6.26) will follow from

$$\Pr(r_J^T b_J \in \mathbf{F}' + \gamma) < 1 - \Omega(1) \tag{6.27}$$

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

and

$$\Pr(r^T b = \gamma | r_J^T b_J \in \mathbf{F}' + \gamma) < (1 + n^{-\Omega(1)})/k.$$

The second inequality is an application of Lemma 4.4 similar to that giving (6.25) which we omit, just noting that we use $|J| \leqslant x + s = o(n)$.

For (6.27) first notice that, since the number of r_js chosen in Stage 1(a) is $x = o(n)$, the probability that these include all s r_js lying outside $\mathbf{G}$ – *i.e.*, that Stage 2 is vacuous – is $o(1)$.

Suppose on the other hand that Stage 2 is not vacuous, and condition on the outcome of Stage 1 as well as on the value – but not the index – of the r_j chosen in Stage 2. Say the set of coordinates involved in Stage 1 is J', $r_{J'}^T b_{J'}$ is equal to γ', and the value chosen in Stage 2 is α. Then, letting j^* be the (as yet unknown) index chosen in Stage 2 (so $J = J' \cup \{j^*\}$), we have $r_J^T b_J = \gamma' + \alpha b_{j^*}$. So $r_J^T b_J \in \mathbf{F}' + \gamma$ (the event in (6.27)) if and only if

$$b_{j^*} \in \alpha^{-1}(\mathbf{F}' + \gamma - \gamma') \tag{6.28}$$

(an event depending on the random index j^*).

Now $b_{[m]\setminus J'}$ contains at least $m/(2k) - \log n = \Omega(n)$ copies of each $\beta \in \mathbf{F}'$, so that (under our conditioning)

$$\Pr((6.28) \text{ fails }) \geqslant \frac{|\mathbf{F}' \setminus \alpha^{-1}(\mathbf{F}' + \gamma - \gamma')|\Omega(n)}{m - |J'|},$$

which is $\Omega(1)$ by (6.14) and the fact that $|J'| = o(n)$. Thus we have (6.27).

Finally we turn to the proof of (6.24). We need to show, assuming $t > t_0$, that

$$\text{a.s. } Lb \notin \mathbf{F}\underline{1} \text{ for all } b \in \mathscr{B}^0. \tag{6.29}$$

Notice that, since $t > t_0$, (6.18) gives $n/\log n > (1-\rho)\ell/2$, which implies

$$\delta = o(1/n) \tag{6.30}$$

(actually $\delta = O((n \log n)^{-1})$).

Call an entry r_j^i of L *bad* if $r_j^i \in \mathbf{F} \setminus \mathbf{G}$ and at least one other entry in either the ith row or jth column of L is also in $\mathbf{F} \setminus \mathbf{G}$. Write Y and X for, respectively, the set of bad entries in L and the set of entries from $\mathbf{F} \setminus \mathbf{G}$ which are not bad. Then (writing E for expectation) $\mathsf{E}|Y| < \ell m\delta(\ell + m)\delta \leqslant 2\delta^2 n^3$ and $\mathsf{E}|X \cup Y| = \ell m\delta$, and it follows, using Proposition 5.6 (with (6.16) to guarantee $\ell m\delta = \omega(1)$) and (6.30), that a.s.

$$|Y| < o(\delta n^2) \quad \text{and} \quad |X| > \delta n^2/2. \tag{6.31}$$

It is thus enough to show (6.29) under conditioning on any specification of the non-$\mathbf{G}$ entries of L for which (6.31) holds; so for the rest of our discussion we fix such a specification. We may assume, without loss of generality, that the entries in X are $r_1^1, \ldots, r_u^u$, say with $r_i^i = \alpha_i$ ($\in \mathbf{F} \setminus \mathbf{G}$) for $i \in [u] =: S$.

For $y = \alpha m/k \in [m/(2k), m/k]$, set

$$\mathscr{B}(y) = \left\{ b \in \mathscr{B}^0 : \min_{\gamma \in \mathbf{F}'} |b^{-1}(\gamma)| = y \right\}.$$

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

(So $\mathscr{B}^0$ is the union of these $\mathscr{B}(y)$s.)

Fix $y = \alpha m/k$ as above and $b \in \mathscr{B}(y)$. Again (as in the proof of (6.25)) regard elements of $\mathbf{F}'$ as d-dimensional $\mathbf{G}$-vectors and b as a $d \times m$ $\mathbf{G}$-matrix N. Then $L_R b \in \mathbf{F}\underline{1}$ is equivalent to

$$L_R c \in \mathbf{G}\underline{1} \quad \text{for all} \quad c^T \in \text{row}(N).$$

That each vector in $\mathbf{G}^d$ appears at least y times as a column of N is easily seen to imply $\text{row}(N) \subseteq \mathscr{C}_{\mathbf{G}}(m, \alpha m/|\mathbf{G}|)$.

So we may apply Lemma 4.4 using $s = \alpha m/|\mathbf{G}|$ (resp. $\alpha m/|\mathbf{G}| - 1$) for rows in R (resp. S) to conclude that, under our conditioning,

$$\Pr(Lb \in \mathbf{F}\underline{1}) \leqslant \Pr(L_{R\cup S} b \in \mathbf{F}\underline{1}) \leqslant k(k^{-1} + o(n^{-\alpha}))^{t+u}. \tag{6.32}$$

(So we sacrifice the $o(\delta n^2)$ rows indexed by $[\ell] \setminus (R \cup S)$. As elsewhere we have a factor – in this case k – on the right-hand side for the choice of the common value of the entries of $L_{R\cup S} b$.)

To get from (6.32) to (6.29), we need bounds on the $|\mathscr{B}(y)|$s which are significantly less than k^m. For α not too close to 1, the large deviation bound of Proposition 5.6 suffices here, but for larger α we need to exploit the entries from X.

Notice first that we cannot have $L_{R\cup S} b \in \mathbf{F}\underline{1}$ unless $b_i \in \alpha_i^{-1}\mathbf{F}'$ for each $i \in [u]$, since for $i' \in R$, $(r^{i'})^T b \in \mathbf{F}'$, while $(r^i)^T b \in \mathbf{F}' + \alpha_i b_i$. So, recalling (see (6.15)) that $|\alpha_i^{-1}\mathbf{F}' \cap \mathbf{F}'| \leqslant k/p$, we find that the number of restrictions $b_{[u]}$ which allow even the *possibility* that $L_{R\cup S} b \in \mathbf{F}\underline{1}$ is at most $(k/p)^u$.

On the other hand, $b \in \mathscr{B}(y)$ implies that some $\gamma \in \mathbf{F}'$ appears at most y times in $b_{[m]\setminus[u]}$, so the number of possibilities for this part of b is (by Proposition 5.6) at most

$$k^{m-u+1} \exp\left[-\frac{(1 - u/m - \alpha)^2 m}{2k(1 - u/m)}\right] =: k^{m-u+1} e^{-Z},$$

provided $\alpha \leqslant 1 - u/m$.

Combining the preceding bounds (and defining $Q(y)$ according to our usual convention), we have

$$\Pr(Q(y)) < \begin{cases} k^{m-t-u+2} p^{-u} \exp[o(n^{1-\alpha}) - Z], & \text{if } \alpha \leqslant 1 - u/m, \\ k^{m-t-u+1} p^{-u} \exp[o(n^{1-\alpha})], & \text{if } \alpha > 1 - u/m. \end{cases} \tag{6.33}$$

We are left with some calculations to show

$$\sum \{\Pr(Q(y)) : m/(2k) \leqslant y \leqslant m/k\} = o(1) \tag{6.34}$$

(which gives (6.24)). Set $\xi = 1 - \alpha \in [0, 1/2]$. (Thus each of y, α and ξ determines the others, and we use whichever is convenient.)

To make use of the bound (6.33), observe first of all that

$$k^{m-t-u} p^{-u} < e^{-\Omega(u)}, \tag{6.35}$$

since $m - t - u < h(n)/(3e) + o(\delta n^2)$ (by (6.3) and (6.31)), $u > \delta n^2/2 > h(n)/2$ (by (6.31) and (6.16)), and $k \leqslant p^e$. (And note that, again by (6.31) and (6.16), the right-hand side of (6.35) is $o(1)$.)

This takes care of those terms in (6.34) for which $\xi \leqslant \log^{-1} n$, since for such ξ the

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

bound in (6.32) is $(1+o(1))k^{1-t-u}$ (and the total number of relevant bs is less than $k^m p^{-u}$), so that (using (6.35))

$$\sum\{\Pr(Q(y)) : \xi \leqslant \log^{-1} n\} < (1+o(1))k^{m-t-u+1}p^{-u} = o(1).$$

Suppose on the other hand that $\xi > \log^{-1} n$. Then, for $u \geqslant \xi m/2$, (6.33) and (6.35) give

$$\Pr(Q(y)) < \exp[-\Omega(u) + o(n^{\xi})] < \exp[-\Omega(\xi n) + o(n^{\xi})],$$

while for $u < \xi m/2$ we have $Z = \Omega(\xi^2 n)$ and

$$\Pr(Q(y)) < \exp[o(n^{\xi}) - \Omega(\xi^2 n)];$$

so in either case (using $\xi > \log^{-1} n$), $\Pr(Q(y)) < \exp[-\Omega(\xi^2 n)]$, and (6.34) follows. □

7. Necessity in Theorem 1.3

Write $I(q)$ for the limiting value $\prod(1-q^{-i})$ in (1.1). To complete the proof of Theorem 1.3, we will show that, if μ is a fixed probability distribution concentrated on some proper affine subfield of $\mathbf{F} = \mathbf{F}_q$, and $N = N_n$ is a random $n \times n$ $\mathbf{F}$-matrix chosen according to μ, then

$$\limsup \Pr(N_n \text{ is nonsingular}) < I(q). \tag{7.1}$$

(We have replaced M by N, preferring another role for M below.)

Let $\mathbf{H} = \alpha\mathbf{G} - \beta$ (with $\mathbf{G}$ a subfield and $\alpha \neq 0$) be the (unique) minimal affine subfield of $\mathbf{F}$ containing $\mathrm{supp}(\mu)$, and set $|\mathbf{G}| = p$. (Since we no longer need the prime field of $\mathbf{F}$, and since we like prime powers to be p or q, we have recycled p, which in particular is no longer required to be prime. We may of course assume $|\mathrm{supp}(\mu)| \geqslant 2$, so that $\mathbf{H}$ is defined.)

It will be convenient to assume $\alpha = 1$, which we may do since the probability in (7.1) is unchanged if we replace μ by σ with $\sigma(\gamma) = \mu(\alpha\gamma)$. We may then assume in addition that β is either 0 or an element of $\mathbf{F} \setminus \mathbf{G}$.

Let μ' be the probability distribution on $\mathbf{G}$ given by $\mu'(\gamma) = \mu(\gamma - \beta)$. Then

$$N = M - \beta J, \tag{7.2}$$

where $M = M_n$ is a random $n \times n$ $\mathbf{G}$-matrix chosen according to μ' (and J is the $n \times n$ matrix of 1s).

Since μ' is not concentrated on any proper affine subfield of $\mathbf{G}$, we know

$$\Pr(M_n \text{ is nonsingular}) \to I(p) < I(q).$$

So (7.1) is immediate if $\beta = 0$. We may therefore assume $\beta \in \mathbf{F} \setminus \mathbf{G}$, which yields a more interesting picture.

Proposition 7.1. (*Assuming* $\beta \neq 0$) *N is nonsingular if and only if either*

(a) *M is nonsingular, or*

(b) $\mathrm{rank}(M) = n-1$, $\underline{1}^T \notin \mathrm{row}(M)$ *and* $\underline{1} \notin \mathrm{col}(M)$.

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

(Note that this is a *deterministic* statement: the proposition holds whenever M, N are as in (7.2) (with M a $\mathbf{G}$-matrix) and $\beta \in \mathbf{F} \setminus \mathbf{G}$.)

Proof. Of course, nonsingularity of N requires $\text{rank}(M) \geqslant n-1$. It is also easy to see that if M has rank $n-1$ and violates (b) then N is singular, since $\underline{1}^T \in \text{row}(M)$ implies $\text{row}(N) \subseteq \text{row}(M)$ and $\underline{1} \in \text{col}(M)$ implies $\text{col}(N) \subseteq \text{col}(M)$, so in either case $\text{rank}(N) \leqslant \text{rank}(M) = n-1$. Hence the more interesting part of the proposition is the sufficiency of the conditions in (a) and (b).

Suppose first that M is nonsingular, and assume, for a contradiction, that $Nx = \underline{0}$ with $x \in \mathbf{F}^n \setminus \{\underline{0}\}$. Then $Mx = \beta Jx = (\beta \sum x_i)\underline{1}$, and x is a nonzero multiple of the unique $y \in \mathbf{G}^n$ for which $My = \underline{1}$. (Note $\sum x_i \neq 0$ since M is nonsingular.) But then $My = \beta Jy$ $(\neq \underline{0})$ gives the desired contradiction, since $y \in \mathbf{G}^n$ implies that the entries of My are in $\mathbf{G}$, while those of βJy are in $\mathbf{F} \setminus \mathbf{G}$.

Finally, suppose M satisfies the conditions in (b) and (again for a contradiction) $Nx = \underline{0}$ for some $\underline{0} \neq x \in \mathbf{F}^n$. As before, we have $Mx = (\beta \sum x_i)\underline{1}$, and this gives a contradiction: either $\sum x_i \neq 0$, in which case $\underline{1} \in \text{col}(M)$, or $\sum x_i = 0$, in which case $Mx = \underline{0}$ implies (since $\text{rank}(M) = n-1$) $\underline{1}^T \in \langle x \rangle^\perp = \text{row}(M)$. □

To complete the proof of (7.1) we need to say something about the probability of the event in Proposition 7.1(b). Actually, the limiting value here seems clear.

Conjecture 7.2.

$$\Pr(\text{rank}(M) = n-1,\ \underline{1}^T \notin \text{row}(M),\ \underline{1} \notin \text{col}(M)) \to (p-1)^{-2} I(p) \quad \textit{as} \quad n \to \infty.$$

(See Theorem 7.3 for motivation.)

While we do not see how to attack this, the machinery developed earlier does give a related statement adequate for our purposes. Set $Q = \{\text{rank}(M) = n-1 \text{ and } \underline{1}^T \notin \text{row}(M)\}$.

Theorem 7.3. $\Pr(Q) \to (p-1)^{-1} I(p) \quad \textit{as} \quad n \to \infty.$

(Of course, this and Conjecture 7.2 apply to a general finite field $\mathbf{G}$.)

Combining Theorem 7.3 with Proposition 7.1, we have (7.1):

$$\Pr(N \text{ nonsingular}) \leqslant \Pr(M \text{ nonsingular}) + \Pr(Q) \to (1 + (p-1)^{-1}) I(p) < I(q),$$

since $q \geqslant p^2$ implies $I(q) \geqslant \prod(1 - p^{-2i}) = I(p) \prod(1 + p^{-i})$.

Proof of Theorem 7.3. This is similar to the proof of Theorem 1.2 given in Section 3 (now applying our various lemmas with $\mathbf{G}$ in place of $\mathbf{F}$). As in that argument, let $\ell = n - h''(n)$, let r^i be the ith row of M, and $M^{(i)}$ the matrix whose rows are $r^1, \ldots, r^i$. Let $\mathscr{C} = \mathscr{C}_{\mathbf{G}}(n, s_0)$ with s_0 again as in (3.2). Set

$$R = \{M^{(\ell)} b \neq \underline{0} \text{ for all } b \in \mathbf{G}^n \setminus (\mathscr{C} \cup \{\underline{0}\})\},$$

and define

$$Q_\ell = R \wedge \{\text{rank}(M^{(\ell)}) = \ell \text{ and } \underline{1}^T \notin \text{row}(M^{(\ell)})\}.$$

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

Then

$$\Pr(Q_\ell) = 1 - o(1) \tag{7.3}$$

by Lemmas 3.2 and 6.1.

For $\ell < i, j \leqslant n$ define

$$Q'_{ij} = \begin{cases} \{r^j \notin \langle r^1, \dots, r^{j-1}, \underline{1}\rangle\}, & \text{if } j \neq i, \\ \{r^j \in \langle r^1, \dots, r^{j-1}\rangle\}, & \text{if } j = i, \end{cases}$$

and

$$Q_{ij} = Q_l \wedge \bigwedge_{\ell < k \leqslant j} Q'_{ij},$$

and for convenience set $Q_i := Q_{in}$, $Q_{i\ell} := Q_\ell$.

Then on Q_ℓ, Q is the disjoint union of $Q_{\ell+1}, \dots, Q_n$, so by (7.3),

$$\Pr(Q) = \sum_{i=\ell+1}^{n} \Pr(Q_i | Q_\ell) - o(1).$$

Now $\Pr(Q_i|Q_\ell) = \prod_{j=\ell+1}^{n} \Pr(Q_{ij}|Q_{i,j-1})$, and we assert that

$$\Pr(Q_{ij}|Q_{i,j-1}) = \begin{cases} (1 + o(1/n))(1 - p^{-(n-j)}), & \text{if } j < i, \\ (1 + o(1/n))p^{-(n-i+1)}, & \text{if } j = i, \\ (1 + o(1/n))(1 - p^{-(n-j+1)}), & \text{if } j > i. \end{cases} \tag{7.4}$$

To see this, note that $Q_{i,j-1}$ implies $\underline{1}^T \notin \mathrm{row}(M^{(j-1)})$, so that on $Q_{i,j-1}$, $\langle r^1, \dots, r^{j-1}, \underline{1}\rangle^\perp$ is a subspace of $\mathbf{F}^n$ contained in $\mathscr{C} \cup \{\underline{0}\}$ and having dimension $n - j$ if $j \leqslant i$ and $n - j + 1$ if $j > i$. The assertion then follows from Lemma 3.1.

Thus the probability of Q_i is asymptotically

$$p^{-(n-i+1)} \prod_{\ell < j < i} (1 - p^{-(n-j)}) \prod_{i < j \leqslant n} (1 - p^{-(n-j+1)}) \to p^{-(n-i+1)} I(p),$$

and

$$\Pr(Q) \to \sum_{i \geqslant 1} p^{-i} I(p) = (p-1)^{-1} I(p). \qquad \square$$

Acknowledgement

We would like to thank an anonymous referee for a very conscientious report as well as for telling us about [6]. The referee's detailed suggestions greatly improved the presentation of the results. We also thank David Lieberman for telling us about [3].

References

[1] Alon, N., Spencer. J. H. and Erdős, P. (1992) *The Probabilistic Method*, Wiley, New York.
[2] Blömer, J., Karp, R. and Welzl, E. (1997) The rank of sparse random matrices over finite fields. *Random Structures and Algorithms* **10** 407–419.
[3] Charlap, L. S., Rees, H. D. and Robbins, D. P. (1990) The asymptotic probability that a random biased matrix is invertible. *Discrete Math.* **82** 153–163.

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press

[4] Cooper, C. (2000) On the rank of random matrices. *Random Struct. Alg.* **16** 209–232.
[5] Kovalenko, I. N. and Levitskaya, A. A. (1975) Limiting behavior of the number of solutions of a system of random linear equations over a finite field and a finite ring. *Dokl. Akad. Nauk SSSR* **221** 778–781. In Russian.
[6] Kovalenko, I. N., Levitskaya, A. A. and Savchuk, M. N. (1986) *Selected Problems in Probabilistic Combinatorics*, Naukova Dumka, Kiev. In Russian.

https://doi.org/10.1017/S096354830100462X Published online by Cambridge University Press