Combinatorics, Probability and Computing (2024), **33**, pp. 643–707
doi:10.1017/S096354832400021X

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

The full rank condition for sparse random matrices

Amin Coja-Oghlan[1], Pu Gao[2], Max Hahn-Klimroth[1], Joon Lee[1], Noela Müller[3] and Maurice Rolvien[1]

[1]Faculty of Computer Science, TU Dortmund, Dortmund, Germany, [2]Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON, Canada, and [3]Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, the Netherlands
Corresponding author: Amin Coja-Oghlan; Email: amin.coja-oghlan@tu-dortmund.de

(Received 7 February 2022; revised 7 February 2024; accepted 7 May 2024; first published online 20 September 2024)

Abstract
We derive a sufficient condition for a sparse random matrix with given numbers of non-zero entries in the rows and columns having full row rank. The result covers both matrices over finite fields with independent non-zero entries and $\{0, 1\}$-matrices over the rationals. The sufficient condition is generally necessary as well.

Keywords: Random matrix; rank

2020 MSC Codes: Primary: 60B20; Secondary: 15B52

1. Introduction

1.1 Background and motivation

Few subjects in combinatorics have had as profound an impact on other disciplines as combinatorial random matrix theory. Prominent applications include powerful error correcting codes called low-density parity check codes [47], data compression [1, 52] and hashing [19]. Needless to mention, random combinatorial matrices are of keen interest to statistical physicists, too [40]. It therefore comes as no surprise that the subject has played a central role in probabilistic combinatorics since the early days [31–34]. The current state of affairs is that the theory of dense random matrices is significantly more advanced than that of sparse ones with a bounded average number of non-zero entries per row or column [50, 51]. This is in part because concentration techniques apply more easily in the dense case. Another reason is that the study of sparse random matrices is closely tied to the investigation of satisfiability thresholds of random constraint satisfaction problems, an area where many fundamental questions still await a satisfactory solution [4].

Perhaps the most basic question to be asked about any random matrix model is whether the resulting matrix will likely have full rank. This paper contributes a succinct sufficient condition that covers a broad range of sparse random matrix models. As we will see, the condition is essentially necessary as well. The main result can be seen as a satisfiability threshold theorem as the full rank property is equivalent to a random linear system of equations possessing a solution w.h.p. This formulation generalises a number of prior results such as the satisfiability threshold theorem

*Amin Coja-Oghlan is supported by DFG CO 646/3 and DFG CO 646/5. Pu Gao is supported by NSERC RGPIN-04173-2019. Max Hahn-Klimroth is supported by DFG CO 646/5. Noela Müller is supported by NWO Gravitation grant NETWORKS-024.002.003. An extended abstract version of this article appears in Proc. RANDOM 2023.

for the random k-XORSAT problem, one of the most intensely studied random constraint satisfaction problems (e.g. [2, 19, 21, 28, 44]). In addition, the main theorem covers other important random matrix models, including those that low-density parity check codes rely on [47].

The classical approach to tackling the full rank problem is the second moment method [3, 4]. This technique was pioneered in the seminal work on the k-XORSAT threshold of Dubois and Mandler [21]. Characteristic of this approach is the emergence of complicated analytic optimisation problems that encode entropy-probability trade-offs resulting from large deviations problems. Tackling these optimisation problems turns out to be rather challenging even in relatively simple special cases such as random k-XORSAT, as witnessed by the intricate calculations that Pittel and Sorkin [44] and Goerdt and Falke [23] had to go through. For the general model that we investigate here this proof technique thus appears futile.

We therefore pursue a totally different proof strategy, largely inspired by ideas from spin glass theory [40, 41]. In statistical physics jargon, the second moment method constitutes an 'annealed' computation. This means that we effectively average over all random matrices, including atypical specimens apt to boost the average. By contrast, the present work relies on a 'quenched' strategy based on a coupling argument that implicitly discards such pathological events. In effect, we will show that a truncated moment calculation confined to certain benign 'equitable' solutions suffices to determine the satisfiability threshold. This part of the proof is an extension of prior work of (some of) the authors on the normalised rank and variations on the random k-XORSAT problem [6, 10]. In addition, to actually compute the truncated second moment we need to determine the precise expected number of equitable solutions. To this end, we devise a new proof ingredient that combines local limit theorem techniques with algebraic ideas, particularly the combinatorial analysis of certain integer lattices. This technique can be seen as a generalisation of an argument of Huang [27] for the study of adjacency matrices of d-regular random graphs.

Let us proceed to present the main results of the paper. The first theorem deals with random matrices over finite fields. As an application we obtain a result on sparse $\{0, 1\}$-matrices over the rationals.

1.2 Results

We work with the comprehensive random matrix model from [10]. Hence, let $\boldsymbol{d} \geq 0$, $\boldsymbol{k} \geq 3$ be independent integer-valued random variables with $\mathbb{P}(\boldsymbol{d} = 0) < 1$ and $\mathbb{E}[\boldsymbol{d}^{2+\eta}] + \mathbb{E}\left[\boldsymbol{k}^{2+\eta}\right] < \infty$ for an arbitrarily small $\eta > 0$. Let $(\boldsymbol{d}_i, \boldsymbol{k}_i)_{i\geq 1}$ be independent copies of $(\boldsymbol{d}, \boldsymbol{k})$ and set $d = \mathbb{E}[\boldsymbol{d}]$, $k = \mathbb{E}[\boldsymbol{k}]$. Moreover, let $\mathfrak{d}$ and $\mathfrak{k}$ be the greatest common divisors of the support of $\boldsymbol{d}$ and $\boldsymbol{k}$, respectively. Further, let $n > 0$ be an integer divisible by $\mathfrak{k}$ and let $\boldsymbol{m}$ be a Poisson variable with mean dn/k, independent of $(\boldsymbol{d}_i, \boldsymbol{k}_i)_{i\geq 1}$. Routine arguments reveal that the event

$$\sum_{i=1}^{n} \boldsymbol{d}_i = \sum_{j=1}^{\boldsymbol{m}} \boldsymbol{k}_j \tag{1.1}$$

occurs with probability $\Omega(n^{-1/2})$ for such n [10, Proposition 1.10]. Given (1.1), we then define the simple random bipartite graph $\mathbb{G} = \mathbb{G}_n(\boldsymbol{d}, \boldsymbol{k})$ on a set $\{a_1 \ldots, a_{\boldsymbol{m}}\}$ of *check nodes* and a set $\{x_1, \ldots, x_n\}$ of *variable nodes* as a uniformly random simple graph such that the degree of a_i equals $\boldsymbol{k}_i$ and the degree of x_j equals $\boldsymbol{d}_j$, for all i, j. The existence of such a graph is proven in [10, Proposition 1.10]. Following coding theory jargon, we refer to $\mathbb{G}$ as the *Tanner graph*. The edges of $\mathbb{G}$ are going to mark the positions of the non-zero entries of the random matrix. The entries themselves will depend on whether we deal with a finite field or the rationals.

1.2.1 Finite fields

Suppose that $q \geq 2$ is a prime power, let $\mathbb{F}_q$ signify the field with q elements and let $\boldsymbol{\chi}$ be a random variable that takes values in the set $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$ of units of $\mathbb{F}_q$. Moreover, let $(\boldsymbol{\chi}_{i,j})_{i,j\geq 1}$ be copies of χ, mutually independent and independent of the d_i, k_i, m and $\mathbb{G}$. Finally, let $\mathbb{A} = \mathbb{A}_n(d, k, \chi)$ be the $m \times n$-matrix with entries

$$\mathbb{A}_{i,j} = \mathbb{1}\left\{a_i x_j \in E(\mathbb{G})\right\} \cdot \chi_{i,j}.$$

Hence, the i-th row of $\mathbb{A}$ contains k_i non-zero entries and the j-th column contains d_j non-zero entries.

The following theorem provides a sufficient condition for $\mathbb{A}$ having full row rank. The condition comes in terms of the probability generating functions $D(x)$ and $K(x)$ of d and k. Since $\mathbb{E}[d^2] + \mathbb{E}[k^2] < \infty$, we may define

$$\Phi : [0, 1] \to \mathbb{R}, \qquad z \mapsto D\left(1 - K'(z)/k\right) - \frac{d}{k}\left(1 - K(z) - (1 - z)K'(z)\right). \tag{1.2}$$

Theorem 1.1. *Let $d \geq 0, k \geq 3$ and q be a fixed prime power such that $\mathfrak{d} = \gcd(\mathrm{supp}(d))$ and q are coprime. If*

$$\Phi(z) < \Phi(0) \qquad \text{for all } 0 < z \leq 1, \tag{1.3}$$

then $\mathbb{A}$ has full row rank over $\mathbb{F}_q$ w.h.p.

Observe that the function Φ does not depend on q. Hence, neither does (1.3).

The sufficient condition (1.3) is generally necessary, too. Indeed, ref. [10, Theorem 1.1] determines the likely value of the *normalised* rank of $\mathbb{A}$:

$$\frac{\mathrm{rk}(\mathbb{A})}{n} \xrightarrow{\mathbb{P}} 1 - \max_{z \in [0,1]} \Phi(z) \qquad \text{as } n \to \infty. \tag{1.4}$$

Since $k \geq 3$, definition (1.2) ensures that $\Phi(0) = 1 - d/k$ and thus $n\Phi(0) \sim n - m$ w.h.p. Hence, (1.4) implies that $\mathrm{rk}(\mathbb{A}) \leq m - \Omega(n)$ w.h.p. unless $\Phi(z)$ attains its maximum at $z = 0$. In other words, $\mathbb{A}$ has full row rank *only if* $\Phi(z) \leq \Phi(0)$ for all $0 < z \leq 1$. Indeed, in Section 1.3 we will discover examples that require a strict inequality as in (1.3). The condition that q and $\mathfrak{d}$ be coprime is generally necessary as well, as we will see in Example 1.7 below.

Let us emphasise that (1.4) does not guarantee that $\mathbb{A}$ has full row rank w.h.p. even if (1.3) is satisfied. Rather due to the normalisation of the l.h.s. (1.4) only implies the much weaker statement $\mathrm{rk}(\mathbb{A}) = m - o(n)$ w.h.p. Hence, in the case that (1.3) is satisfied, Theorem 1.1 improves over the asymptotic estimate (1.4) rather substantially. Unsurprisingly, this stronger result also requires a more delicate proof strategy.

We finally remark that condition (1.3) in combination with (1.4) enforces that $d \leq k$. Moreover, if $d = k$, then $\Phi(0) = 0 \leq D(0) = \Phi(1)$, such that condition (1.3) also cannot be satisfied for such d, k. Thus, $d < k$ for all matrices to which Theorem 1.1 applies, and the subject of Theorem 1.1 is the rank of rectangular matrices with asymptotically more columns than rows. For sparse and square Bernoulli matrices, Glasgow, Kwan, Sah and Sawhney [24] recently provided precise combinatorial descriptions of the exact real rank. The method of [24] applies to both symmetric and asymmetric Bernoulli matrices and relates the real rank to the Karp-Sipser core of the associated graph models. Theorem 1.1 also does not make a quantitative statement about the rate of convergence. While such a quantification could in principle be obtained from our proof, we do not expect it to be very close to optimal and have therefore not pursued this.

1.2.2 Zero-one matrices over the rationals

Apart from matrices over finite fields, the rational rank of sparse random $\{0, 1\}$-matrices has received a great deal of attention [50, 51]. The random graph $\mathbb{G}$ naturally induces a $\{0, 1\}$-matrix, namely the $m \times n$-biadjacency matrix $\mathbb{B} = \mathbb{B}(\mathbb{G})$. Explicitly, $\mathbb{B}_{ij} = \mathbb{1}\{a_i x_j \in E(\mathbb{G})\}$. As an application of Theorem 1.1 we obtain the following result.

Corollary 1.2. *If (1.3) is satisfied then the random matrix $\mathbb{B}$ has full row rank over $\mathbb{Q}$ w.h.p.*

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Since (1.4) holds for random matrices over the rationals as well, Corollary 1.2 is optimal to the extent that $\mathbb{B}$ fails to have full row rank w.h.p. if $\max_{x\in[0,1]} \Phi(x) > \Phi(0)$. Moreover, in Example 1.4 we will see that $\mathbb{B}$ does not generally have full rank w.h.p. unless $x = 0$ is the unique maximiser of Φ.

1.2.3 Fixed-degree sequences

In Section 2.3, we consider a more general model for $\mathbb{A}$, where the sequences $(d_i)_{i\geq 1}$ and $(k_j)_{j\geq 1}$ are specified instead of being obtained by taking i.i.d. copies of $\boldsymbol{d}$ and $\boldsymbol{k}$. Under analogous conditions like (1.3) together with some additional 'smoothness' conditions for $d_1, \ldots, d_n$ and $k_1, \ldots, k_m$, we also show that the matrix $\underline{\mathbb{A}}$ corresponding to the fixed-degree setting has full row rank (see Proposition 2.1).

1.3 Examples

To illustrate the power of Theorem 1.1 and Corollary 1.2 we consider a few instructive special cases of distributions $\boldsymbol{d}, \boldsymbol{k}, \boldsymbol{\chi}$.

Example 1.3 (random k-XORSAT). In random k-XORSAT we are handed a number of independent random constraints c_i of the type

$$c_i = y_{i1} \text{ XOR } \cdots \text{ XOR } y_{ik}, \tag{1.5}$$

where each y_{ij} is either one of n available Boolean variables $x_1, \ldots, x_n$ or a negation $\neg x_1, \ldots, \neg x_n$. The obvious question is to determine the satisfiability threshold, that is, the maximum number of random constraints that can be satisfied simultaneously w.h.p.

Because Boolean XOR boils down to addition over $\mathbb{F}_2$, this problem can be rephrased as the full rank problem for the random matrix $\mathbb{A}$ with $q = 2$, $\boldsymbol{k} = k$ fixed to a deterministic value and $\boldsymbol{d} \sim \text{Po}(d)$ for a parameter $d > 0$. To elaborate, because the constraints c_i are drawn uniformly and independently, we can think of each as tossing k balls randomly into n bins that represent $x_1, \ldots, x_n$. If there are $\boldsymbol{m} \sim \text{Po}(dn/k)$ constraints c_i, the joint distribution of the variable degrees coincides with the distribution of $(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n)$ subject to the condition (1.1). Furthermore, the random negation patterns of the constraints (1.5) amount to choosing a random right-hand side vector $\boldsymbol{y}$ for which we are to solve $\mathbb{A}x = \boldsymbol{y}$.

Since the generating functions of $\boldsymbol{d}, \boldsymbol{k}$ work out to be $D(z) = \exp(d(z-1))$ and $K(z) = z^k$, we obtain

$$\Phi_{d,k}(z) = \exp(-dz^{k-1}) - \frac{d}{k}\left(1 - kz^{k-1} + (k-1)z^k\right).$$

Thus, Theorem 1.1 implies that for a given $k \geq 3$ the threshold of d up to which random k-XORSAT is satisfiable w.h.p. equals the largest d such that

$$\Phi_{d,k}(z) < \Phi_{d,k}(0) = 1 - d/k \qquad \text{for all } 0 < z \leq 1. \tag{1.6}$$

A few lines of calculus verify that (1.6) matches the formulas for the k-XORSAT threshold derived by combinatorial methods tailored to this specific case [19, 21, 41, 44]. Theorem 1.1 also encompasses the generalisations to other finite fields $\mathbb{F}_q$ from [6, 23]. (For $d = 6.5$ and $k = 7$ see the left of Fig. 1.)

Example 1.4 (identical distributions). An interesting scenario arises when $\boldsymbol{d}, \boldsymbol{k}$ are identically distributed. For example, suppose that $\mathbb{P}[\boldsymbol{d} = 3] = \mathbb{P}[\boldsymbol{d} = 4] = \mathbb{P}[\boldsymbol{k} = 3] = \mathbb{P}[\boldsymbol{k} = 4] = 1/2$. Thus, $D(z) = K(z) = (z^3 + z^4)/2$ and

$$\Phi(z) = \frac{256z^{12} + 768z^{11} + 864z^{10} - 1808z^9 - 4959z^8 - 3780z^7 + 6111z^6 + 10584z^5 - 3234z^4 - 4802z^3}{4802}.$$

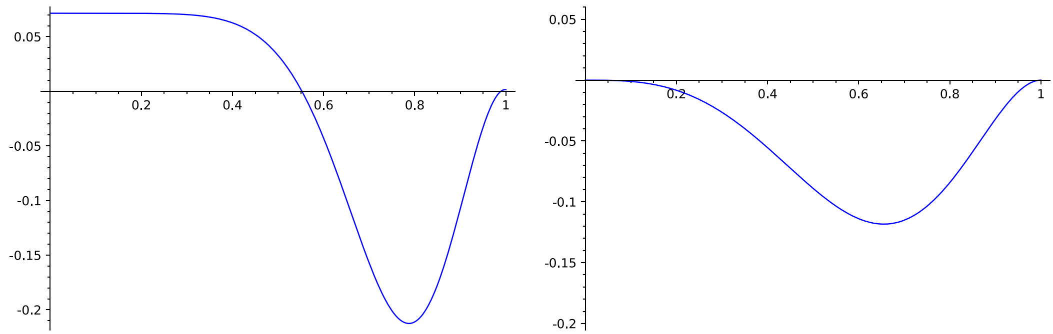

Figure 1. Left: The function Φ from Example 1.3 with $D(z) = \exp(6.5(z-1))$ and $K(z) = z^7$. Right: The function Φ from Example 1.4 with $D(z) = K(z) = (z^3 + z^4)/2$.

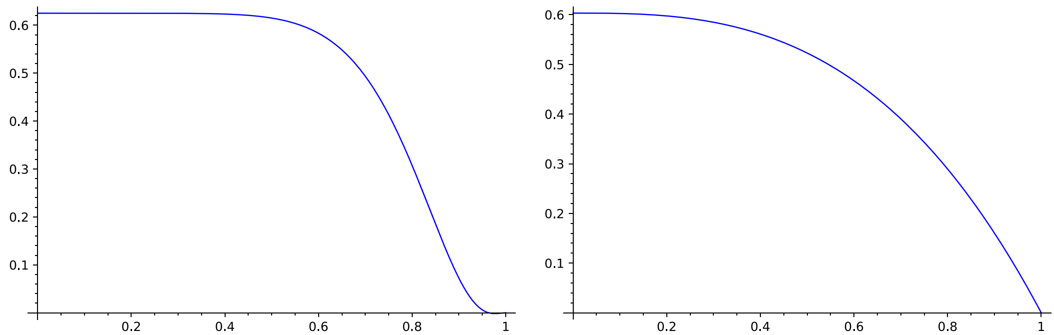

Figure 2. Left: The function Φ from Example 1.5 with $D(z) = z^3$, $K(z) = z^8$. Right: The function Φ from Example 1.6 with $D(z) = \sum_{\ell=1}^{\infty} \zeta(3.5)^{-1} z^{\ell} \ell^{-3.5}$ and $K(x) = x^3$.

This function attains two identical maxima, namely $\Phi(0) = \Phi(1) = 0$ (See the right of Fig. 1). Since the degrees $\boldsymbol{k}_i, \boldsymbol{d}_i$ are chosen independently subject only to (1.1), the probability that $\mathbb{A}$ has more rows than columns works out to be $1/2 + o(1)$. As a consequence, $\mathbb{A}$ cannot have full row rank w.h.p. This example shows that the condition that 0 be the *unique* maximiser of $\Phi(x)$ is generally necessary to ensure that $\mathbb{A}$ has full row rank. The same applies to the rational rank of $\mathbb{B}$.

Example 1.5 (fixed $\boldsymbol{d}, \boldsymbol{k}$). Suppose that both $\boldsymbol{d} = d, \boldsymbol{k} = k \geq 3$ are constants rather than genuinely random. Then

$$\Phi(z) = \left(1 - z^{k-1}\right)^d - \frac{d}{k}\left(1 - kz^{k-1} + (k-1)z^k\right).$$

Clearly, $\mathbb{A}$ cannot have full row rank unless $d \leq k$, while Theorem 1.1 implies that $\mathbb{A}$ has full row rank w.h.p. if $d < k$ (See the left of Fig. 2). This result was previously established via the second moment method [42]. But in the critical case $d = k$ the function $\Phi(z)$ attains its identical maxima at $z = 0$ and $z = 1$. Specifically, $0 = \Phi(0) = \Phi(1) > \Phi(z)$ for all $0 < z < 1$. Hence, Theorem 1.1 does not cover this special case. Nonetheless, Huang [27] proved that the random $\{0, 1\}$-matrix $\mathbb{B}$ has full rational rank w.h.p. The proof is based on a delicate moment computation in combination with a precise local expansion around the equitable solutions.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Example 1.6 (power laws). Let $\mathbb{P}(\boldsymbol{d}=\ell) \propto \ell^{-\alpha}$ for some $\alpha > 3$ and $\boldsymbol{k}=k \geq 3$. Thus,

$$D(z)=\frac{1}{\zeta(\alpha)}\sum_{\ell=1}^{\infty}\frac{z^\ell}{\ell^\alpha}, \quad K(z)=z^k,$$

$$\Phi(z)=D\left(1-z^{k-1}\right)-\frac{\zeta^{-1}(\alpha)\zeta(\alpha-1)}{k}\left(1-kz^{k-1}+(k-1)z^k\right).$$

Since

$$\Phi'(z)=-(k-1)z^{k-2}D'\left(1-z^{k-1}\right)+\frac{\zeta^{-1}(\alpha)\zeta(\alpha-1)}{k}\left(k(k-1)(z^{k-1}-z^{k-2})\right)<0,$$

the function $\Phi(z)$ is strictly decreasing on $(0, 1)$. Therefore, (1.3) is satisfied (For $\alpha=3.5$ and $k=3$ see the right of Fig. 2).

Example 1.7 (zero row sums). Theorem 1.1 requires the assumption that q and the g.c.d. $\mathfrak{d}$ of the support of $\boldsymbol{d}$ be coprime. This assumption is indeed necessary. To see this, consider the case that $q=2$, $\boldsymbol{\chi}=1$, $\boldsymbol{d}=4$ and $\boldsymbol{k}=8$ deterministically. Then the rows of $\mathbb{A}$ always sum to zero. Hence, $\mathbb{A}$ cannot have full row rank.

2. Overview

In contrast to much of the prior work on the rank problem, random k-XORSAT and random constraint satisfaction problems generally, the proofs of the main results do not rely on an 'annealed' second moment computation. Such arguments appear to be far too susceptible to large deviation effects to extend to as general a random matrix model as we deal with here. Instead, we proceed by way of a 'quenched' argument that enables us to discard pathological events. As a result, it suffices to carry out the moment calculation in the particularly benign case of 'equitable' solutions.

This proof strategy draws on but substantially generalises tools that were developed towards the approximate rank formula (1.4) and variations on random k-XORSAT [6, 10]. In addition, to actually prove that $\mathbb{A}$ has full rank with *high* probability we will need to carry out a meticulous, asymptotically exact calculation of the expected number of equitable solutions. A key element of this analysis will be a delicate analysis of the lattices generated by certain integer vectors that encode conceivable equitable solutions. This part of the proof, which generalises a part of Huang's argument for the adjacency matrices of random d-regular graphs [27], combines local limit techniques with a whiff of linear algebra.

To describe the proof strategy in detail let us first explore the 'annealed' path, discover its pitfalls and then apply the lessons learned to develop a workable 'quenched' strategy. The bulk of the proof deals with the random matrix model from Section 1.2.1 over the finite field $\mathbb{F}_q$; the rational case from Corollary 1.2 comes out as an easy consequence.

In order to reduce fluctuations we are going to condition on the σ-algebra $\mathfrak{A}$ generated by $\boldsymbol{m}$, $(\boldsymbol{k}_i)_{i\geq 1}$, $(\boldsymbol{d}_i)_{i\geq 1}$ and by the numbers $\boldsymbol{m}(\chi_1, \ldots, \chi_\ell)$ of checks of degree $\ell \geq 3$ with coefficients $\chi_1, \ldots, \chi_\ell \in \mathbb{F}_q^*$. We write $\mathbb{P}_{\mathfrak{A}}=\mathbb{P}[\,\cdot\mid\mathfrak{A}]$ and $\mathbb{E}_{\mathfrak{A}}=\mathbb{E}[\,\cdot\mid\mathfrak{A}]$ for brevity.

2.1 Moments and deviations

We already alluded to how the full rank problem for the random matrix $\mathbb{A}$ over $\mathbb{F}_q$ can be viewed as a random constraint satisfaction problem. Indeed, suppose we draw a right-hand side vector $\boldsymbol{y}\in\mathbb{F}_q^{\boldsymbol{m}}$ independently of $\mathbb{A}$. Then $\mathbb{A}$ has full row rank w.h.p. iff the random linear system $\mathbb{A}x=\boldsymbol{y}$ admits a solution w.h.p. For if $\mathrm{rk}\mathbb{A}<\boldsymbol{m}$, then the image $\mathbb{A}\mathbb{F}_q^n$ is a proper subspace of $\mathbb{F}_q^{\boldsymbol{m}}$ and thus the random linear system $\mathbb{A}x=\boldsymbol{y}$ has a solution with probability at most $1/q$. Naturally, the random linear system is nothing but a random constraint satisfaction problem with $\boldsymbol{m}$ constraints and n variables.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Over the past two decades the second moment method has emerged as the default approach to pinpointing satisfiability thresholds of random constraint satisfaction problems [3, 4]. Indeed, one of the first success stories was the random 3-XORSAT problem, which boils down directly to a full rank problem over $\mathbb{F}_2$ [21]. In fact, as we saw in Example 1.3, to mimic 3-XORSAT we just set $q = 2$, $\boldsymbol{d} = \mathrm{Po}(d)$ for some $d > 0$ and $\boldsymbol{k} = 3$ deterministically. In addition, draw $\boldsymbol{y} \in \mathbb{F}_2^{\boldsymbol{m}}$ uniformly and independently of everything else.

We try the second moment method on the number $\boldsymbol{Z} = \boldsymbol{Z}(\mathbb{A}, \boldsymbol{y})$ of solutions to $\mathbb{A}x = \boldsymbol{y}$ given $\mathfrak{A}$. Since $\boldsymbol{y}$ is independent of $\mathbb{A}$, for any fixed vector $x \in \mathbb{F}_2^n$ the event $\mathbb{A}x = \boldsymbol{y}$ has probability $2^{-\boldsymbol{m}}$. Consequently,

$$\mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}] = 2^{n-\boldsymbol{m}}. \tag{2.1}$$

Hence, (2.1) recovers the obvious condition that we cannot have more rows than columns. Since $\boldsymbol{m} \sim \mathrm{Po}(dn/3)$, (2.1) boils down to $d < 3$.

The second moment method now rests on the hope that we may be able to show that $\mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}^2] \sim \mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}]^2$. Then Chebyshev's inequality would imply $\boldsymbol{Z} \sim \mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}]$ w.h.p., and thus, in light of (2.1), that $\mathbb{A}x = \boldsymbol{y}$ has a solution w.h.p.

Concerning the computation of $\mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}^2]$, because the set of solutions is either empty or a translation of the kernel, we obtain

$$\begin{aligned}\mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}^2] &= \sum_{\sigma,\tau \in \mathbb{F}_q^n} \mathbb{P}_{\mathfrak{A}}\left[\mathbb{A}\sigma = \mathbb{A}\tau = \boldsymbol{y}\right] = \sum_{\sigma,\tau \in \mathbb{F}_q^n} \mathbb{P}_{\mathfrak{A}}\left[\mathbb{A}\sigma = \boldsymbol{y}\right] \mathbb{P}_{\mathfrak{A}}\left[\sigma - \tau \in \ker \mathbb{A}\right] \\ &= \mathbb{E}_{\mathfrak{A}}\left[\boldsymbol{Z}\right] \mathbb{E}_{\mathfrak{A}} |\ker \mathbb{A}|. \end{aligned} \tag{2.2}$$

To calculate the expected kernel size we notice that the probability that a vector x is in the kernel depends on its Hamming weight. For instance, the zero vector always belongs to the kernel, while the all-ones vector $\mathbb{1}$ does not w.h.p. More systematically, invoking inclusion/exclusion, we find that for a vector x of Hamming weight w we have $\mathbb{P}_{\mathfrak{A}}\left[x \in \ker \mathbb{A}\right] \sim \left[(1 + (1 - 2w/n)^3)/2\right]^{\boldsymbol{m}}$. Since the total number of such vectors comes to $\binom{n}{w}$, we obtain

$$\mathbb{E}_{\mathfrak{A}} |\ker \mathbb{A}| = \sum_{w=0}^{n} \binom{n}{w} \left(\frac{1 + (1 - 2w/n)^3}{2}\right)^{\boldsymbol{m}}. \tag{2.3}$$

Taking logarithms, invoking Stirling's formula and parametrising $w = zn$, we simplify (2.3) to

$$\log\, \mathbb{E}_{\mathfrak{A}} |\ker \mathbb{A}| \sim n \cdot \max_{z \in [0,1]} -z \log z - (1 - z) \log (1 - z) + \frac{\boldsymbol{m}}{n} \log \frac{1 + (1 - 2z)^3}{2} \quad \text{(cf. [21])}. \tag{2.4}$$

If we substitute $z = 1/2$ into (2.4), the expression further simplifies to $(n - \boldsymbol{m}) \log 2$. Hence, if the maximum is attained at another value $z \neq 1/2$, then (2.4) yields $\mathbb{E}_{\mathfrak{A}} |\ker \mathbb{A}| \gg 2^{n-\boldsymbol{m}}$ and the second moment method fails.

Figure 3 displays (2.4) for $d = 2.5$ and $d = 2.7$. While for $d = 2.5$ the function takes its maximum at $z = 1/2$, for $d = 2.7$ the maximum is attained at $z \approx 0.085$. However, the true random 3-XORSAT threshold is $d \approx 2.75$ [21]. Thus, the naive second moment calculation falls short of the real threshold.

How so? The expression (2.4) does not determine the 'likely' but the expected size of the kernel, a value prone to large deviations effects. Indeed, because the number of vectors in the kernel scales exponentially with n, an exponentially unlikely event that causes an exceptionally large kernel may end up dominating $\mathbb{E}_{\mathfrak{A}} |\ker \mathbb{A}|$. Precisely such an event manifests itself in the left local maximum in Fig. 3. Moreover, as we approach the satisfiability threshold such large deviations issues are

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

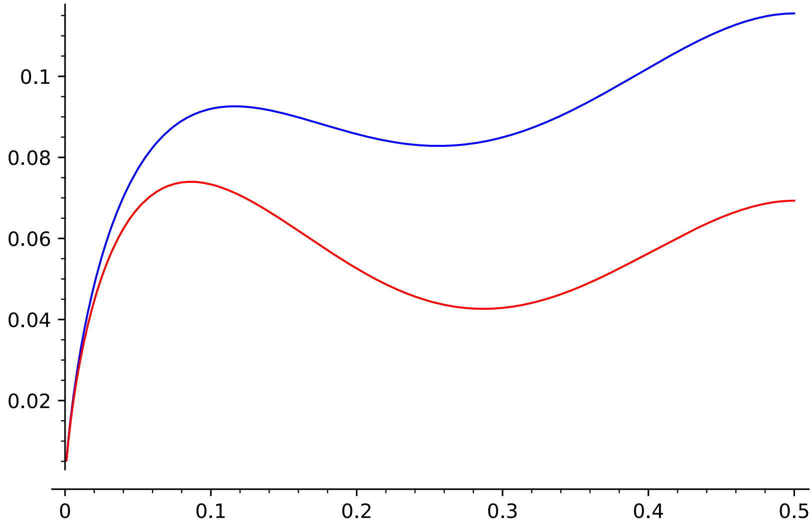

Figure 3. The r.h.s. of (2.4) for $d = 2.5$ (blue) and $d = 2.7$ (red) in the interval $[0, \frac{1}{2}]$.

compounded by a diminishing error tolerance. Indeed, while for $d = 2.5$ the value at $z = 1/2$ just swallows the spurious maximum, this is no longer the case for $d = 2.7$.

For random k-XORSAT Dubois and Mandler managed to identify the precise large deviations effect at work. It stems from fluctuations of a densely connected sub-graph of $\mathbb{G}$ called the 2-core, obtained by iteratively pruning nodes of degree less than two along with their neighbours (if any). Dubois and Mandler pinpointed the 3-XORSAT threshold by applying the second moment method to the minor $\mathbb{A}^{(2)}$ induced by $\mathbb{G}^{(2)}$ while conditioning on the 2-core having its typical dimensions.

The technical difficulty is that the rows of $\mathbb{A}^{(2)}$ are no longer independent. Indeed, $\mathbb{A}^{(2)}$ is distributed as a random matrix with a truncated Poisson $\boldsymbol{d}^{(2)} \sim \mathrm{Po}_{\geq 2}(d')$ with $d' = d'(d, k) > 0$ as the distribution of the variable degrees. Unfortunately, the given-degrees model leads to a fairly complicated moment computation. Instead of the humble one-dimensional problem from (2.4) we now face parameters $(z_i)_{i \geq 2}$ that gauge the fraction of variables of each possible degree i set to one. Additionally, on the constraint side we need to keep track of the number of equations with zero and with two variables set to one. Of course, these variables are tied together through the constraint that the total Hamming weight on the variable side matches that on the constraint side.

With a deal of diligence Dubois and Mandler managed to solve this optimisation problem. However, even just the step on to check degrees $k > 3$ turns out to be tricky because now we need to keep track of all the possible ways in which a k-ary parity constraint can be satisfied [19, 44]. Yet even these difficulties are eclipsed by those that result from merely advancing to fields of size $q = 3$ [23].

Not to mention entirely general degree distributions $\boldsymbol{d}, \boldsymbol{k}$ and general fields $\mathbb{F}_q$ as in Theorem 1.1. The ensuing optimisation problem comes in terms of variables $(z_i)_{i \in \mathrm{supp}\boldsymbol{d}}$ that range over the space $\mathscr{P}(\mathbb{F}_q)$ of probability distributions on $\mathbb{F}_q$. Additionally, there is a second set of variables $(\hat{z}_{\chi_1,\ldots,\chi_\ell})_{\ell \in \mathrm{supp}\boldsymbol{k},\, \chi_1,\ldots,\chi_\ell \in \mathrm{supp}\boldsymbol{\chi}}$ to go with the rows of $\mathbb{A}$ whose non-zero entries are precisely $\chi_1, \ldots, \chi_\ell$. These variables range over probability distributions on solutions $\sigma \in \mathbb{F}_q^\ell$ to $\chi_1\sigma_1 + \cdots + \chi_\ell\sigma_\ell = 0$. In terms of these variables we would need to solve

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

$$\max \sum_{\sigma\in\mathbb{F}_q} \mathbb{E}\left[(\boldsymbol{d}-1)z_{\boldsymbol{d}}(\sigma)\log z_{\boldsymbol{d}}(\sigma)\right]$$

$$-\frac{d}{k}\mathbb{E}\left[\sum_{\substack{\sigma_1,\ldots,\sigma_{\boldsymbol{k}}\in\mathbb{F}_q\\ \boldsymbol{\chi}_{1,1}\sigma_1+\cdots+\boldsymbol{\chi}_{1,\boldsymbol{k}}\sigma_{\boldsymbol{k}}=0}} \hat{z}_{\boldsymbol{\chi}_{1,1},\ldots,\boldsymbol{\chi}_{1,\boldsymbol{k}}}(\sigma_1,\ldots,\sigma_{\boldsymbol{k}})\log\hat{z}_{\boldsymbol{\chi}_{1,1},\ldots,\boldsymbol{\chi}_{1,\boldsymbol{k}}}(\sigma_1,\ldots,\sigma_{\boldsymbol{k}})\right] \tag{2.5}$$

$$\text{s.t.}\quad \mathbb{E}[\boldsymbol{d}z_{\boldsymbol{d}}(\tau)]=\mathbb{E}\left[\sum_{\substack{\sigma_1,\ldots,\sigma_{\boldsymbol{k}}\in\mathbb{F}_q\\ \boldsymbol{\chi}_{1,1}\sigma_1+\cdots+\boldsymbol{\chi}_{1,\boldsymbol{k}}\sigma_{\boldsymbol{k}}=0}} \boldsymbol{k}\mathbb{1}\{\sigma_1=\tau\}\,\hat{z}_{\boldsymbol{\chi}_{1,1},\ldots,\boldsymbol{\chi}_{1,\boldsymbol{k}}}(\sigma_1,\ldots,\sigma_{\boldsymbol{k}})\right]\quad\text{for all } \tau\in\mathbb{F}_q.$$

On an high level, (2.5) is not so different from (2.4): The first summand in (2.5) corresponds to the number of vectors with a specified number of components of each field element, taking into account the different numbers of non-zero entries of the columns. The remaining part corresponds to the probability that any such vector satisfies all equations, taking into account the number of field elements of each type in a random equation. Finally, while the frequencies of the field elements appear decoupled for rows and columns in the first line of (2.5), the second line ensures that only compatible frequencies are considered after all. As in random 3-XORSAT, a simple calculation shows that the value of (2.5) evaluated at the 'equitable' solution

$$z_i(\sigma)=q^{-1}\qquad \hat{z}_{\chi_1,\ldots,\chi_\ell}(\sigma_1,\ldots,\sigma_\ell)=q^{1-\ell}\qquad \text{for all } i,\chi_1,\ldots,\chi_\ell \tag{2.6}$$

hits the value $(1-d/k)\log q$, which matches the normalised first moment $n^{-1}\log\mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}]$.

In summary, the second moment method hardly seems like a promising path towards Theorem 1.1. Not only does (2.5) seem unwieldy as even for very special cases of $\boldsymbol{d},\boldsymbol{k}$ an analytic solution remains elusive [23]. Even worse, just in the case of 'unabridged' random k-XORSAT large deviations effects may cause spurious maxima. In effect, even if we could miraculously figure out the precise conditions for (2.5) being attained at the uniform solution, this would hardly determine for what $\boldsymbol{d},\boldsymbol{k}$ the random matrix $\mathbb{A}$ actually has full row rank w.h.p.

2.2 Quenching and truncating

The large deviations issues ultimately result from our attempt at computing the mean of $|\ker\mathbb{A}|$, a (potentially) exponential quantity. The mathematical physics prescription is to compute the expectation of its logarithm instead [40]. In the present algebraic setting this comes down to computing the mean of the nullity $\mathrm{nul}\mathbb{A}=\dim\ker\mathbb{A}$, or equivalently of the rank $\mathrm{rk}\mathbb{A}=n-\mathrm{nul}\mathbb{A}$. This 'quenched average' is always of order $O(n)$ and therefore immune to large deviations effects. In fact, even if on some unfortunate event of exponentially small probability $\exp(-\Omega(n))$ the kernel of $\mathbb{A}$ were quite large, the ensuing boost to $\mathbb{E}_{\mathfrak{A}}[\mathrm{nul}\mathbb{A}]$ remains negligible.

Yet computing the quenched average $\mathbb{E}_{\mathfrak{A}}[\mathrm{nul}\mathbb{A}]$ does not suffice to prove Theorem 1.1. Indeed, (1.4) already provides an asymptotic formula for $\mathbb{E}_{\mathfrak{A}}[\mathrm{nul}\mathbb{A}]$. But as we saw due to the normalisation on the l.h.s. (1.4) merely implies that $\mathrm{rk}\mathbb{A}=\boldsymbol{m}-o(n)$ w.h.p. To actually prove that $\mathrm{rk}\mathbb{A}=\boldsymbol{m}$ w.h.p. we will combine the quenched computation with a truncated moment argument calculation. Specifically, we will harness an enhanced version of (1.4) to prove that under the assumptions of Theorem 1.1 the only combinatorially meaningful solutions to (2.5) asymptotically coincide with the equitable solution (2.6), around which we will subsequently expand (2.5) carefully.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

To carry this programme out, let $\boldsymbol{x}_{\mathbb{A}} = (\boldsymbol{x}_{\mathbb{A},i})_{i\in[n]} \in \mathbb{F}_q^n$ be a random vector from the kernel of $\mathbb{A}$. Consider the event

$$\mathfrak{O} = \left\{ A : \sum_{\sigma,\tau\in\mathbb{F}_q} \sum_{i,j=1}^{n} \left| \mathbb{P}\left[\boldsymbol{x}_{A,i} = \sigma,\ \boldsymbol{x}_{A,j} = \tau \mid A\right] - q^{-2} \right| = o(n^2) \right\}. \tag{2.7}$$

Then by Chebyshev's inequality on $\mathfrak{O}$ w.h.p. we have

$$\sum_{i=1}^{n} \mathbb{1}\left\{\boldsymbol{d}_i = \ell,\ \boldsymbol{x}_{\mathbb{A},i} = \sigma\right\} = \mathbb{P}\left[\boldsymbol{d} = \ell\right] n/q + o(n) \qquad \text{for all } \sigma \in \mathbb{F}_q,\ \ell \in \operatorname{supp}\boldsymbol{d}.$$

Hence, on $\mathfrak{O}$ the only combinatorially relevant value of $z_\ell(\sigma)$ from (2.5) is the uniform $1/q$ for every ℓ, σ, because for every ℓ asymptotically almost all kernel vectors set about an equal number of variables of degree ℓ to each of the q possible values. Thanks to this observation will prove that w.h.p.

$$\mathbb{E}_{\mathfrak{A}}\left[\boldsymbol{Z} \cdot \mathbb{1}\left\{\mathbb{A} \in \mathfrak{O}\right\}\right] \sim \mathbb{E}_{\mathfrak{A}}\left[\boldsymbol{Z}\right] \sim q^{n-\boldsymbol{m}} \qquad \text{and} \tag{2.8}$$

$$\mathbb{E}_{\mathfrak{A}}\left[\boldsymbol{Z}^2 \cdot \mathbb{1}\left\{\mathbb{A} \in \mathfrak{O}\right\}\right] \sim \mathbb{E}_{\mathfrak{A}}\left[\boldsymbol{Z}\right]^2, \tag{2.9}$$

provided that (1.3) is satisfied. Theorem (1.1) will turn out to be an easy consequence of (2.8)–(2.9) and Corollary 1.2 of Theorem 1.1.

Thus, the challenge is to prove (2.8)–(2.9). Specifically, while the second asymptotic equality in (2.8) is easy, the proof of the first is where we require knowledge of the 'quenched average' (1.4). In fact, instead of just applying (1.4) as is we will need to perform a 'quenched' computation for a slightly enhanced random matrix from scratch. Second, the key challenge towards the proof of (2.9) is to obtain an exact asymptotic equality here, rather than the weaker estimate $\mathbb{E}_{\mathfrak{A}}\left[\boldsymbol{Z}^2 \cdot \mathbb{1}\left\{\mathbb{A} \in \mathfrak{O}\right\}\right] = O(\mathbb{E}_{\mathfrak{A}}\left[\boldsymbol{Z}\right]^2)$. This will require a meticulous expansion of the second moment around the uniform solution, which will involve the detailed analysis of the lattices generated by integer vectors that encode conceivable values of $z_i, \hat{z}_{\chi_1,\ldots,\chi_\ell}$ from (2.5).

2.3 Specified $d_1, \ldots, d_n$ and $k_1, \ldots, k_m$

Given two positive integers n and $m = m(n)$, consider now two arrays $(d_i^{(n)})_{1\le i\le n}$ and $(k_i^{(m)})_{1\le i\le m}$ of non-negative integers such that for all n,

$$\sum_{i=1}^{n} d_i^{(n)} = \sum_{j=1}^{m} k_j^{(m)}.$$

We aim to find conditions on the arrays $(d_i^{(n)})_{1\le i\le n}$ and $(k_i^{(m)})_{1\le i\le m}$ and the sequence $(m(n))_{n\ge1}$ which guarantee full row rank for the corresponding matrix in this fixed-degree setting as $n \to \infty$ and are analogous to (1.3). Throughout, we abbreviate $m(n) = m$. Let $\boldsymbol{d}_n$ denote a uniformly chosen element from the sequence $(d_i^{(n)})_{1\le i\le n}$ and $\boldsymbol{k}_n$ a uniformly chosen element from the sequence $(k_i^{(m)})_{1\le i\le m}$. Assume that $(d_i^{(n)})_{1\le i\le n}$ and $(k_i^{(m)})_{1\le i\le m}$ satisfy the following conditions in terms of the uniformly chosen degrees $\boldsymbol{d}_n$ and $\boldsymbol{k}_n$:

(P1) There exist (integer-valued) random variables $\boldsymbol{d}, \boldsymbol{k} \ge 0$ with $\mathbb{P}(\boldsymbol{d} = 0) < 1$ and $\mathbb{P}(\boldsymbol{k} \ge 3) = 1$ such that $\boldsymbol{d}_n \xrightarrow{d} \boldsymbol{d}$ and $\boldsymbol{k}_n \xrightarrow{d} \boldsymbol{k}$;

(P2) $\mathbb{E}[\boldsymbol{d}], \mathbb{E}[\boldsymbol{k}] < \infty$ and $\mathbb{E}[\boldsymbol{d}_n] \to \mathbb{E}[\boldsymbol{d}], \mathbb{E}[\boldsymbol{k}_n] \to \mathbb{E}[\boldsymbol{k}]$ as $n \to \infty$.

(P3) $\mathbb{E}[\boldsymbol{d}^2], \mathbb{E}[\boldsymbol{k}^2] < \infty$ and $\mathbb{E}[\boldsymbol{d}_n^2] \to \mathbb{E}[\boldsymbol{d}^2], \mathbb{E}[\boldsymbol{k}_n^2] \to \mathbb{E}[\boldsymbol{k}^2]$ as $n \to \infty$.

(P4) For some $\eta > 0$, $\mathbb{E}[\boldsymbol{d}^{2+\eta}] < \infty$ and $\mathbb{E}[\boldsymbol{d}_n^{2+\eta}] \to \mathbb{E}[\boldsymbol{d}^{2+\eta}]$.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

(P5) $m \sim \mathbb{E}[\boldsymbol{d}]\, n/\mathbb{E}[\boldsymbol{k}]$.

(P6) For all m and $j \in [m]$, $k_j^{(m)} \geq 3$.

(P7) For all n, $\gcd(\mathrm{supp}(\boldsymbol{d})) = \gcd(d_i^{(n)})_{1 \leq i \leq n}$.

Conditions (P1)-(P3) correspond to standard regularity conditions for the non-bipartite version of the configuration model (see Condition 7.8 in [25], for example).

Let $D(x)$ and $K(x)$ denote the probability generating functions for $\boldsymbol{d}$ and $\boldsymbol{k}$, respectively. We also abbreviate $d = \mathbb{E}[\boldsymbol{d}]$ and $k = \mathbb{E}[\boldsymbol{k}]$ as before. We may then define

$$\Phi : [0, 1] \to \mathbb{R}, \qquad z \mapsto D\left(1 - K'(z)/k\right) - \frac{d}{k}\left(1 - K(z) - (1 - z)K'(z)\right). \tag{2.10}$$

Finally, to make the difference to the i.i.d. degree case apparent, we denote the random matrix constructed by generating a uniformly random simple Tanner graph based on the fixed-degree sequences $(d_i^{(n)})_{1 \leq i \leq n}$ and $(k_i^{(m)})_{1 \leq i \leq m}$ by $\underline{\mathbb{A}}$. Again, the non-zero entries of $\underline{\mathbb{A}}$ are i.i.d. copies of the random variable $\boldsymbol{\chi}$.

Proposition 2.1. *Suppose that* $(d_i^{(n)})_{1 \leq i \leq n}$ *and* $(k_i^{(m)})_{1 \leq i \leq m}$ *satisfy* $\sum_{i=1}^n d_i^{(n)} = \sum_{j=1}^m k_j^{(n)}$ *for all* n *and properties (P1)-(P7). Let* $\mathfrak{d} = \gcd(\mathrm{supp}(\boldsymbol{d}))$. *If* q *and* $\mathfrak{d}$ *are coprime, and*

(a) $\underline{\mathbb{A}} \in \mathfrak{O}$ *w.h.p.;*

(b) $\Phi(z) < \Phi(0)$ *for all* $0 < z \leq 1$*;*

then $\underline{\mathbb{A}}$ *has full row rank over* $\mathbb{F}_q$ *w.h.p.*

Remark 2.2. As mentioned above, conditions (P1)-(P3) are natural when considering the configuration model on general specified degree sequences $(d_i^{(n)})_{1 \leq i \leq n}$ and $(k_i^{(m)})_{1 \leq i \leq m}$. In particular, these are sufficient and necessary conditions to allow translation of results from the pairing model. A more detailed discussion and references can be found below Lemma 3.6. Conditions (P4) and (P7) are needed in the proof of a local limit theorem for the random vector $\rho_{\boldsymbol{\sigma}} \in \mathbb{Z}^{\mathbb{F}_q}$, where $\rho_{\boldsymbol{\sigma}}(s) := \sum_{i=1}^n d_i \mathbb{1}\{\boldsymbol{\sigma}_i = s\}$ for a uniformly random $\boldsymbol{\sigma} \in \mathbb{F}_q^n$ and $s \in \mathbb{F}_q$. While $m = \Theta(n)$ is essential throughout the whole proof, the precise asymptotics in (P5) are only used in the final conclusion in the proof of Proposition 2.1. Finally, we chiefly employ condition (P6) in the proof of Claim 7.12.

We first prove Proposition 2.1. Then, we prove Theorem 1.1 by showing that w.h.p. $\mathbb{A} \in \mathfrak{O}$ if $\boldsymbol{m} \sim \mathrm{Po}(dn/k)$, $(\boldsymbol{d}_i)_{i \geq 1}$ and $(\boldsymbol{k}_j)_{j \geq 1}$ are i.i.d. copies of $\boldsymbol{d}$ and $\boldsymbol{k}$. Theorem 1.1 then follows from Proposition 2.1.

In the current case, $\mathfrak{A}$ simply is the σ-algebra generated by the numbers $\boldsymbol{m}(\chi_1, \ldots, \chi_\ell)$ of equations of degree $\ell \geq 3$ with coefficients $\chi_1, \ldots, \chi_\ell \in \mathbb{F}_q^*$, since all degrees are deterministic. When $\mathfrak{A}$ is used as a subscript, it serves as a notation that suppresses explicit mentioning of m, $(d_i^{(n)})_{i=1}^n$ and $(k_i^{(m)})_{1 \leq i \leq m}$. As discussed above, it suffices to prove (2.8) (with $\boldsymbol{m} = m$) and (2.9) in the more general model $\underline{\mathbb{A}}$. We observe that (2.8) follows immediately by hypothesis (a) of Proposition 2.1, as w.h.p.,

$$\mathbb{E}_{\mathfrak{A}}\left[\boldsymbol{Z} \cdot \mathbb{1}\left\{\underline{\mathbb{A}} \in \mathfrak{O}\right\}\right] \\ = \sum_{A \in \mathfrak{O}} \sum_{\sigma \in \mathbb{F}^n} \sum_{y \in \mathbb{F}^m} \mathbb{P}_{\mathfrak{A}}\left[\underline{\mathbb{A}} = A, \boldsymbol{y} = y\right] \mathbb{1}\left\{A\sigma = y\right\} = q^{n-m} \sum_{A \in \mathfrak{O}} \mathbb{P}_{\mathfrak{A}}\left[\underline{\mathbb{A}} = A\right] \sim q^{n-m}. \tag{2.11}$$

Thus, to complete the proof for Proposition 2.1 it suffices to prove (2.9).

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

2.4 The truncated first moment

We start our discussion by verifying condition (a) of Proposition 2.1 for i.i.d. $\boldsymbol{d}_i$ and $\boldsymbol{k}_j$. Hence, let us restrict to the 'i.i.d. version' of $\mathbb{A}$, that is, $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n$ and $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_{\boldsymbol{m}}$ are i.i.d. copies of $\boldsymbol{d}$ and $\boldsymbol{k}$. Although we know the approximate nullity (1.4) of $\mathbb{A}$ already, this does not suffice to actually prove that $\mathfrak{O}$ is a 'likely' event. To this end we need to study a slightly modified matrix instead. Specifically, for an integer $t \geq 0$ obtain $\mathbb{A}_{[t]}$ from $\mathbb{A}$ by adding t more rows that contain precisely three non-zero entries. The positions of these non-zero entries are chosen uniformly, mutually independently and independently of everything else, and the non-zero entries themselves are independent copies of $\boldsymbol{\chi}$. We require the following lower bound on the rank of $\mathbb{A}_{[t]}$.

Proposition 2.3. *If* (1.3) *is satisfied then there exists* $\delta_0 = \delta_0(\boldsymbol{d}, \boldsymbol{k}) > 0$ *such that for all* $0 < \delta < \delta_0$ *we have*

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{nul}\mathbb{A}_{[\lfloor \delta n\rfloor]}] \leq 1 - \frac{d}{k} - \delta. \tag{2.12}$$

The proof of Proposition 2.3 relies on the Aizenman-Sims-Starr scheme, a coupling argument inspired by spin glass theory [5]. The technique was also used in ref. [10] to prove the rank formula (1.4). While we mostly follow that proof strategy and can even reuse some of the intermediate deliberations, a subtle modification is required to accommodate the additional ternary equations. The details can be found in Section 4.

How does Proposition 2.3 facilitate the proof of (2.8)? Assuming (1.3), we obtain from (1.4) that $\mathrm{nul}\mathbb{A}/n \sim 1 - d/k$ w.h.p. Hence, (2.12) shows that nearly each one of the additional ternary rows added to $\mathbb{A}_{[\lfloor \delta n\rfloor]}$ reduces the nullity. We are going to argue that this is possible only if $\mathbb{A} \in \mathfrak{O}$ w.h.p.

To see this, let us think about the kernel of a general $M \times N$ matrix A over $\mathbb{F}_q$ for a short moment. Draw $\boldsymbol{x}_A = (\boldsymbol{x}_{A,i})_{i\in[N]} \in \ker A$ uniformly at random. For any given coordinate $\boldsymbol{x}_{A,i}$, $i \in [N]$ there are two possible scenarios: either $\boldsymbol{x}_{A,i} = 0$ deterministically, or $\boldsymbol{x}_{A,i}$ is uniformly distributed over $\mathbb{F}_q$. (This is because if we multiply $\boldsymbol{x}_A$ by a scalar $t \in \mathbb{F}_q$ we obtain $t\boldsymbol{x}_A \in \ker A$.) We therefore call coordinate i *frozen* if $x_i = 0$ for all $x \in \ker A$ and unfrozen otherwise. Let $\mathfrak{F}(A)$ be the set of frozen coordinates.

If $\mathbb{A}$ had many frozen coordinates then adding an extra random row with three non-zero entries could hardly decrease the nullity w.h.p. For if all three non-zero coordinates fall into the frozen set, then we get the new equation 'for free', that is, $\mathrm{nul}\mathbb{A}_{[1]} = \mathrm{nul}\mathbb{A}$. Thus, Proposition 2.3 implies that $|\mathfrak{F}(\mathbb{A})| = o(n)$ w.h.p. We conclude that $\boldsymbol{x}_{\mathbb{A},i}$ is uniformly distributed over $\mathbb{F}_q$ for all but $o(n)$ coordinates $i \in [n]$. However, this does not yet imply that $\boldsymbol{x}_{\mathbb{A},i}, \boldsymbol{x}_{\mathbb{A},j}$ are independent for most i, j, as required by $\mathfrak{O}$. Yet a more careful argument based on the 'pinning lemma' from [10] does. The proof of the following statement can be found in Section 5.

Proposition 2.4. *Assume that* (1.3) *is satisfied. Then w.h.p.* $\mathbb{A} \in \mathfrak{O}$.

2.5 Expansion around the equitable solution

In this part, we consider the more general model $\underline{\mathbb{A}}$, that is, the model in Proposition 2.1. As outlined earlier, given that we know (2.8), we can establish (2.9) by expanding (2.5) around the uniform distribution (2.6). At first glance, this may not seem entirely immediate because (2.8) only appears to fix the variables $(z_i(\sigma))_{i,\sigma}$ of (2.5) that correspond to the variable nodes. But thanks to a certain inherent symmetry property the optimal $\hat{z}_{\chi_1,\ldots,\chi_\ell}$ to go with the check nodes end up being nearly equitable as well. This observation by itself now suffices to show without further ado that

$$\mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}^2 \cdot \mathbb{1}\{\underline{\mathbb{A}} \in \mathfrak{O}\}] = O\left(\mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z} \cdot \mathbb{1}\{\underline{\mathbb{A}} \in \mathfrak{O}\}]^2\right). \tag{2.13}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Yet the estimate (2.13) is not quite precise enough to complete the proof of Proposition 2.1. Indeed, to apply Chebyshev's inequality we would need asymptotic equality as in (2.9) rather than just an $O(\cdot)$-bound; Huang [27] faced the same issue in the case $\boldsymbol{d}=\boldsymbol{k}$ constant and q prime. The proof of this seemingly innocuous improvement actually constitutes one of the main technical obstacles that we need to surmount.

As a first step, using a careful local expansion we will show that the dominant contribution to the second moment actually comes from $(z_\ell)_\ell$ such that

$$\sum_{\ell\geq 0}\sum_{i=1}^{n}\frac{\mathbb{1}\{d_i=\ell\}}{n}\sum_{\sigma\in\mathbb{F}_q}|z_\ell(\sigma)-q^{-1}|=O(n^{-1/2}). \tag{2.14}$$

But even once we know (2.14) a critical issue remains because we allow general distributions of degrees $d_1,\dots,d_n,k_1,\dots,k_m$ and matrix entries $\boldsymbol{\chi}$. In effect, to estimate the kernel size accurately we need to investigate the conceivable frequencies of field values that can lead to solutions. Specifically, for an integer $k_0\geq 3$ and $\chi_1,\dots,\chi_{k_0}\in\mathbb{F}_q^*$ let

$$\mathscr{S}_q(\chi_1,\dots,\chi_{k_0})=\left\{\sigma\in\mathbb{F}_q^{k_0}:\sum_{i=1}^{k_0}\chi_i\sigma_i=0\right\} \tag{2.15}$$

comprise all solutions to a linear equation with coefficients $\chi_1,\dots,\chi_{k_0}\in\mathbb{F}_q$. For each $\sigma\in\mathscr{S}_q(\chi_1,\dots,\chi_{k_0})$ the vector

$$\hat{\sigma}=\left(\sum_{i=1}^{k_0}\mathbb{1}\{\sigma_i=s\}\right)_{s\in\mathbb{F}_q^*}\in\mathbb{Z}^{\mathbb{F}_q^*} \tag{2.16}$$

tracks the frequencies with which the various non-zero field elements appear. Depending on the coefficients $\chi_1,\dots,\chi_{k_0}$, the frequency vectors $\hat{\sigma}$ may be confined to a proper sub-grid of the integer lattice. For example, in the case $q=k_0=3$ and $\chi_1=\chi_2=\chi_3=1$ they span the sub-lattice spanned by $\binom{1}{1}$ and $\binom{0}{3}$. The following proposition characterises the lattice spanned by the frequency vectors for general k_0 and $\chi_1,\dots,\chi_{k_0}$.

Proposition 2.5. *Let $k_0\geq 3$, let $\chi_1,\dots,\chi_{k_0}\in\mathbb{F}_q^*$ and let $\mathfrak{M}_q(\chi_1,\dots,\chi_{k_0})\subseteq\mathbb{Z}^{\mathbb{F}_q^*}$ be the $\mathbb{Z}$-module generated by the frequency vectors $\hat{\sigma}$ for $\sigma\in\mathscr{S}_q(\chi_1,\dots,\chi_{k_0})$. Then $\mathfrak{M}_q(\chi_1,\dots,\chi_{k_0})$ has a basis $(\mathfrak{b}_1,\dots,\mathfrak{b}_{q-1})$ of non-negative integer vectors with $\|\mathfrak{b}_i\|_1\leq 3$ for all $1\leq i\leq q-1$ such that* $\det_{\mathbb{Z}}\begin{pmatrix}\mathfrak{b}_1 & \cdots & \mathfrak{b}_{q-1}\end{pmatrix}=q^{\mathbb{1}\{\chi_1=\cdots=\chi_{k_0}\}}$.

A vital feature of Proposition 2.5 is that the module basis consists of non-negative integer vectors with small ℓ_1-norm. In effect, the basis vectors are 'combinatorially meaningful' towards our purpose of counting solutions. Perhaps surprisingly, the proof of Proposition 2.5 turns out to be rather delicate, with details depending on whether q is a prime or a prime power, among other things. The details can be found in Section 6.

In addition to the sub-grid constraints imposed by the linear equations themselves, we need to take a divisibility condition into account. Indeed, for any assignment $\sigma\in\mathbb{F}_q^n$ of values to variables, the frequencies of the various field elements $s\in\mathbb{F}_q$ are divisible by the g.c.d. $\mathfrak{d}$ of $d_1,\dots,d_n$, that is,

$$\mathfrak{d}\mid\sum_{i=1}^{n}d_i\mathbb{1}\{\sigma_i=s\}\qquad\text{for all } s\in\mathbb{F}_q. \tag{2.17}$$

To compute the expected kernel size we need to study the intersection of the sub-grid (2.17) with the grid spanned by the frequency vectors $\hat{\sigma}$ for $\sigma\in\mathscr{S}_q(\boldsymbol{\chi}_{1,1},\dots,\boldsymbol{\chi}_{1,k})$. Specifically, by way of

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

estimating the number of assignments represented by each grid point and calculating the ensuing satisfiability probability, we obtain the following.

Proposition 2.6. *Assume that q and $\mathfrak{d}$ are coprime. Then* (2.9) *holds w.h.p.*

We prove Proposition 2.6 in Section 7. Combining Propositions 2.3–2.6, we now establish the main theorems.

Proof of Proposition 2.1. Assumption (1.3) implies that $1-d/k=\Phi(0)>\Phi(1)=\mathbb{P}(\boldsymbol{d}=0)\geq 0$. Combining (2.11) and Proposition 2.6, we obtain (2.8)–(2.9) for the matrix $\underline{\mathbb{A}}$. Hence, Chebyshev's inequality and assumption (P5) imply that $\boldsymbol{Z}\geq(1-o(1))q^{n-m} = q^{n(1-d/k+o(1))}>0$ w.h.p. Consequently, the random linear system $\underline{\mathbb{A}}x=\boldsymbol{y}$ has a solution w.h.p., and thus $\mathrm{rk}\underline{\mathbb{A}}=m$ w.h.p. □

Proof of Theorem 1.1. This is now an immediate consequence of Proposition 2.4 and Proposition 2.1. □

Proof of Corollary 1.2. Let q be a prime that does not divide $\mathfrak{d}$ and let $\boldsymbol{\chi}=1$ deterministically. Obtain the matrix $\bar{\mathbb{B}}\in\mathbb{F}_q^{\boldsymbol{m}\times n}$ by reading the $\{0,1\}$-entries of $\mathbb{B}$ as elements of $\mathbb{F}_q$. Then the distribution of $\bar{\mathbb{B}}$ coincides with the distribution of the random $\mathbb{F}_q$-matrix $\mathbb{A}$. Hence, Theorem 1.1 implies that $\bar{\mathbb{B}}$ has full row rank w.h.p.

Suppose that indeed $\mathrm{rk}\bar{\mathbb{B}}=\boldsymbol{m}$. We claim that then the rows of $\mathbb{B}$ are linearly independent. Indeed, assume that $z^\top\mathbb{B}=0$ for some vector $z=(z_1,\ldots,z_{\boldsymbol{m}})^\top\in\mathbb{Z}^{\boldsymbol{m}}$. Factoring out $\gcd(z_1,\ldots,z_{\boldsymbol{m}})$ if necessary, we may assume that the vector $\bar{z}\in\mathbb{F}_q^{\boldsymbol{m}}$ with entries $\bar{z}_i=z_i+q\mathbb{Z}$ is non-zero. Since $z^\top\mathbb{B}=0$ implies that $\bar{z}^\top\bar{\mathbb{B}}=0$, the rows of $\bar{\mathbb{B}}$ are linearly dependent, in contradiction to our assumption that $\bar{\mathbb{B}}$ has full row rank. □

2.6 Discussion and related work

The present proof strategy draws on the prior work [6, 10] on the rank of random matrices. Specifically, toward the proof of Proposition 2.3 we extend the Aizenman-Sims-Starr technique from [10] and to prove Proposition 2.4 we generalise an argument from [6]. Additionally, the expansion around the centre carried out in the proof of Proposition 2.6 employs some of the techniques developed in the study of satisfiability thresholds, particularly the extensive use of local limit theorems and auxiliary probability spaces [12, 13].

The principal new proof ingredient is the asymptotically precise analysis of the second moment by means of the study of the sub-grids of the integer lattice induced by the constraints as sketched in Section 2.5. This issue was absent in the prior literature on variations on random k-XORSAT [6, 10, 15] and on other random constraint satisfaction problems [12, 13]. However, in the study of the random regular matrix from Example 1.5 Huang [27] faced a similar issue in the special case $\boldsymbol{d}=\boldsymbol{k}$ constant and $\boldsymbol{\chi}=1$ deterministically. Proposition 2.5, whose proof is based on a combinatorial investigation of lattices in the general case, constitutes a generalisation of the case Huang studied. A further feature of Proposition 2.5 absent in ref. [27] is the explicit ℓ_1-bound on the basis vectors. This bound facilitates the proof of Proposition 2.6, which ultimately carries out the expansion around the equitable solution.

Satisfiability thresholds of random constraint satisfaction problems have been studied extensively in the statistical physics literature via a non-rigorous technique called the 'cavity method'. The cavity method comes in two installments: the simpler 'replica symmetric ansatz' associated with the Belief Propagation message passing scheme, and the more intricate 'replica symmetry breaking ansatz'. The proof of Theorem 1.1 demonstrates that the former renders the correct prediction as to the satisfiability threshold of random linear equations. By contrast, in quite a few problems, notoriously random k-SAT, replica symmetry breaking occurs [14, 20].

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

An intriguing question for future work might be to understand the 'critical' case of Φ that attain their global max at 0 and another point left open by Theorem 1.1. While Example 1.4 shows that it cannot generally be true that $\mathbb{A}$ has full row rank w.h.p., the regular case where $\boldsymbol{d} = \boldsymbol{k} = d$ are fixed to the same constant provides an intriguing example. For this scenario Huang proved that the random $\{0, 1\}$-matrix $\mathbb{B}$ has full rank w.h.p. [27]. The proof, based effectively on a moment computation over finite fields and local limit techniques, also applies to the adjacency matrices of random d-regular graphs.

While the present paper deals with sparse random matrices with a bounded average number of non-zero entries in each row and column, the case of dense random matrices has received a great deal of attention, too. Komlós [34] first shows that dense square random $\{0, 1\}$-matrices are regular over the rationals w.h.p.; Vu [50] suggested an alternative proof. The computation of the exponential order of the singularity probability subsequently led to a series of intriguing articles [30, 48, 49]. By contrast, the singularity probability of a dense square matrix over a finite field converges to a value strictly between zero and one [35, 36, 38, 39].

Apart from the sparse and dense case, the regime of intermediate densities has been studied as well. Balakin [7] and Blömer, Karp and Welzl [8] dealt with the rank of such random matrices of intermediate densities over finite fields. In addition, Costello and Vu [16, 17] studied the rational rank of random symmetric matrices of an intermediate density.

Indeed, an interesting open problem appears to be the extension of the present methods to the symmetric case. In particular, it would be interesting to see if the present techniques can be used to add to the line of works on the adjacency matrices of random graphs, which have been approached by means of techniques based on local weak convergence or Littlewood-Offord techniques [9, 22]. Several core ideas of [10] have recently been used to study the asymptotic rank of a special class of symmetric random matrices [26].

2.7 Organisation

After some preliminaries in Section 3 we begin with the proof of Proposition 2.3 in Section 4. The proof relies on an Aizenman-Sims-Starr coupling argument, some details of which are deferred to Section 8. Section 5 deals with the proof of Proposition 2.4. Subsequently we prove Proposition 2.5 in Section 6, thereby laying the ground for the proof of Proposition 2.6 in Section 7.

3. Preliminaries

Unsurprisingly, the proofs of the main results involve a few concepts and ideas from linear algebra. We mostly follow the terminology from [10], summarised in the following definition.

Definition 3.1 ([10, Definition 2.1]). *Let A be an $m \times n$-matrix over a field $\mathbb{F}$.*

- *A set $\emptyset \neq I \subseteq [n]$ is a **relation** of A if there exists a row vector $y \in \mathbb{F}^{1\times m}$ such that $\emptyset \neq \mathrm{supp}(yA) \subseteq I$.*
- *If $I = \{i\}$ is a relation of A, then we call i **frozen** in A. Let $\mathfrak{F}(A)$ be the set of all frozen $i \in [n]$ and let*

$$\mathfrak{f}(A) = |\mathfrak{F}(A)|/n.$$

- *A set $I \subseteq [n]$ is a **proper relation** of A if $I \setminus \mathfrak{F}(A)$ is a relation of A.*
- *For $\delta > 0$, $\ell \geq 1$ we say that A is **(δ, ℓ)-free** if there are no more than δn^{ℓ} proper relations $I \subseteq [n]$ of size $|I| = \ell$.*

Thus, a relation is set of column indices such that the support of a non-zero linear combination yA of rows of A is contained in that set of indices. Of course, every single row induces a relation on

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

the column indices where it has non-zero entries. An important special case is a relation consisting of one coordinate i only. If such a relation exists, then $x_i = 0$ for all vectors $x \in \ker A$, which is why we call such a coordinate i frozen. Furthermore, a proper relation is a relation that is not just built up of frozen variables. Finally, we introduce the term (δ, ℓ)-free to express that A has 'relatively few' relations of size ℓ as we will generally employ this term for bounded ℓ and small $\delta > 0$.

The following observation will aid the Aizenman-Sims-Starr coupling argument, where we will need to study the effect of adding a few extra rows and columns to a random matrix.

Lemma 3.2 ([10, Lemma 2.5]). *Let A, B, C be matrices of size $m \times n$, $m' \times n$ and $m' \times n'$, respectively, and let $I \subseteq [n]$ be the set of all indices of non-zero columns of B. Moreover, obtain B_* from B by replacing for each $i \in I \cap \mathfrak{F}(A)$ the i-th column of B by zero. Unless I is a proper relation of A we have*

$$\operatorname{nul}\begin{pmatrix} A & 0 \\ B & C \end{pmatrix} - \operatorname{nul} A = n' - \operatorname{rk}(B_* \ C). \tag{3.1}$$

Apart from Lemma 3.2 we will harness an important trick called the 'pinning operation'. The key insight is that for *any* given matrix we can diminish the number of short proper relations by simply expressly freezing a few random coordinates. The basic idea behind the pinning operation goes back to the work of Montanari [43] and has been used in other contexts [11, 46]. The version of the construction that we use here goes as follows.

Definition 3.3 ([10, Definition 2.3]). *Let A be an $m \times n$ matrix and let $\theta \geq 0$ be an integer. Let $\boldsymbol{i}_1, \boldsymbol{i}_2, \ldots, \boldsymbol{i}_\theta \in [n]$ be uniformly random and mutually independent column indices. Then the matrix $A[\theta]$ is obtained by adding θ new rows to A such that for each $j \in [\theta]$ the j-th new row has precisely one non-zero entry, namely a one in the $\boldsymbol{i}_j$-th column.*

Proposition 3.4 ([10, Proposition 2.4]). *For any $\delta > 0$, $\ell > 0$ there exists $\Theta_0 = \Theta_0(\delta, \ell) > 0$ such that for all $\Theta > \Theta_0$ and for any matrix A over any field $\mathbb{F}$ the following is true. With $\boldsymbol{\theta} \in [\Theta]$ chosen uniformly at random we have $\mathbb{P}\left[A[\boldsymbol{\theta}] \text{ is } (\delta, \ell)\text{-free}\right] > 1 - \delta$.*

At first sight, it might appear surprising that Proposition 3.4 does not depend on the matrix A at all. It is here where the randomness in the number of added unit rows $\boldsymbol{\theta}$ comes into play: On a heuristic level, the proof of Proposition 3.4 is based on tracing the effect of adding unit rows over a sufficiently large number of steps. Throughout this process, irrespective of the underlying matrix A, there cannot be too many steps where the expected increase in the size of the set of frozen variables is large, since their number is trivially bounded above by n. Thus, when choosing a uniformly random number of unit rows to append, we have to be truly unlucky to hit exactly one of these few steps. On the other hand, a multitude of proper linear relations at any given point increases the chances to freeze a large number of variables upon addition of one more unit row, and therefore there also cannot be too many such moments throughout the process of adding unit rows. Of course, the precise details of the proof are more involved, and we refer the interested reader to [10].

As a fairly immediate application of Proposition 3.4 we conclude that if the pinning operation applied to a random matrix over a finite field leaves us with few frozen variables, a decorrelation condition akin to the event $\mathfrak{O}$ from (2.7) will be satisfied. For a matrix A we continue to denote by $\boldsymbol{x}_A$ a uniformly random vector from $\ker A$.

Corollary 3.5 ([6, Lemma 4.2]). *For any $\zeta > 0$ and any prime power $q > 0$ there exist $\xi > 0$ and $\Theta_0 > 0$ such that for any $\Theta > \Theta_0$ for large enough n the following is true. Let A be a $m \times n$-matrix over $\mathbb{F}_q$. Suppose that for a uniformly random $\boldsymbol{\theta} \in [\Theta]$ we have $\mathbb{E}|\mathfrak{F}(A[\boldsymbol{\theta}])| < \xi n$. Then*

$$\sum_{\sigma,\tau \in \mathbb{F}_q} \sum_{i,j=1}^{n} \mathbb{E}\left|\mathbb{P}\left[x_i = \sigma,\ x_j = \tau \mid A\right] - q^{-2}\right| < \zeta n^2.$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

As mentioned earlier, at a key junction of the moment computation we will need to estimate the number of integer lattice points that satisfy certain linear relations. The following elementary estimate will prove useful.

Lemma 3.6. [37, p. 135] *Let $\mathfrak{M} \subseteq \mathbb{R}^\ell$ be a $\mathbb{Z}$-module with basis $b_1, \ldots, b_\ell$. Then*

$$\lim_{r\to\infty} \frac{|\{x \in \mathfrak{M} : \|x\| \leq r\}|}{vol\left(\left\{x \in \mathbb{R}^\ell : \|x\| \leq r\right\}\right)} = \frac{1}{|\det(b_1 \cdots b_\ell)|}.$$

The definition of the random Tanner graph in Section 1.2.1 provides that $\mathbb{G}$ is simple. Commonly it is easier to conduct proofs for an auxiliary random multi-graph drawn from a pairing model and then lift the results to the simple random graph. This is how we proceed as well. Given (1.1) we let $\boldsymbol{G}$ be the random bipartite graph on the set $\{x_1, \ldots, x_n\}$ of variable nodes and $\{a_1, \ldots, a_{\boldsymbol{m}}\}$ of check nodes generated by drawing a perfect matching $\boldsymbol{\Gamma}$ of the complete bipartite graph on

$$\bigcup_{i=1}^{n} \{x_i\} \times [\mathbf{d}_i] \qquad \text{and} \qquad \bigcup_{i=1}^{\mathbf{m}} \{a_i\} \times [k_i]$$

and contracting the sets $x_i \times [\mathbf{d}_i]$ and $a_i \times [\boldsymbol{k_i}]$ of variable/check clones. We also let $\boldsymbol{A}$ be the random matrix to go with this random multi-graph. Hence,

$$\boldsymbol{A}_{ij} = \boldsymbol{\chi}_{i,j} \sum_{u=1}^{k_i} \sum_{v=1}^{\mathbf{d}_j} \mathbb{1}\{\{(a_i, u), (x_j, v)\} \in \boldsymbol{\Gamma}\}.$$

Similarly, given fixed-degree sequences $(d_1, \ldots, d_n)$ and $(k_1, \ldots, k_m)$ with $\sum_{i=1}^n d_i = \sum_{j=1}^m k_j$, we may define a random bipartite graph $\underline{\boldsymbol{G}}$ and the corresponding matrix $\underline{\boldsymbol{A}}$. The deviating notation only emphasises that the underlying degrees have been fixed in contrast to the i.i.d. model. Moreover, if the degree sequences $(d_1, \ldots, d_n)$ and $(k_1, \ldots, k_m)$ satisfy (P3), then routine arguments (e.g. see [29]) show that $\underline{\boldsymbol{G}}$ is simple with non-vanishing probability.

Proposition 3.7 ([25, Theorem 7.12]). *Suppose that the degree sequences $(d_1, \ldots, d_n)$ and $(k_1, \ldots, k_m)$ satisfy (P3). Then,* $\mathbb{P}\left[\underline{\boldsymbol{G}} \text{ is simple}\right] = \Omega(1)$.

If $(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n)$ and $(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_{\boldsymbol{m}})$ are i.i.d. copies of $\boldsymbol{d}$ and $\boldsymbol{k}$ with $\mathbb{E}[\boldsymbol{d}^2] + \mathbb{E}\left[\boldsymbol{k}^2\right] < \infty$ as in Section 1.2.1, then a standard Azuma-Hoeffding argument shows that w.h.p. they satisfy (P3).

Corollary 3.8 ([10, Lemma 4.3]). $\mathbb{P}\left[\boldsymbol{G} \textit{ is simple} \mid \sum_{i=1}^n \boldsymbol{d}_i = \sum_{i=1}^{\boldsymbol{m}} \boldsymbol{k}_i\right] = \Omega(1)$.

When working with the random graphs $\mathbb{G}$ or $\boldsymbol{G}$ we occasionally encounter the size-biased versions $\hat{\boldsymbol{d}}, \hat{\boldsymbol{k}}$ of the degree distributions defined by

$$\mathbb{P}\left[\hat{\boldsymbol{d}} = \ell\right] = \ell\mathbb{P}\left[\boldsymbol{d} = \ell\right]/d, \qquad \mathbb{P}\left[\hat{\boldsymbol{k}} = \ell\right] = \ell\mathbb{P}\left[\boldsymbol{k} = \ell\right]/k \qquad (\ell \geq 0). \tag{3.2}$$

In particular, these distributions occur in the Aizenman-Sims-Starr coupling argument. In that context we will also need the following crude but simple tail bound.

Lemma 3.9 ([10, Lemma 1.11]). *Let $(\boldsymbol{\lambda}_i)_{i\geq 1}$ be a sequence of independent copies of an integer-valued random variable $\boldsymbol{\lambda} \geq 0$ with $\mathbb{E}\left[\boldsymbol{\lambda}^r\right] < \infty$ for some $r > 2$. Further, let s be a sequence such that $s = \Theta(n)$. Then for all $\delta > 0$,*

$$\mathbb{P}\left[\left|\sum_{i=1}^{s} (\boldsymbol{\lambda}_i - \mathbb{E}\left[\boldsymbol{\lambda}\right])\right| > \delta n\right] = o(1/n).$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Finally, throughout the article we use the common $O(\cdot)$-notation to refer to the limit $n \to \infty$. In addition, we will sometimes need to deal with another parameter $\varepsilon > 0$. In such cases we use $O_\varepsilon(\cdot)$ and similar symbols to refer to the double limit $\varepsilon \to 0$ after $n \to \infty$.

4. Proof of Proposition 2.3

4.1 Overview

The first ingredient of the proof of Proposition 2.3 is a coupling argument inspired by the Aizenman-Sims-Starr scheme from mathematical physics [5], which also constituted the main ingredient of the proof of the approximate rank formula (1.4) from [10]. Indeed, the coupling argument here is quite similar to that from [10], with some extra bells and whistles to accommodate the additional ternary equations. We therefore defer that part of the proof to Section 8. The Aizenman-Sims-Starr argument leaves us with a variational formula for a lower bound on the rank of $\mathbb{A}_{[\lfloor \delta n \rfloor]}$. The second proof ingredient is to solve this variational problem. Harnessing the assumption (1.3), we will obtain the explicit expression for the rank provided by Proposition 2.3.

Let us come to the details. As explained in Section 3, we will have an easier time working with the pairing model versions $\boldsymbol{G}, \boldsymbol{A}$ of the Tanner graph and the random matrix. Moreover, to facilitate the coupling argument we will need to poke a few holes, known as 'cavities' in physics jargon, into the random matrix. More precisely, we will slightly reduce the number of check nodes and tolerate a small number of variable nodes x_i of degree less than $\boldsymbol{d}_i$. The cavities will provide the flexibility needed to set up the coupling argument. Finally, to be able to assume that the matrices we are dealing with are (δ, ℓ)-free with probability close to one, we also add a random, but bounded number of unary checks $p_1, \ldots, p_{\boldsymbol{\theta}}$, as described in Proposition 3.4. While this measure does not affect the asymptotic rank, quite crucially, it enables our bound on the rank difference in the coupling argument of Section 8.

Formally, let $\varepsilon, \delta \in (0, 1)$ and let $\Theta \geq 0$ be an integer. Ultimately Θ will depend on ε but not on n or δ. We then construct the random matrix $\boldsymbol{A}[n, \varepsilon, \delta, \Theta]$ as follows. Let

$$\boldsymbol{m}_\varepsilon \sim \mathrm{Po}((1-\varepsilon)dn/k), \qquad \boldsymbol{m}_\delta \sim \mathrm{Po}(\delta n), \qquad \boldsymbol{\theta} \sim \mathrm{unif}([\Theta]). \tag{4.1}$$

The Tanner multi-graph $\mathbf{G}[n, \varepsilon, \delta, \Theta]$ has variable nodes $x_1, \ldots, x_n$ and check nodes $a_1, \ldots, a_{\boldsymbol{m}_\varepsilon}, t_1, \ldots, t_{\boldsymbol{m}_\delta}, p_1, \ldots, p_{\boldsymbol{\theta}}$. To connect them draw a random maximum matching $\boldsymbol{\Gamma}[n, \varepsilon]$ of the complete bipartite graph with vertex classes

$$V_1 = \bigcup_{i=1}^{\boldsymbol{m}_\varepsilon} \{a_i\} \times [\boldsymbol{k}_i] \qquad \text{and} \qquad V_2 = \bigcup_{j=1}^{n} \left\{x_j\right\} \times [\boldsymbol{d}_j].$$

For every matching edge $\{(a_i, h), (x_j, \ell)\} \in \boldsymbol{\Gamma}[n, \varepsilon]$, $h \in [\boldsymbol{k}_i]$, $\ell \in [\boldsymbol{d}_j]$, between a clone of x_j and a clone of a_i we insert an a_i-x_j-edge into $\mathbf{G}[n, \varepsilon, \delta, \Theta]$. Moreover, the check nodes $t_1, \ldots, t_{\boldsymbol{m}_\delta}$ each independently and uniformly choose three neighbouring variables $\mathbf{i}_{i,1}, \mathbf{i}_{i,2}, \mathbf{i}_{i,3}$ with replacement among $\{x_1, \ldots, x_n\}$. Further, check node p_ℓ for $\ell \in [\boldsymbol{\theta}]$ is adjacent to x_ℓ only. Finally, to obtain the random $(\boldsymbol{\theta} + \boldsymbol{m}_\varepsilon + \boldsymbol{m}_\delta) \times n$-matrix $\boldsymbol{A}[n, \varepsilon, \delta, \Theta]$ from $\mathbf{G}[n, \varepsilon, \delta, \Theta]$ we let

$$\boldsymbol{A}[n, \varepsilon, \delta, \Theta]_{p_i, x_h} = \mathbb{1}\left\{i = h\right\} \qquad (i \in [\boldsymbol{\theta}], h \in [n]), \tag{4.2}$$

$$\boldsymbol{A}[n, \varepsilon, \delta, \Theta]_{a_i, x_h} = \boldsymbol{\chi}_{i,h} \sum_{\ell=1}^{\boldsymbol{k}_i} \sum_{s=1}^{\boldsymbol{d}_h} \mathbb{1}\{(x_h, s), (a_i, \ell)\} \in \boldsymbol{\Gamma}[n, \varepsilon]\} \qquad (i \in [\boldsymbol{m}_\varepsilon], h \in [n]), \tag{4.3}$$

$$\boldsymbol{A}[n, \varepsilon, \delta, \Theta]_{t_i, x_h} = \boldsymbol{\chi}_{\boldsymbol{m}_\varepsilon + i, h} \sum_{\ell=1}^{3} \mathbb{1}\{\mathbf{i}_{i,\ell} = h\} \qquad (i \in [\boldsymbol{m}_\delta], h \in [n]). \tag{4.4}$$

Applying the Aizenman-Sims-Starr scheme to the matrix $\boldsymbol{A}[n, \varepsilon, \delta, \Theta]$, we obtain the following variational bound.

Proposition 4.1. *There exist* $\delta_0 > 0$, $\Theta_0(\varepsilon) > 0$ *such that for all* $0 < \delta < \delta_0$ *and any* $\Theta = \Theta(\varepsilon) \geq \Theta_0(\varepsilon)$ *we have*

$$\limsup_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}\left[n, \varepsilon, \delta, \Theta\right])\right] \leq \max_{\alpha, \beta \in [0,1]} \Phi(\alpha)$$
$$+ \left(\exp\left(-3\delta\beta^2\right) - 1\right) D(1 - K'(\alpha)/k) - \delta + 3\delta\beta^2 - 2\delta\beta^3. \tag{4.5}$$

The proof of Proposition 4.1, carried out in Section 8 in detail, resembles that of the rank formula (1.4), except that we have to accommodate the additional ternary checks t_i. Their presence is the reason why the optimisation problem on the r.h.s. comes in terms of two variables α, β rather than a single variable as (1.4).

To complete the proof of Proposition 2.3 we need to solve the optimisation problem (4.5). This is the single place where we require that $\Phi(z)$ takes its unique global max at $z = 0$, which ultimately implies that the optimiser of (4.5) is $\alpha = \beta = 0$. This fact in turn implies the following.

Proposition 4.2. *For any* $\boldsymbol{d}, \boldsymbol{k}$ *that satisfy* (1.3) *there exists* $\delta_0 > 0$ *such that for all* $0 < \delta < \delta_0$ *we have*

$$\max_{\alpha, \beta \in [0,1]} \Phi(\alpha) + \left(\exp\left(-3\delta\beta^2\right) - 1\right) D(1 - K'(\alpha)/k) - \delta + 3\delta\beta^2 - 2\delta\beta^3 = 1 - \frac{d}{k} - \delta.$$

The proof of Proposition 4.2 can be found in Section 4.2. Finally, in Section 4.3 we will see that Proposition 2.3 is an easy consequence of Propositions 4.1 and 4.2.

4.2 Proof of Proposition 4.2

Let

$$\tilde{\Phi}_\delta(\alpha, \beta) = \Phi(\alpha) + \left(\exp\left(-3\delta\beta^2\right) - 1\right) D(1 - K'(\alpha)/k) - \delta + 3\delta\beta^2 - 2\delta\beta^3 \quad (\alpha, \beta \in [0, 1]).$$

Assuming (1.3), we are going to prove that for small enough δ,

$$\max_{\alpha, \beta \in [0,1]} \tilde{\Phi}_\delta(\alpha, \beta) = \tilde{\Phi}_\delta(0, 0) = 1 - \frac{d}{k} - \delta, \tag{4.6}$$

whence the assertion is immediate.

The C^1-function $\tilde{\Phi}_\delta$ attains its maximum either at a boundary point of the compact domain $[0, 1]^2$ or at a point where the partial derivatives vanish. Beginning with the former, we consider four cases.

Case 1: $\alpha = 0$ We have

$$\tilde{\Phi}_\delta(0, \beta) = \tilde{\Phi}_\delta(0, 0) + 3\delta\beta^2 - 2\delta\beta^3 - (1 - \exp\left(-3\delta\beta^2\right)). \tag{4.7}$$

Expanding the exponential function, we see that $3\delta\beta^2 - 2\delta\beta^3 - (1 - \exp\left(-3\delta\beta^2\right)) = -2\delta\beta^3 + O_\delta(\delta^2\beta^4)$. Since $-2\delta\beta^3 + O_\delta(\delta^2\beta^4)$ is non-positive for all $\beta \in [0, 1]$, (4.7) yields $\max_\beta \tilde{\Phi}_\delta(0, \beta) = \tilde{\Phi}_\delta(0, 0)$ for all small enough $\delta > 0$.

Case 2: $\beta = 0$ The assumption (1.3) ensures that Φ is maximised in 0. Therefore, as $\tilde{\Phi}_\delta(\alpha, 0) = \Phi(\alpha) - \delta$, the maximum on $\{(\alpha, 0) : \alpha \in [0, 1]\}$ is attained in $\alpha = 0$.

Case 3: $\alpha = 1$ We obtain

$$\tilde{\Phi}_\delta(1,\beta) = \Phi(1) + \left(\exp\left(-3\delta\beta^2\right) - 1\right) D(0) - \delta + 3\delta\beta^2 - 2\delta\beta^3 = \exp\left(-3\delta\beta^2\right) D(0) + \delta(3\beta^2 - 2\beta^3 - 1).$$

Again, expanding the exponential, we see that for sufficiently small δ, $\tilde{\Phi}_\delta(1,\beta) \le \tilde{\Phi}_\delta(1,0) = \Phi(1) - \delta$. Thanks to assumption (1.3), this yields $\max_\beta \tilde{\Phi}_\delta(1,\beta) = \tilde{\Phi}_\delta(0,0)$ for all small enough $\delta > 0$.

Case 4: $\beta = 1$ We have

$$\tilde{\Phi}_\delta(\alpha,1) = \Phi(\alpha) - (1 - \exp(-3\delta)\,)D\left(1 - \frac{K'(\alpha)}{k}\right). \tag{4.8}$$

Because D and K' are continuous on $[0,1]$ due to the assumption $\mathbb{E}[\boldsymbol{d}^2] + \mathbb{E}[\boldsymbol{k}^2] < \infty$, for any $\zeta > 0$ there exists $\hat{\alpha} > 0$ such that $D(1 - K'(\alpha)/k) > 1 - \zeta$ for all $0 < \alpha < \hat{\alpha}$. Therefore, (4.8) shows that for small enough $\delta > 0$ and $0 < \alpha < \hat{\alpha}$ we have $\tilde{\Phi}_\delta(\alpha,1) < \tilde{\Phi}_\delta(\alpha,0) \le \tilde{\Phi}_\delta(0,0)$. On the other hand, for $\hat{\alpha} \le \alpha \le 1$ the difference $\Phi(\alpha) - \Phi(0)$ is uniformly negative because of our assumption (1.3) that Φ attains its unique global maximum at $\alpha = 0$. Hence, for δ small enough and $\hat{\alpha} \le \alpha \le 1$ we obtain $\tilde{\Phi}_\delta(\alpha,1) < \tilde{\Phi}_\delta(0,0)$.

Combining Cases 1–4, we obtain

$$\max_{(\alpha,\beta)\in\partial[0,1]^2} \tilde{\Phi}_\delta(\alpha,\beta) = \tilde{\Phi}_\delta(0,0). \tag{4.9}$$

Moving on to the interior of $[0,1]^2$, we calculate the derivatives

$$\begin{aligned}\frac{\partial\tilde{\Phi}_\delta}{\partial\alpha} &= \Phi'(\alpha) + \left(1 - \exp\left(-3\delta\beta^2\right)\right)\frac{K''(\alpha)}{k} D'(1 - K'(\alpha)/k)\\ &= \frac{K''(\alpha)}{k}\left(d(1-\alpha) - \exp\left(-3\delta\beta^2\right) D'(1 - K'(\alpha)/k)\right),\\ \frac{\partial\tilde{\Phi}_\delta}{\partial\beta} &= 6\delta\beta\left(1 - \beta - \exp\left(-3\delta\beta^2\right) D(1 - K'(\alpha)/k)\right).\end{aligned}$$

Hence, potential maximisers (α,β) in the interior of $[0,1]^2$ satisfy

$$d(1-\alpha) = D'(1 - K'(\alpha)/k)\exp\left(-3\delta\beta^2\right) \qquad \text{and} \qquad 1 - \beta = \exp\left(-3\delta\beta^2\right) D(1 - K'(\alpha)/k). \tag{4.10}$$

Substituting (4.10) into $\tilde{\Phi}_\delta$, we obtain

$$\begin{aligned}\tilde{\Phi}_\delta(\alpha,\beta) &= \Phi(\alpha) - \delta + \left(\exp\left(-3\delta\beta^2\right) - 1\right) D(1 - K'(\alpha)/k) + 3\delta\beta^2 - 2\delta\beta^3\\ &= \Phi(\alpha) - \delta + (1-\beta)(1 - \exp\left(3\delta\beta^2\right)\,) + 3\delta\beta^2 - 2\delta\beta^3\\ &\le \Phi(\alpha) - \delta - 3\delta\beta^2(1-\beta) + 3\delta\beta^2 - 2\delta\beta^3 = \Phi(\alpha) - \delta + \delta\beta^3.\end{aligned} \tag{4.11}$$

To estimate the r.h.s. we consider the cases of small and large α separately. Specifically, by continuity for any $\zeta > 0$ there is $0 < \hat{\alpha} < \delta$ such that $D(1 - K'(\alpha)/k) > 1 - \zeta$ for all $0 < \alpha < \hat{\alpha}$.

Case 1: $0 < \alpha < \hat{\alpha}$ Since $D(1 - K'(\alpha)/k) > 1 - \zeta$, (4.10) implies that for $\beta > 0$

$$1 - \beta > (1 - 3\delta\beta^2)(1-\zeta) = 1 - \zeta - 3\delta\beta^2(1-\zeta).$$

In particular, small $\hat{\alpha}$ implies that also β is small. More precisely, after choosing δ, ζ small enough, we may assume that $\beta < \hat{\beta}$ for any fixed $\hat{\beta} > 0$. In this case, we may thus restrict to solutions $(\alpha, \beta) \in (0, 1)^2$ to (4.10) where *both* coordinates are sufficiently small. Also here, we distinguish three cases that all lead to contradictions.

(A) If the solution satisfies $\alpha = \beta$, consider the function

$$x \mapsto 1 - x - \exp\left(-3\delta x^2\right) D(1 - K'(x)/k)$$

whose zeros determine the solutions to the right equation in (4.10) under the assumption $\alpha = \beta$. Its value is zero at $x = 0$ and it has derivative

$$-1 + 6\delta x \exp\left(-3\delta x^2\right) D(1 - K'(x)/k) + \exp\left(-3\delta x^2\right) D'(1 - K'(x)/k)\frac{K''(x)}{k},$$

which is negative in a neighbourhood of $x = 0$. Thus (α, α) cannot be a solution to (4.10) for $\alpha \in (0, \hat{\alpha})$.

(B) Assume now that $\alpha < \beta$. Then the right equation of (4.10) yields

$$1 - \beta > \exp\left(-3\delta\beta^2\right) D\left(1 - K'(\beta)/k\right) > \left(1 - 3\delta\beta^2\right)\left(1 - \frac{d}{k}K'(\beta)\right).$$

Now since $\boldsymbol{k} \geq 3$, $K'(\beta) = O_\beta(\beta^2)$. But then the above equation yields a contradiction for β small enough and thus $(\alpha, \beta) \in (0, \hat{\alpha}) \times (0, \hat{\beta})$ with $\alpha < \beta$ is no possible solution.

(C) Finally, if $\alpha > \beta$, the left equation of (4.10) yields

$$d\,(1 - \alpha) > \exp\left(-3\delta\alpha^2\right) D'\left(1 - K'(\alpha)/k\right) > d\left(1 - 3\delta\alpha^2\right)\left(1 - \frac{\mathbb{E}\left[\boldsymbol{d}^2\right]}{dk}K'(\alpha)\right).$$

Now since $\boldsymbol{k} \geq 3$, $K'(\alpha) = O_\alpha(\alpha^2)$. But then the above equation yields a contradiction for α small enough and thus $(\alpha, \beta) \in (0, \hat{\alpha}) \times (0, \hat{\beta})$ with $\alpha > \beta$ is no possible solution.

Hence, (4.10) has no solution with $0 < \alpha < \hat{\alpha}$.

Case 2: $\hat{\alpha} \leq \alpha < 1$ because $\Phi(\alpha) < \Phi(0)$ for all $0 < \alpha \leq 1$, (4.11) shows that we can choose δ small enough so that $\tilde{\Phi}_\delta(\alpha, \beta) < \tilde{\Phi}_\delta(0, 0)$ for all $\alpha \geq \hat{\alpha}$ and all $\beta \in [0, 1]$.

Combining both cases and recalling (4.9), we obtain (4.6).

4.3 Proof of Proposition 2.3

Combining Propositions 4.1 and 4.2, we see that

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}\left[n, \varepsilon, \delta, \Theta\right])\right] \leq 1 - \frac{d}{k} - \delta + o_\varepsilon(1). \tag{4.12}$$

The only (small) missing piece is that we still need to extend this result to the original random matrix $\mathbb{A}_{[\lfloor\delta n\rfloor]}$ based on the simple random factor graph $\mathbb{G}$. To this end we apply the following lemma.

Lemma 4.3 ([10, Lemma 4.8]). *For any fixed* $\Theta > 0$ *there exists a coupling of* $\boldsymbol{A}$ *and* $\boldsymbol{A}\left[n, \varepsilon, 0, \Theta\right]$ *such that*

$$\mathbb{E}|nul\boldsymbol{A} - nul\boldsymbol{A}\left[n, \varepsilon, 0, \Theta\right]| = O_\varepsilon(\varepsilon n).$$

Let $\boldsymbol{A}_{[\lfloor\delta n\rfloor]}$ be the matrix obtained from $\boldsymbol{A}$ by adding $\lfloor\delta n\rfloor$ random ternary equations. Combining (4.12) with Corollary 4.3, we obtain

$$\frac{1}{n}\mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}_{[\lfloor\delta n\rfloor]})\right] \leq 1 - \frac{d}{k} - \delta + o(1). \tag{4.13}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Furthermore, since changing a single edge of the Tanner graph $\boldsymbol{G}$ or a single entry of $\boldsymbol{A}$ can change the rank by at most one, the Azuma–Hoeffding inequality shows that $\mathrm{nul}(\boldsymbol{A}_{[\lfloor \delta n \rfloor]})$ is tightly concentrated. Thus, (4.13) implies

$$\mathbb{P}\left[\frac{1}{n}\mathrm{nul}(\boldsymbol{A}_{[\lfloor \delta n \rfloor]}) \leq 1 - \frac{d}{k} - \delta + o(1)\right] = 1 - o(1/n). \tag{4.14}$$

Finally, combining (4.14) with Proposition 3.7, we conclude that

$$\mathbb{P}\left[\frac{1}{n}\mathrm{nul}(\mathbb{A}_{[\lfloor \delta n \rfloor]}) \leq 1 - \frac{d}{k} - \delta + o(1)\right] = 1 - o(1/n),$$

which implies the assertion because $\mathrm{nul}(\mathbb{A}_{[\lfloor \delta n \rfloor]}) \leq n$ deterministically.

5. Proof of Proposition 2.4

We now go on to prove that if the matrix $\mathbb{A}[\boldsymbol{\theta}_0]$ obtained from $\mathbb{A}$ by adding a few random unary checks had many frozen coordinates, then the nullity of $\mathbb{A}_{[\lfloor \delta n \rfloor]}$ would be greater than permitted by Proposition 2.3; we use an argument similar to [6, proof of Proposition 2.7]. Invoking Corollary 3.5 will then complete the proof of Proposition 2.4.

Lemma 5.1. *Assume that for some* $\Theta > 0$ *and* $\boldsymbol{\theta}_0 \sim \mathrm{unif}([\Theta])$ *we have*

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\,|\mathfrak{F}(\mathbb{A}[\boldsymbol{\theta}_0])| > 0.$$

Then for all $\delta > 0$ *we have*

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\mathrm{nul}(\mathbb{A}_{[\lfloor \delta n \rfloor]})\right] > 1 - \frac{d}{k} - \delta.$$

Proof. For an integer $\ell \geq 0$ obtain $\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0]$ from $\mathbb{A}[\boldsymbol{\theta}_0]$ by adding ℓ random ternary equations. Since $\mathrm{nul}\mathbb{A}_{[\lfloor \delta n \rfloor]} \geq \mathrm{nul}\mathbb{A}_{[\lfloor \delta n \rfloor]}[\boldsymbol{\theta}_0] \geq \mathrm{nul}\mathbb{A}_{[\lfloor \delta n \rfloor]} - \boldsymbol{\theta}_0$, for any fixed $\Theta > 0$,

$$\mathbb{E}\left|\mathrm{nul}\mathbb{A}_{[\lfloor \delta n \rfloor]}[\boldsymbol{\theta}_0] - \mathrm{nul}\mathbb{A}_{[\lfloor \delta n \rfloor]}\right| = O(1). \tag{5.1}$$

For fixed large n, we now estimate the nullity of $\mathbb{A}_{[\delta n]}[\boldsymbol{\theta}_0]$ under the assumption that

$$\mathbb{P}\left[|\mathfrak{F}(\mathbb{A}[\boldsymbol{\theta}_0])| > \zeta n\right] > \zeta \qquad \text{for some } \zeta > 0. \tag{5.2}$$

Because adding equations can only increase the set of frozen variables, we have $\mathfrak{F}(\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0]) \subseteq \mathfrak{F}(\mathbb{A}_{[\ell+1]}[\boldsymbol{\theta}_0])$ for all $\ell \geq 0$. Therefore, (5.2) implies that

$$\mathbb{P}\left[\left|\mathfrak{F}(\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0])\right| > \zeta n\right] > \zeta \qquad \text{for all } \ell \geq 0. \tag{5.3}$$

We now claim that for any $\delta > 0$

$$\frac{1}{n}\mathbb{E}\left[\mathrm{nul}\mathbb{A}_{[\delta n]}[\boldsymbol{\theta}_0]\right] \geq 1 - d/k - \delta + \delta\zeta^4 + o(1). \tag{5.4}$$

To prove (5.4) it suffices to show that for any $\ell \geq 0$,

$$\mathbb{E}\left[\mathrm{nul}\mathbb{A}_{[\ell+1]}[\boldsymbol{\theta}_0] - \mathrm{nul}\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0]\right] \geq \zeta^4 - 1. \tag{5.5}$$

Indeed, we obtain (5.4) from (5.5) and the nullity formula $n^{-1}\mathbb{E}[\mathrm{nul}\mathbb{A}_{[0]}[\boldsymbol{\theta}_0]] = n^{-1}\mathbb{E}[\mathrm{nul}\mathbb{A}] + o(1) = 1 - d/k + o(1)$ from (1.4) by writing a telescoping sum.

To establish (5.5) we observe that $\mathrm{nul}\mathbb{A}_{[\ell+1]}[\boldsymbol{\theta}_0] - \mathrm{nul}\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0] \geq -1$ because we obtain $\mathbb{A}_{[\ell+1]}[\boldsymbol{\theta}_0]$ from $\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0]$ by adding a single ternary equation. Furthermore, if $|\mathfrak{F}(\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0])| \geq \zeta n$, then with probability at least ζ^3 all three variables of the new ternary equation are frozen in $\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0]$, in which case $\mathrm{nul}\mathbb{A}_{[\ell+1]}[\boldsymbol{\theta}_0] = \mathrm{nul}\mathbb{A}_{[\ell]}[\boldsymbol{\theta}_0]$. Hence, (5.4) follows from (5.5), which follows from (5.3). Finally, combining (5.1) and (5.4) completes the proof. □

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Proof of Proposition 2.4. The proposition follows from Proposition 2.3, Corollary 3.5 and Lemma 5.1. □

6. Proof of Proposition 2.5

The proof proceeds very differently depending on whether the coefficients $\chi_1, \ldots, \chi_{k_0}$ are identical or not. The following two lemmas summarise the analyses of the two cases.

Lemma 6.1. *For any prime power q and any $\chi \in \mathbb{F}_q^*$ the $\mathbb{Z}$-module $\mathfrak{M}_q(\chi, \chi, \chi)$ possesses a basis $(\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1})$ of non-negative integer vectors $\mathfrak{b}_i \in \mathbb{Z}^{\mathbb{F}_q^*}$ for all $i \in [q-1]$ such that*

$$\|\mathfrak{b}_i\|_1 \leq 3 \quad \text{and} \quad \sum_{s \in \mathbb{F}_q^*} \mathfrak{b}_{i,s} s = 0 \quad \text{in } \mathbb{F}_q \text{ for all } i \in [q-1], \quad \text{and} \quad \det_{\mathbb{Z}} \begin{pmatrix} \mathfrak{b}_1 & \cdots & \mathfrak{b}_{q-1} \end{pmatrix} = q.$$

Furthermore, for any $k_0 > 3$ we have $\mathfrak{M}_q\underbrace{(\chi, \ldots, \chi)}_{k_0 \text{ times}} = \mathfrak{M}_q(\chi, \chi, \chi)$.

Lemma 6.2. *Suppose that q is a prime power, that $k_0 \geq 3$ and that $\chi_1, \ldots, \chi_{k_0} \in \mathbb{F}_q^*$ satisfy $|\{\chi_1, \ldots, \chi_{k_0}\}| \geq 2$. Then*

$$\mathfrak{M}_q(\chi_1, \ldots, \chi_{k_0}) = \mathbb{Z}^{\mathbb{F}_q^*}.$$

Furthermore, $\mathfrak{M}_q(\chi_1, \ldots, \chi_{k_0})$ possesses a basis $(\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1})$ of non-negative integer vectors $\mathfrak{b}_i \in \mathbb{Z}^{\mathbb{F}_q^*}$ such that

$$\|\mathfrak{b}_i\|_1 \leq 3 \quad \text{and} \quad \sum_{s \in \mathbb{F}_q^*} \mathfrak{b}_{i,s} s = 0 \quad \text{in } \mathbb{F}_q \text{ for all } i \in [q-1].$$

Clearly, Proposition 2.5 is an immediate consequence of Lemmas 6.1 and 6.2 . We proceed to prove the former in Section 6.1 and the latter in Section 6.2.

6.1 Proof of Lemma 6.1

Because we can just factor out any scalar, it suffices to consider the module

$$\mathfrak{M} = \mathfrak{M}_q\underbrace{(1, \ldots, 1)}_{k_0 \text{ times}}.$$

Being a submodule of the free $\mathbb{Z}$-module $\mathbb{Z}^{\mathbb{F}_q^*}$, $\mathfrak{M}$ is free, but it is not entirely self-evident that a basis with the additional properties stated in Lemma 6.1 exists. Indeed, while it is easy enough to come up with $q-1$ linearly independent vectors in $\mathfrak{M}$ that all have ℓ_1-norm bounded by 3, it is more difficult to show that these vectors generate $\mathfrak{M}$. In the proof of Lemma 6.1, we sidestep this difficulty by working with two sets of vectors $\mathscr{B}_1$ and $\mathscr{B}_2$. The first set $\mathscr{B}_1$ is easily seen to generate $\mathfrak{M}$, while $\mathscr{B}_2$ is a set of linearly independent vectors in $\mathfrak{M}$ with ℓ_1-norms bounded by 3. To argue that $\mathscr{B}_2$ generates $\mathfrak{M}$, too, it then suffices to show that the determinant of the change of basis matrix equals one.

To interpret the bases as subsets of $\mathbb{Z}^{q-1}$ rather than $\mathbb{Z}^{\mathbb{F}_q^*}$ in the following, we fix some notation for the elements of $\mathbb{F}_q$. Throughout this section, we let $q = p^\ell$ for a prime p and $\ell \in \mathbb{N}$. If $\ell = 1$, we regard $\mathbb{F}_q$ as the set $\{0, \ldots, p-1\}$ with $\mod p$ arithmetic. If $\ell \geq 2$, the field elements can be written as

$$\{a_0 + a_1\mathbb{X} + \ldots + a_{\ell-1}\mathbb{X}^{\ell-1} : a_j \in \mathbb{F}_p \text{ for } j = 0, \ldots, \ell - 1\},$$

with mod $g(\mathbb{X})$ arithmetic for a prime polynomial $g(\mathbb{X}) \in \mathbb{F}_p[\mathbb{X}]$ of degree ℓ. Exploiting this representation of the field elements as polynomials, we define the length $\mathrm{len}(a_0 + a_1\mathbb{X} + \ldots + a_{\ell-1}\mathbb{X}^{\ell-1})$ of an element of $\mathbb{F}_q$ to be the number of its non-zero coefficients. Finally, let

$$\mathbb{F}_q^{(\geq 2)} = \left\{h \in \mathbb{F}_q : \mathrm{len}(h) \geq 2\right\} \tag{6.1}$$

be the set of all elements of $\mathbb{F}_q$ with length at least two. Of course, if $\ell = 1$, $\mathbb{F}_q^{(\geq 2)}$ is empty.

Recall that we view $\mathfrak{M}$ as a subset of $\mathbb{Z}^{\mathbb{F}_q^*}$ that is generated by the vectors

$$\left(\sum_{i=1}^{k_0} \mathbb{1}\left\{\sigma_i = s\right\}\right)_{s\in\mathbb{F}_q^*}, \quad \sigma \in \mathscr{S}_q(1, \ldots, 1).$$

In the above representation, the generators are indexed by $\mathbb{F}_q^*$ rather than by the set $[q-1]$. But to carry out the determinant calculation, it is immensely useful to represent both $\mathscr{B}_1$ and $\mathscr{B}_2$ as matrices with a convenient structure. Hence, there is ambiguity in the choice of a bijection $f : \mathbb{F}_q^* \to \{1, \ldots, q-1\}$ that maps the non-zero elements of $\mathbb{F}_q$ to coordinates in $\mathbb{Z}^{\mathbb{F}_q^*}$. To put a clear structure to the matrices in this subsection, we will soon choose f in a particular way. With the above notation, we will from now on fix a bijection f that is monotonically decreasing with respect to the length function on $\mathbb{F}_q^*$: If $\mathrm{len}(h_1) < \mathrm{len}(h_2)$ for $h_1, h_2 \in \mathbb{F}_q^*$, then $f(h_1) > f(h_2)$. More precisely, f maps the $(p-1)^\ell$ elements in $\mathbb{F}_q^*$ of maximal length ℓ to the interval $[(p-1)^\ell]$, the $\ell(p-1)^{\ell-1}$ elements of length $\ell - 1$ to the interval $\{(p-1)^\ell + 1, \ldots, (p-1)^\ell + \ell(p-1)^{\ell-1}\}$, and so on. For elements of length one, we further specify that

$$f(a\mathbb{X}^i) = q - 1 - (\ell - i)(p-1) + a \quad \text{for } i \in \{0, \ldots \ell - 1\} \text{ and } a \in [p-1].$$

For our purposes, there is no need to fully specify the values of f within sets of constant length greater than one, but one could follow the lexicographic order, for example. The benefit of such an ordering will become apparent in the next two subsections.

6.1.1 First basis $\mathscr{B}_1$

The idea behind the first set $\mathscr{B}_1$ is that it consists of vectors whose coordinates can be easily seen to correspond to element statistics of a valid solution while ignoring the ℓ_1-restriction formulated in Lemma 6.1. We build $\mathscr{B}_1$ from frequency vectors of solutions of the form

$$\left(a_0 + a_1\mathbb{X} + \ldots + a_{\ell-1}\mathbb{X}^{\ell-1}\right) + \sum_{i=0}^{\ell-1} a_i \cdot ((p-1)\mathbb{X}^i) = 0.$$

That is, we take any element $a_0 + a_1\mathbb{X} + \ldots + a_{\ell-1}\mathbb{X}^{\ell-1}$ from $\mathbb{F}_q^*$ and cancel it by a linear combination of elements from $\{p-1, (p-1)\mathbb{X}, \ldots, (p-1)\mathbb{X}^{\ell-1}\} \subseteq \mathbb{F}_q^*$. Formally, let $e_1, \ldots, e_{q-1}$ denote the canonical basis of $\mathbb{Z}^{q-1}$. The set of statistics of all frequency vectors of the form described above then reads

$$\mathscr{B}_1 = \left\{ e_{f(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i)} + \sum_{i=0}^{\ell-1} a_i e_{f(-\mathbb{X}^i)} : \sum_{i=0}^{\ell-1} a_i\mathbb{X}^i \in \mathbb{F}_q^* \right\}.$$

A moment of thought shows that $|\mathscr{B}_1| = q - 1$. Indeed, it is helpful to notice that for any $h \in \mathbb{F}_q^* \setminus \{-1, \ldots, -\mathbb{X}^{\ell-1}\}$, there is exactly one element with a non-zero position in coordinate $f(h)$, and this coordinate is 1. That is, there is basically exactly one element in $\mathscr{B}_1$ associated with each element of $\mathbb{F}_q^*$. Generally, the elements of $\mathscr{B}_1$ can be ordered to yield a lower triangular matrix M_q. To sketch this matrix, we first consider the case $\ell = 1$. In this case, with our choice of indexing

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

$$M_p = \begin{array}{c} \\ 1 \\ 2 \\ \\ \\ \vdots \\ \\ \\ p-1 \end{array}\!\!\begin{array}{c} \begin{array}{cccccccc} 1 & 2 & & . & . & . & . & p-1 \end{array} \\ \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ 1 & 2 & 3 & \cdots & & p-2 & p \end{pmatrix} \end{array}.$$

Figure 4. The matrix M_p.

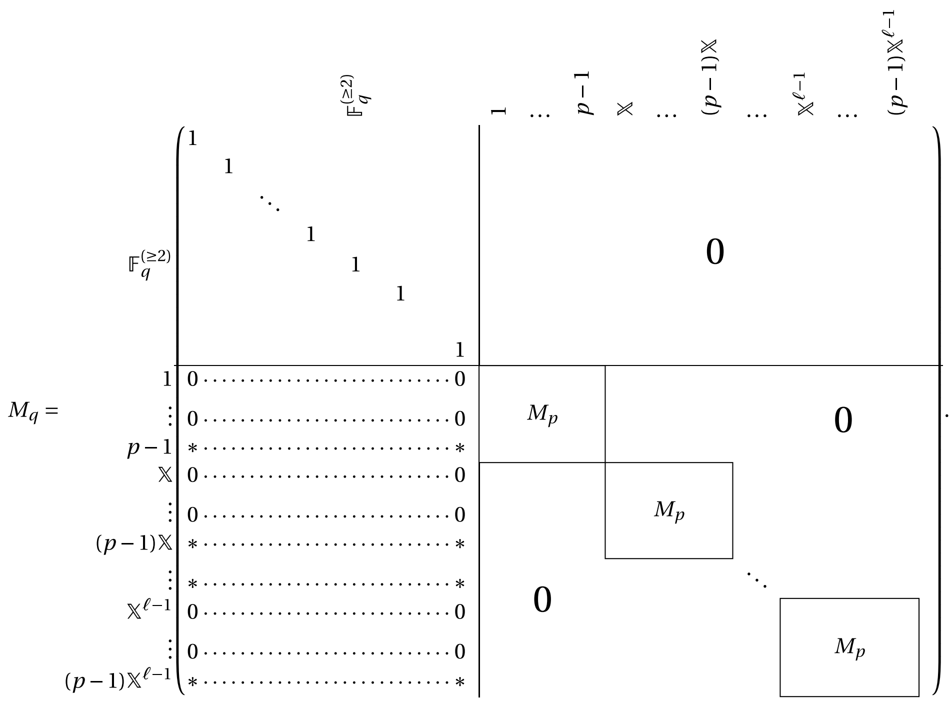

Figure 5. The matrix M_q for $\ell \geq 2$.

function f, the elements of $\boldsymbol{B}_1$ can be ordered to give the matrix displayed in Fig. 4. For the case of fields of prime order, this basis is already implicitly mentioned in [27].

Note that this reduces to $M_2 = (2)$ in the case $p = 2$. In this representation, rows are indexed by the field elements they represent, while columns are indexed by the field elements they are associated with. For $\ell \geq 2$, we can use the matrix M_p for the compact representation of M_q displayed in Fig. 5.

In the matrix M_q, the upper left block is an identity matrix of the appropriate dimension, the upper right is a zero matrix, the lower left is a matrix that only has non-zero entries in rows that correspond to $-1, \ldots, -\mathbb{X}^{\ell-1}$ while the lower right is a block diagonal matrix whose blocks are

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

given by M_p. In particular, M_p is a lower triangular matrix. Because M_p has determinant p the following is immediate.

Claim 6.3. *We have* $\det(M_q) = p^\ell = q$.

Let $\mathfrak{B}_1$ denote the $\mathbb{Z}$-module generated by the elements of $\mathscr{B}_1$. Then the lower triangular structure of M_q also implies the following.

Claim 6.4. *The rank of* $\mathscr{B}_1$ *is* $q-1$.

The following lemma shows that the module $\mathfrak{M}$ is contained in $\mathfrak{B}_1$.

Lemma 6.5. *The* $\mathbb{Z}$*-module* $\mathfrak{M}$ *is contained in the* $\mathbb{Z}$*-module* $\mathfrak{B}_1$.

Proof. We show that each element of $\mathfrak{M}$ can be written as a linear combination of elements of $\mathscr{B}_1$. To this end it is sufficient to show that every frequency vector of a solution to an equation with exactly k_0 non-zero entries and all-one coefficients can be written as a linear combination of the elements of $\mathscr{B}_1$. Let thus $x \in \mathbb{N}^{q-1}$ be such a frequency vector, that is, $\sum_{i=1}^{q-1} x_i f^{-1}(i) = 0$ in $\mathbb{F}_q$. Before we state a linear combination of x in terms of $\mathscr{B}_1$, observe that for each $j \in [q-1] \setminus \{q-1-(\ell-1)(p-1), q-1-(\ell-2)(p-1), \ldots, q-1\}$, there is exactly one basis vector with a non-zero entry in position j. Moreover, the entry of this basis vector in position j is 1. On the other hand, the basis vectors corresponding to the remaining ℓ columns $q-1-(\ell-1)(p-1), q-1-(\ell-2)(p-1), \ldots, q-1$ of M_q are actually integer multiples of the standard unit vectors, as

$$e_{f((p-1)\mathbb{X}^i)} + (p-1)e_{f(-\mathbb{X}^i)} = pe_{f((p-1)\mathbb{X}^i)}$$

for $i = 0, \ldots, \ell-1$. With these observations, the only valid candidate for a linear combination of x in terms of the elements of $\mathscr{B}_1$ is given by

$$x = \sum_{\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i \in \mathbb{F}_q^* \setminus \{-1,\ldots,-\mathbb{X}^{\ell-1}\}} x_{f(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i)} \left(e_{f(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i)} + \sum_{j=0}^{\ell-1} a_j e_{f(-\mathbb{X}^j)} \right)$$
$$+ \sum_{j=0}^{\ell-1} \frac{x_{f(-\mathbb{X}^j)} - \sum_{\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i \in \mathbb{F}_q^* \setminus \{-1,\ldots,-\mathbb{X}^{\ell-1}\}} a_j x_{f(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i)}}{p} \cdot pe_{f(-\mathbb{X}^j)}.$$

It remains to argue why the coefficients of the basis vectors $pe_{f(-1)}, \ldots, pe_{f(-\mathbb{X}^{\ell-1})}$ in the second sum are integers. At this point, we will use that x is a solution statistic: Because

$$\sum_{\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i \in \mathbb{F}_q^*} x_{f(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i)} \sum_{j=0}^{\ell-1} a_j\mathbb{X}^j = 0 \qquad \text{in} \quad \mathbb{F}_q$$

and the additive group $(\mathbb{F}_q, +)$ is isomorphic to $((\mathbb{F}_p)^\ell, +)$, all 'components' in the above sum must be zero and thus

$$\sum_{\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i \in \mathbb{F}_q^*} x_{f(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i)} a_j = 0 \qquad \text{in} \quad \mathbb{F}_p$$

for all $j = 0, \ldots, \ell-1$. However, isolating the contribution from $\{-1, \ldots, -\mathbb{X}^{\ell-1}\}$ yields

$$0 = \sum_{\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i \in \mathbb{F}_q^*} x_{f(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i)} a_j = -x_{f(-\mathbb{X}^j)} + \sum_{\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i \in \mathbb{F}_q^* \setminus \{-1,\ldots,-\mathbb{X}^{\ell-1}\}} a_j x_{f(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i)} \qquad \text{in} \quad \mathbb{F}_p, \tag{6.3}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

as the coefficient a_j of $\mathbb{X}^j$ in $-\mathbb{X}^i$ is zero unless $i=j$. Therefore, the right-hand side in (6.3) is divisible by p and the claim follows. □

6.1.2 Second basis $\mathscr{B}_2$

In this subsection, we define a candidate set for the vectors $(\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1})$ in the statement of Lemma 6.1. That is, we define a set $\mathscr{B}_2$ all whose elements have non-negative components and ℓ_1-norm at most three. In other words, we are looking for solutions to

$$x_1 + \ldots + x_{k_0} = 0 \tag{6.4}$$

with at most three different non-zero components.

Here again, our construction basically associates one basis vector to each element of $\mathbb{F}_q^*$. However, due to the ℓ_1-restriction, there is less freedom in choosing the remaining non-zero coordinates. Our approach to design a set that satisfies this restriction while retaining a representation in a convenient block lower triangular matrix structure is to distinguish between elements of length one and of length at least two. We will therefore construct $\mathscr{B}_2$ via two sets $\mathscr{B}^{(1)}$ and $\mathscr{B}^{(\geq 2)}$ such that $\mathscr{B}_2$ is given as

$$\mathscr{B}_2 = \mathscr{B}^{(1)} \cup \mathscr{B}^{(\geq 2)}. \tag{6.5}$$

Let us start with an element $h = \sum_{i=0}^{\ell-1} a_i \mathbb{X}^i$ of length at least two in $\mathbb{F}_q$. Assume that its leading coefficient is a_r for $r \in [\ell - 1]$. If a variable in (6.4) takes value h, we may cancel its contribution to the equation by subtracting the two elements $a_r\mathbb{X}^r$ and $h - a_r\mathbb{X}^r$, both of which are shorter than h:

$$\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i - a_r\mathbb{X}^r - \left(\sum_{i=0}^{\ell-1} a_i\mathbb{X}^i - a_r\mathbb{X}^r\right) = 0.$$

This solution corresponds to the vector

$$e_{f(h)} + e_{f(-a_r\mathbb{X}^r)} + e_{f(-h+a_r\mathbb{X}^r)}.$$

This idea for field elements $h \in \mathbb{F}_q^{(\geq 2)}$ of length at least two then yields the $q - 1 - \ell(p-1)$ integer vectors

$$\mathscr{B}^{(\geq 2)} = \left\{ e_{f(h)} + e_{f(-a_r\mathbb{X}^r)} + e_{f(-h+a_r\mathbb{X}^r)} : r \in [\ell - 1] \text{ and } h = \sum_{i=0}^{r} a_i\mathbb{X}^i \in \mathbb{F}_q^{(\geq 2)} \text{ with } a_r \neq 0 \right\}.$$

For a field element h of length one, an analogous shortening operation would correspond to the vector

$$e_{f(h)} + e_{f(-h)}.$$

If $p = 2$, this procedure applied to all field elements of length one yields ℓ distinct vectors and we are done. However, if $p > 2$, employing this idea for all elements of length one would only lead to $\ell(p-1)/2$ rather than $\ell(p-1)$ additional vectors, as h and $-h$ are distinct and obviously give rise to the same statistic. As a consequence, for $p > 2$, we need to deviate from the above construction and come up with a modified 'short-solution' scheme. Let $h = a_r\mathbb{X}^r$ be an element of length one. If $a_r \in \{1, \ldots (p-1)/2\}$, we simply associate the vector $e_{f(h)} + e_{f(-h)}$ to it, as indicated. If on the other hand $a_r \in \{(p+1)/2, \ldots, p-1\}$, we let h correspond to the vector

$$e_{f(h)} + e_{f(-\mathbb{X}^r)} + e_{f(-h+\mathbb{X}^r)}.$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

$$A_p = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots & \vdots & & & \iddots & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & & \iddots & \iddots & 0 \\ \vdots & & \ddots & \ddots & 0 & \vdots & \iddots & \iddots & \iddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 1 & 2 & 0 & \cdots & \cdots & 0 \\ \vdots & & \iddots & \iddots & 0 & 0 & 1 & \ddots & & \vdots \\ \vdots & \iddots & \iddots & \iddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \iddots & \iddots & & \vdots & 0 & \cdots & 0 & 1 & 0 \\ 1 & 0 & \cdots & \cdots & 0 & 1 & \cdots & \cdots & 1 & 2 \end{pmatrix}.$$

Figure 6. The matrix A_p.

With this, for $p > 2$, the part of $\mathscr{B}_2$ that corresponds to field elements of length one is given by the set

$$\mathscr{B}^{(1)} = \bigcup_{r=0}^{\ell-1} \Big(\big\{ e_{f(a_r \mathbb{X}^r)} + e_{f(-a_r \mathbb{X}^r)} : a_r \in [(p-1)/2] \big\} \\ \cup \big\{ e_{f(-a_r \mathbb{X}^r)} + e_{f(-\mathbb{X}^r)} + e_{f(a_r \mathbb{X}^r + \mathbb{X}^r)} : a_r \in [(p-1)/2] \big\} \Big). \tag{6.6}$$

If $p = 2$, in line with the above discussion, we simply let

$$\mathscr{B}^{(1)} = \bigcup_{r=0}^{\ell-1} \big\{ 2 e_{f(\mathbb{X}^r)} \big\}. \tag{6.7}$$

Again, a moment of thought shows that in any case, $|\mathscr{B}_2| = |\mathscr{B}_1| = q - 1$. Let $\mathfrak{B}_2$ denote the $\mathbb{Z}$-module generated by the elements of $\mathscr{B}_2$. Our choice of $\mathscr{B}_2$ has the advantage that again, its elements may be represented in a block lower triangular matrix. For this representation, it is instructive to consider the case $\ell = 1$ first. In this case and with our choice of f, the elements of $\mathscr{B}_2$ can be arranged as the columns of a matrix A_p as in Fig. 6.

Here, as in the construction of M_p, column i corresponds to the unique vector associated to $i \in \mathbb{F}_q$. In the special case $p = 2$, this matrix reduces to

$$A_2 = (2).$$

For $\ell \geq 2$, the elements of $\mathscr{B}_2$ may then be visualised in the matrix from Fig. 7.

In A_q, column $i \in [q-1]$ corresponds to the unique vector that is associated with the field element $f^{-1}(i)$. Moreover, at this point, a moment of appreciation of our indexing choice f is in place: Because f is monotonically decreasing with respect to length, there are no entries above the diagonal in the first $|\mathbb{F}_q^{(\geq 2)}|$ columns, as we only cancel field elements by strictly shorter ones. Moreover, the remaining $\ell(p-1)$ columns are governed by a simple block structure. As a concrete example, (6.8) with $p = 7$ reads

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

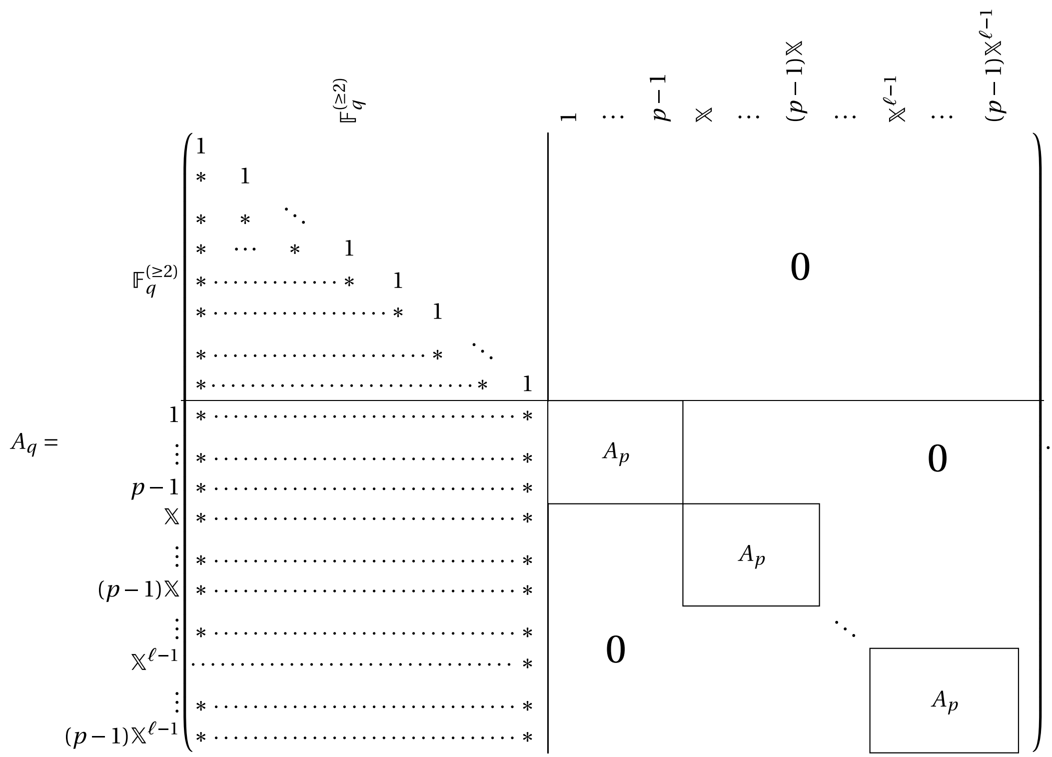

Figure 7. The matrix A_q for $\ell \geq 2$.

$$A_7 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 2 \end{pmatrix}$$

and A_7 would be used as a block matrix in any field of order 7^ℓ as shown in (6.9).

As each element of $\mathscr{B}_2$ corresponds to a solution with at most $3 \leq k_0$ non-zero components, we obtain the following.

Claim 6.6. *The $\mathbb{Z}$-module $\mathfrak{B}_2$ is contained in the $\mathbb{Z}$-module $\mathfrak{M}$.*

Thus far we know $\mathfrak{B}_2 \subseteq \mathfrak{M} \subseteq \mathfrak{B}_1$. Moreover, $\mathscr{B}_2$ has the desired ℓ_1-property. On the other hand, in comparison to $\mathscr{B}_1$, it is less clear that $\mathscr{B}_2$ generates $\mathfrak{M}$. It thus remains to show that in fact $\mathfrak{B}_2 = \mathfrak{B}_1$. We will do so by using the following fact, which is an immediate consequence of the adjugate matrix representation of the inverse matrix.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Fact 6.7. *If M is a free $\mathbb{Z}$-module with basis $x_1, \ldots, x_n$, a set of elements $y_1, \ldots, y_n \in M$ is a basis of M if and only if the change of basis matrix (c_{ij}) has determinant ± 1.*

We will apply Fact 6.7 to $M = \mathfrak{B}_1$ with $\{x_1, \ldots, x_n\} = \mathscr{B}_1$ and $\{y_1, \ldots, y_n\} = \mathscr{B}_2$. Let $C_q \in \mathbb{Z}^{(q-1)\times(q-1)}$ be the matrix whose entries comprise the coefficients when we express the elements of $\mathscr{B}_2$ by $\mathscr{B}_1$ (recall that $\mathfrak{B}_2 \subseteq \mathfrak{B}_1$) when we order the elements of $\mathscr{B}_1, \mathscr{B}_2$ as done in the construction of M_q and A_q. Thus $A_q = M_q C_q$. As

$$\det(A_q) = \det(M_q C_q) = \det(M_q) \cdot \det(C_q),$$

we do not need to compute C_q explicitly to apply Fact 6.7, but instead it suffices to compute $\det(M_q)$ and $\det(A_q)$. From Claim 6.3, $\det(M_q)$ is already known. Moreover, for A_q, the computation will not be too hard, as A_q is a block lower triangular matrix. Therefore, we are just left to calculate the determinant of the non-trivial diagonal blocks.

Lemma 6.8. *For any prime p we have* $\det(A_p) = p$.

Proof. The case $p = 2$ is immediate. We thus assume that $p > 2$ in the following. We transform A_p into a lower triangular matrix by elementary column operations. To this end, let $a_1, \ldots, a_{q-1}$ be the columns of A_p. The first $(p+1)/2$ columns already have the right form, so we do not alter this part of the matrix. For any $j = (p+3)/2, \ldots, p-1$, subtract column a_{p+1-j} from column a_j. This yields the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & \iddots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & -1 & 0 & 0 \\ 0 & 0 & \iddots & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \ddots & -1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 2 \end{pmatrix}.$$

Next, we swap column $(p+1)/2$ successively with columns $(p+3)/2, \ldots$ up to $p-1$, yielding

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & \iddots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 2 \\ 0 & 0 & \iddots & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 2 & 1 \end{pmatrix}.$$

This changes the determinant by a factor of $(-1)^{(p-3)/2}$. Finally, in order to erase the entry 2 in row $(p+1)/2$ and column $p-1$, we add twice the sum of columns $(p+1)/2, \ldots, p-2$ to column $p-1$. We thus obtain the matrix

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & \iddots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & \iddots & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 2 & p \end{pmatrix}.$$

with determinant $(-1)^{(p-3)/2}p$. Multiplying with $(-1)^{(p-3)/2}$ from the column swaps yields the claim. □

Corollary 6.9. *For any prime p and $\ell \geq 1$, we have* $\det(A_q) = q$.

Finally, Claim 6.3 and Corollary 6.9 imply that $\det(C_q) = 1$. Thus, by Fact 6.7, $\mathscr{B}_2$ is a basis of $\mathfrak{B}_1$, which implies that $\mathfrak{B}_1 = \mathfrak{B}_2 = \mathfrak{M}$. The column vectors $\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1}$ of A_q therefore enjoy the properties stated in Lemma 6.1.

6.2 Proof of Lemma 6.2

Assume w.l.o.g. that $\chi_1 = 1$. Moreover, by assumption, the set $\{\chi_1, \ldots, \chi_{k_0}\}$ contains at least two different elements, and so we may also assume that $\chi_3 \neq 1$ (recall that $k_0 \geq 3$).

We define $(\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1})$ by distinguishing between three cases:

Case 1: $p = 2$ and $\chi_2 = 1$.

Denote the order of χ_3^{-1} in $(\mathbb{F}_q^*, \cdot)$ by $\mathfrak{o}$, so that the elements $1, \chi_3^{-1}, \ldots, \chi_3^{-(\mathfrak{o}-1)}$ are pairwise distinct. Since $p = 2$ and $\mathfrak{o} \mid q-1$, $\mathfrak{o}$ is an odd number. Moreover, because $\chi_3^{-1} \neq 1$, $\mathfrak{o} \geq 3$. We now partition $\mathbb{F}_q^*$ into orbits of the action of $(\{1, \chi_3^{-1}, \ldots, \chi_3^{-(\mathfrak{o}-1)}\}, \cdot)$ on $\mathbb{F}_q^*$ such that

$$\mathbb{F}_q^* = \dot{\bigcup}_{j=1}^{(q-1)/\mathfrak{o}} \mathfrak{O}_j,$$

where each orbit $\mathfrak{O}_j$ contains exactly $\mathfrak{o}$ elements. Suppose that $\mathscr{O}_j = \{g_1^{(j)}, \ldots, g_{\mathfrak{o}}^{(j)}\}$, where the elements are indexed such that $g_{i+1}^{(j)} = \chi_3^{-1} g_i^{(j)}$.

To each $\mathfrak{O}_j$, we associate a set of potential basis vectors whose union over different j then yields the full set $(\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1})$. More precisely, the set corresponding to $\mathfrak{O}_j$ is defined as

$$\mathscr{B}_j = \bigcup_{i=1}^{\mathfrak{o}-1} \left\{ e_{g_i^{(j)}} + e_{g_{i+1}^{(j)}} \right\} \cup \left\{ e_{g_1^{(j)}} + e_{g_2^{(j)}} + e_{g_2^{(j)} + g_3^{(j)}} \right\}.$$

In this definition, we have used that for $\chi_1 = -\chi_2 = 1$ and any $h \in \mathbb{F}_q$,

$$\chi_1 \cdot h + \chi_2 \cdot 0 + \chi_3 \cdot \chi_3^{-1} h = 0 \qquad \text{as well as} \qquad \chi_1 \cdot h + \chi_2 \cdot \chi_3^{-1} h + \chi_3 \cdot (\chi_3^{-1} h + \chi_3^{-2} h) = 0.$$

Note that the element

$$g_2^{(j)} + g_3^{(j)} = (1 + \chi_3^{-1}) g_2^{(j)}$$

is non-zero and distinct from both $g_2^{(j)}$ and $g_3^{(j)}$. It might be one of $g_1^{(j)}, g_4^{(j)}, \ldots, g_{\mathfrak{o}}^{(j)}$.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

We next argue that the union of the different $\mathscr{B}_j$ generates $\mathbb{Z}^{\mathbb{F}_q^*}$. By linear transformation and using that $\mathfrak{o}$ is odd, $\mathscr{B}_j$ has the same span as

$$\left\{e_{g_1^{(j)}} + e_{g_2^{(j)}}, e_{g_1^{(j)}} - e_{g_3^{(j)}}, e_{g_1^{(j)}} + e_{g_4^{(j)}}, \ldots, e_{g_1^{(j)}} - e_{g_{\mathfrak{o}}^{(j)}}\right\} \cup \left\{e_{g_1^{(j)}+g_2^{(j)}}\right\}.$$

Now, there are two cases.

1. For all $j \in [(q-1)/\mathfrak{o}]$, $g_2^{(j)} + g_3^{(j)} \in \{g_1^{(j)}, g_4^{(j)}, \ldots, g_{\mathfrak{o}}^{(j)}\}$. In this case, either $e_{g_2^{(j)}+g_3^{(j)}} = e_{g_1^{(j)}}$, or we can subtract $e_{g_2^{(j)}+g_3^{(j)}}$ from or add it to the element $e_{g_1^{(j)}} \pm e_{g_2^{(j)}+g_3^{(j)}}$ to obtain $e_{g_1^{(j)}}$. After isolating $e_{g_1^{(j)}}$, a straightforward linear transformation yields a set of $\mathfrak{o}$ distinct unit vectors whose non-zero components are given by $\mathfrak{O}_j$. Thus, the union over all $\mathscr{B}_j$ constitutes a set of linearly independent elements that generates $\mathbb{Z}^{\mathbb{F}_q^*}$.
2. For all $j \in [(q-1)/\mathfrak{o}]$, $g_2^{(j)} + g_3^{(j)} \notin \{g_1^{(j)}, g_4^{(j)}, \ldots, g_{\mathfrak{o}}^{(j)}\}$. In this case, consider the union $\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \mathscr{B}_j$, which has the same span as

$$\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \left\{e_{g_1^{(j)}} + e_{g_2^{(j)}}, e_{g_1^{(j)}} - e_{g_3^{(j)}}, e_{g_1^{(j)}} + e_{g_4^{(j)}}, \ldots, e_{g_1^{(j)}} - e_{g_{\mathfrak{o}}^{(j)}}\right\} \cup \left\{e_{g_1^{(j)}+g_2^{(j)}}\right\}.$$

   Since for each j, the element $g_1^{(j)} + g_2^{(j)}$ must be contained in some $\mathfrak{O}_{j'}$ for $j \neq j'$, as in case (1), $e_{g_1^{(j)}+g_2^{(j)}}$ can be used to isolate $e_{g_1^{(j')}}$. After isolating $e_{g_1^{(j')}}$ for all j', these elements can be straightforwardly used to linearly transform the union over all $\mathscr{B}_j$ into the standard basis $(e_h)_{h\in\mathbb{F}_q^*}$ of $\mathbb{Z}^{\mathbb{F}_q^*}$.

Finally, set $\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \mathscr{B}_j = \{\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1}\}$.

Case 2: $p \neq 2$ and $\chi_2 = -1$.

We proceed almost exactly as before, only the choice of the 'cyclic' basis vectors is different: Denote the order of χ_3^{-1} in $(\mathbb{F}_q^*, \cdot)$ by $\mathfrak{o}$, so that the elements $1, \chi_3^{-1}, \ldots, \chi_3^{-(\mathfrak{o}-1)}$ are pairwise distinct. Then $\mathfrak{o} \mid q-1$, and since $\chi_3^{-1} \neq 1$, $\mathfrak{o} \geq 2$. We now partition $\mathbb{F}_q^*$ into orbits of the action of $(\{1, \chi_3^{-1}, \ldots, \chi_3^{-(\mathfrak{o}-1)}\}, \cdot)$ on $\mathbb{F}_q^*$ such that

$$\mathbb{F}_q^* = \dot{\bigcup}_{j=1}^{(q-1)/\mathfrak{o}} \mathfrak{O}_j,$$

where each orbit $\mathfrak{O}_j$ contains exactly $\mathfrak{o}$ elements. Suppose that $\mathfrak{O}_j = \{g_1^{(j)}, \ldots, g_{\mathfrak{o}}^{(j)}\}$, where the elements are indexed such that $g_{i+1}^{(j)} = \chi_3^{-1} g_i^{(j)}$.

To each $\mathfrak{O}_j$, we associate a set of potential basis vectors whose union over different j then yields the full set $(\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1})$. More precisely, the set corresponding to $\mathfrak{O}_j$ is defined as

$$\mathscr{B}_j = \bigcup_{i=1}^{\mathfrak{o}-1} \left\{e_{g_i^{(j)}} + e_{g_{i+1}^{(j)}}\right\} \cup \left\{e_{g_1^{(j)}} + e_{g_2^{(j)}} + e_{2g_1^{(j)}}\right\}.$$

Here, we have used that for $\chi_1 = -\chi_2 = 1$ and $p \neq 2$,

$$\chi_1 \cdot 0 + \chi_2 \cdot h + \chi_3 \cdot \chi_3^{-1} h = 0 \qquad \text{and} \qquad \chi_1 \cdot h + \chi_2 \cdot 2h + \chi_3 \cdot \chi_3^{-1} h = 0.$$

Note that the element $2g_1^{(j)}$ is distinct from $g_1^{(j)}$. It might be one of $g_2^{(j)}, \ldots, g_{\mathfrak{o}}^{(j)}$.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

We next argue that the union of the different $\mathscr{B}_j$ generates $\mathbb{Z}^{\mathbb{F}_q^*}$. By linear transformation, $\mathscr{B}_j$ has the same span as

$$\left\{e_{g_1^{(j)}} + e_{g_2^{(j)}}, e_{g_1^{(j)}} - e_{g_3^{(j)}}, e_{g_1^{(j)}} + e_{g_4^{(j)}}, \dots, e_{g_1^{(j)}} \pm e_{g_\mathfrak{o}^{(j)}}\right\} \cup \left\{e_{2g_1^{(j)}}\right\}.$$

Now, there are two cases.

1. For all $j \in [(q-1)/\mathfrak{o}]$, $2g_1^{(j)} \in \{g_2^{(j)}, \dots, g_\mathfrak{o}^{(j)}\}$. As in case 1, we can then subtract $e_{2g_2^{(j)}}$ from or add it to $e_{g_1^{(j)}} \pm e_{2g_2^{(j)}}$ to isolate $e_{g_1^{(j)}}$. After isolating $e_{g_1^{(j)}}$, a straightforward linear transformation yields a set of $\mathfrak{o}$ distinct unit vectors whose non-zero components are given by $\mathfrak{O}_j$. Thus, the union over all $\mathscr{B}_j$ constitutes a set of linearly independent elements that generates $\mathbb{Z}^{\mathbb{F}_q^*}$.
2. For all $j \in [(q-1)/\mathfrak{o}]$, $2g_1^{(j)} \notin \{g_2^{(j)}, \dots, g_\mathfrak{o}^{(j)}\}$. In this case, consider the union $\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \mathscr{B}_j$, which has the same span as

$$\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \left\{e_{g_1^{(j)}} + e_{g_2^{(j)}}, e_{g_1^{(j)}} - e_{g_3^{(j)}}, e_{g_1^{(j)}} + e_{g_4^{(j)}}, \dots, e_{g_1^{(j)}} \pm e_{g_\mathfrak{o}^{(j)}}\right\} \cup \left\{e_{2g_1^{(j)}}\right\}.$$

   Since for each j, the element $2g_1^{(j)}$ must be contained in some $\mathfrak{O}_{j'}$ for $j \neq j'$, as in case 1, $e_{2g_1^{(j)}}$ can be used to isolate $e_{g_1^{(j')}}$. After isolating $e_{g_1^{(j')}}$ for all j', these elements can be straightforwardly used to linearly transform the union over all $\mathscr{B}_j$ into the standard basis $(e_h)_{h\in\mathbb{F}_q^*}$ of $\mathbb{Z}^{\mathbb{F}_q^*}$.

In any case, set $\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \mathscr{B}_j = \{\mathfrak{b}_1, \dots, \mathfrak{b}_{q-1}\}$.

Case 3: $\chi_2 \neq -1$.

Denote the order of $-\chi_2^{-1}$ in $(\mathbb{F}_q^*, \cdot)$ by $\mathfrak{o}$, so that the elements $1, -\chi_2^{-1}, \dots, (-\chi_2^{-1})^{\mathfrak{o}-1}$ are pairwise distinct. Then $\mathfrak{o} \mid q-1$, and since $-\chi_2^{-1} \neq 1$, $\mathfrak{o} \geq 2$. We now partition $\mathbb{F}_q^*$ into orbits of the action of $(\{1, -\chi_2^{-1}, \dots, (-\chi_2^{-1})^{\mathfrak{o}-1}\}, \cdot)$ on $\mathbb{F}_q^*$ such that

$$\mathbb{F}_q^* = \dot{\bigcup}_{j=1}^{(q-1)/\mathfrak{o}} \mathfrak{O}_j,$$

where each orbit $\mathfrak{O}_j$ contains exactly $\mathfrak{o}$ elements. Suppose that $\mathfrak{O}_j = \{g_1^{(j)}, \dots, g_\mathfrak{o}^{(j)}\}$, where the elements are indexed such that $g_{i+1}^{(j)} = -\chi_2^{-1} g_i^{(j)}$.

To each $\mathfrak{O}_j$, we associate a set of potential basis vectors whose union over different j then yields the full set $(\mathfrak{b}_1, \dots, \mathfrak{b}_{q-1})$. More precisely, the set corresponding to $\mathfrak{O}_j$ is defined as

$$\mathscr{B}_j = \bigcup_{i=1}^{\mathfrak{o}-1} \left\{e_{g_i^{(j)}} + e_{g_{i+1}^{(j)}}\right\} \cup \left\{e_{g_1^{(j)}} + e_{g_2^{(j)}} + e_{(1-\chi_3)g_1^{(j)}}\right\}.$$

In the above, we have used that for $\chi_1 = 1$,

$$\chi_1 \cdot h + \chi_2 \cdot (-\chi_2^{-1})h + \chi_3 \cdot 0 = 0 \qquad \text{and} \qquad \chi_1 \cdot (1-\chi_3)h + \chi_2 \cdot (-\chi_2^{-1})h + \chi_3 \cdot h = 0.$$

Note that the element $(1-\chi_3)g_1^{(j)}$ is distinct from $g_1^{(j)}$. It might be one of $g_2^{(j)}, \dots, g_\mathfrak{o}^{(j)}$.

We next argue that the union of the different $\mathscr{B}_j$ generates $\mathbb{Z}^{\mathbb{F}_q^*}$. By linear transformation, $\mathscr{B}_j$ has the same span as

$$\left\{e_{g_1^{(j)}} + e_{g_2^{(j)}}, e_{g_1^{(j)}} - e_{g_3^{(j)}}, e_{g_1^{(j)}} + e_{g_4^{(j)}}, \dots, e_{g_1^{(j)}} \pm e_{g_\mathfrak{o}^{(j)}}\right\} \cup \left\{e_{(1-\chi_3)g_1^{(j)}}\right\}.$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Now, there are two cases.

1. For all $j \in [(q-1)/\mathfrak{o}]$, $(1-\chi_3)g_1^{(j)}$ is one of the elements $g_2^{(j)}, \ldots, g_{\mathfrak{o}}^{(j)}$. As in case 1, we can then subtract $e_{(1-\chi_3)g_1^{(j)}}$ from or add it to $e_{g_1^{(j)}} \pm e_{(1-\chi_3)g_1^{(j)}}$ to isolate $e_{g_1^{(j)}}$. After isolating $e_{g_1^{(j)}}$, a straightforward linear transformation yields a set of $\mathfrak{o}$ distinct unit vectors whose non-zero components are given by $\mathfrak{O}_j$. Thus, the union over all $\mathscr{B}_j$ constitutes a set of linearly independent elements that generates $\mathbb{Z}^{\mathbb{F}_q^*}$.
2. For all $j \in [(q-1)/\mathfrak{o}]$, $(1-\chi_3)g_1^{(j)}$ is none of the elements $g_2^{(j)}, \ldots, g_{\mathfrak{o}}^{(j)}$. In this case, consider the union $\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \mathscr{B}_j$, which has the same span as

$$\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \left\{ e_{g_1^{(j)}} + e_{g_2^{(j)}}, e_{g_1^{(j)}} - e_{g_3^{(j)}}, e_{g_1^{(j)}} + e_{g_4^{(j)}}, \ldots, e_{g_1^{(j)}} \pm e_{g_{\mathfrak{o}}^{(j)}} \right\} \cup \left\{ e_{(1-\chi_3)g_1^{(j)}} \right\}.$$

Since for each j, the element $(1-\chi_3)g_1^{(j)}$ must be contained in some $\mathfrak{O}_{j'}$ for $j \neq j'$, as in case (1), $e_{(1-\chi_3)g_1^{(j)}}$ can be used to isolate $e_{g_1^{(j')}}$. After isolating $e_{g_1^{(j')}}$ for all j', these elements can be straightforwardly used to linearly transform the union over all $\mathscr{B}_j$ into the standard basis $(e_h)_{h \in \mathbb{F}_q^*}$ of $\mathbb{Z}^{\mathbb{F}_q^*}$.

In any case, set $\bigcup_{j=1}^{(q-1)/\mathfrak{o}} \mathscr{B}_j = \{\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1}\}$.

7. Proof of Proposition 2.6

7.1 Overview

Recall that Proposition 2.6 concerns the model $\mathbb{A}$ with fixed numbers of non-zero entries per column and row, where both m and the degree sequences $(d_i^{(n)})_{1 \le i \le n}$ and $(k_i^{(m)})_{1 \le i \le m}$ are specified. For the sake of readability, throughout this section, we will omit the superscript from $d_i^{(n)}$ and $k_i^{(m)}$. Let $\mathfrak{A}$ be the σ-algebra generated by the numbers $\boldsymbol{m}(\chi_1, \ldots, \chi_\ell)$ of equations of degree $\ell \ge 3$ with coefficients $\chi_1, \ldots, \chi_\ell \in \mathbb{F}_q^*$. Let $\Delta = \sum_{i=1}^n d_i$ denote the total degree. As before, we let $\underline{\boldsymbol{A}}$ be the random matrix arising from the pairing model in this setting.

The aim in this section is to bound the expected size of the kernel of $\underline{\boldsymbol{A}}$ on $\mathfrak{O}$ from (2.7), that is, $|\ker \underline{\boldsymbol{A}}| \cdot \mathbb{1}\{\underline{\boldsymbol{A}} \in \mathfrak{O}\}$. This is related to Proposition 2.6 through the identities $\boldsymbol{Z}^2 = \boldsymbol{Z} \cdot |\ker \underline{\boldsymbol{A}}|$ and $\mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}^2 \cdot \mathbb{1}\{\underline{\boldsymbol{A}} \in \mathfrak{O}\}] = \mathbb{E}_{\mathfrak{A}}[\boldsymbol{Z}]\mathbb{E}_{\mathfrak{A}}[|\ker \underline{\boldsymbol{A}}| \cdot \mathbb{1}\{\underline{\boldsymbol{A}} \in \mathfrak{O}\}]$. Let us first observe that it suffices to count 'nearly equitable' kernel vectors, in the following sense. For a vector $\sigma \in \mathbb{F}_q^n$ and $s \in \mathbb{F}_q$ define the empirical frequency

$$\rho_\sigma(s) = \sum_{i=1}^n d_i \mathbb{1}\{\sigma_i = s\} \tag{7.1}$$

and let $\rho_\sigma = (\rho_\sigma(s))_{s \in \mathbb{F}_q}$. If $\mathfrak{O}$ occurs, then ρ_σ is nearly uniform for most kernel vectors. Formally, we have the following statement.

Fact 7.1. *For any $\varepsilon > 0$ and n large enough, we have* $\mathbb{1}\{\underline{\boldsymbol{A}} \in \mathfrak{O}\} \cdot |\ker \underline{\boldsymbol{A}}| \le (1+\varepsilon) \left| \left\{ \sigma \in \ker \underline{\boldsymbol{A}} : \|\rho_\sigma - q^{-1}\Delta \mathbb{1}\|_1 < \varepsilon \Delta \right\} \right|$.

Proof. Observe that to prove the claim, it is enough to show that for $\underline{\boldsymbol{A}} \in \mathfrak{O}$, w.h.p. for all $s \in \mathbb{F}_q$, $\sum_{i=1}^n d_i \mathbb{1}\left\{ \boldsymbol{x}_{\underline{\boldsymbol{A}}, i} = s \right\} - \Delta/q < \varepsilon \Delta$. Choose $\delta = \delta(\varepsilon, q) > 0$ small enough. Thanks to condition (P1), $\Delta = \sum_{i=1}^n d_i = \Omega(n)$. Moreover, (P3) ensures that the sequence $(\boldsymbol{d}_n)_n$ is uniformly

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

integrable, such that

$$\Delta > \sqrt{\delta}n \qquad \text{and} \qquad \sum_{i=1}^{n} \mathbb{1}\{d_i > d^*\}d_i < \delta\Delta \tag{7.2}$$

for a large constant d^* and all n large enough. On the other hand, for any degree $\ell \leq d^*$, a random vector $\boldsymbol{x}_{\underline{\boldsymbol{A}}} \in \ker \underline{\boldsymbol{A}}$ satisfies

$$\sum_{s,t\in\mathbb{F}_q} \sum_{i,j=1}^{n} \mathbb{1}\{d_i = d_j = \ell\} \left|\mathbb{P}\left[\boldsymbol{x}_{\underline{\boldsymbol{A}},i} = s,\ \boldsymbol{x}_{\underline{\boldsymbol{A}},j} = t \mid \underline{\boldsymbol{A}}\right] - q^{-2}\right| = o(n^2) \quad \text{for } \underline{\boldsymbol{A}} \in \mathfrak{O}. \tag{7.3}$$

Again by (P1), for all $\ell \in \mathrm{supp}(\boldsymbol{d})$, $\sum_{j=1}^{n} \mathbb{1}\{d_j = \ell\} = \Omega(n)$ and consequently (7.3) shows that

$$\sum_{i=1}^{n} \mathbb{1}\left\{d_i = \ell\right\} \left|\mathbb{P}\left[\boldsymbol{x}_{\underline{\boldsymbol{A}},i} = s|\underline{\boldsymbol{A}}\right] - 1/q\right| = o(n) \qquad \text{for all } s \in \mathbb{F}_q,\ \ell \leq d^* \text{ and } \underline{\boldsymbol{A}} \in \mathfrak{O}. \tag{7.4}$$

Combining (7.2) and (7.4) with the definition (7.1) of ρ_σ completes the proof. □

We proceed to contemplate different regimes of 'nearly equitable' frequency vectors and employ increasingly subtle estimates to bound their contributions. To this end, let $\mathfrak{P}_q$ be the set of all possible frequency vectors, that is,

$$\mathfrak{P}_q = \left\{\rho_\sigma : \sigma \in \mathbb{F}_q^n\right\}.$$

Moreover, for $\varepsilon > 0$ let

$$\mathfrak{P}_q(\varepsilon) = \left\{\rho \in \mathfrak{P}_q : \|\rho - q^{-1}\Delta\mathbb{1}\|_1 < \varepsilon\Delta\right\}.$$

In addition, we introduce

$$\begin{aligned}
\mathscr{Z}_\rho &= \left|\left\{\sigma \in \ker \underline{\boldsymbol{A}} : \rho_\sigma = \rho\right\}\right| && (\rho \in \mathfrak{P}_q),\\
\mathscr{Z}_\varepsilon &= \sum_{\rho\in\mathfrak{P}_q(\varepsilon)} \mathscr{Z}_\rho && (\varepsilon \geq 0),\\
\mathscr{Z}_{\varepsilon,\varepsilon'} &= \mathscr{Z}_{\varepsilon'} - \mathscr{Z}_\varepsilon && (\varepsilon, \varepsilon' \geq 0).
\end{aligned}$$

The following lemma sharpens the $\varepsilon\Delta$ error bound from Fact 7.1 to $\omega n^{-1/2}\Delta$.

Lemma 7.2. *For any small enough $\varepsilon > 0$, for large enough $\omega = \omega(\varepsilon) > 1$ we have $\mathbb{E}_{\mathfrak{A}}\left[\mathscr{Z}_{\omega n^{-1/2},\varepsilon}\right] < \varepsilon q^{n-m}$.*

The proof of Lemma 7.2, which can be found in Section 7.2, is based on an expansion to the second order of the optimisation problem (2.5) around the equitable solution. Similar arguments have previously been applied in the theory of random constraint satisfaction problems, particularly random k-XORSAT (e.g. [4, 6, 21]).

For ρ that are within $O(n^{-1/2}\Delta)$ of the equitable solution such relatively routine arguments do not suffice anymore. Indeed, by comparison to examples of random CSPs that have been studied previously, sometimes by way of the small sub-graph conditioning technique, a new challenge arises. Namely, due to the algebraic nature of our problem the conceivable empirical distributions $\rho_{\boldsymbol{x}}$ given that $\boldsymbol{x} \in \ker \underline{\boldsymbol{A}}$ are confined to a proper sub-lattice of $\mathbb{Z}^q$. The same is true of $\mathfrak{P}_q$ unless $\mathfrak{d} = 1$. Hence, we need to work out how these lattices intersect. Moreover, for $\rho \in \mathfrak{P}_q$ we need to calculate the number of assignments σ such that $\rho_\sigma = \rho$ as well as the probability that such an assignment satisfies all equations. Seizing upon Proposition 2.5 and local limit theorem-type techniques, we will deal with these challenges in Section 7.3, where we prove the following.

Lemma 7.3. *Assume that* $\mathfrak{d}$ *and* q *are coprime. Then for any* $\varepsilon > 0$ *for large enough* $\omega = \omega(\varepsilon) > 1$ *we have* $\mathbb{E}_{\mathfrak{A}}[\mathscr{Z}_{\omega n^{-1/2}}] \leq (1+\varepsilon)q^{n-m}$ *w.h.p.*

Proof of Proposition 2.6. This is an immediate consequence of Fact 7.1, Lemma 7.2 and Lemma 7.3. □

7.2 Proof of Lemma 7.2

As we just saw, on the one hand we need to count $\sigma \in \mathbb{F}_q^n$ such that ρ_σ hits a particular attainable $\rho \in \mathfrak{P}_q(\varepsilon)$. On the other hand, we need to estimate the probability that such a given σ satisfies all equations. The first of these, the entropy term, increases as ρ becomes more equitable. The second, the probability term, takes greater values for non-uniform ρ. Roughly, the more zero entries ρ contains, the better. The thrust of the proofs of Lemmas 7.2 and 7.3 is to show that the drop in entropy is an order of magnitude stronger than the boost to the success probability.

Toward the proof of Lemma 7.2 we can get away with relatively rough bounds, mostly disregarding constant factors. The first claim bounds the entropy term. Instead of counting assignments we will take a probabilistic viewpoint. Hence, let $\boldsymbol{\sigma} \in \mathbb{F}_q^n$ be a uniformly random assignment.

Claim 7.4. *There exists* $C > 0$ *such that w.h.p.* $\mathbb{P}_{\mathfrak{A}}\left[\|\rho_{\boldsymbol{\sigma}} - q^{-1}\Delta\mathbb{1}\|_1 > t\sqrt{\Delta}\right] \leq C\exp(-t^2/C)$ *for all* $t \geq 1$.

Proof. This is an immediate consequence of (P3) and Azuma–Hoeffding. □

Let us move on to the probability term. We proceed indirectly by way of Bayes' rule. Hence, fix $\rho \in \mathfrak{P}_q$ and let $\boldsymbol{\xi} = (\boldsymbol{\xi}_{ij})_{i,j\geq 1}$ be an infinite array of $\mathbb{F}_q$-valued random variables with distribution $\Delta^{-1}\rho$, mutually independent and independent of all other randomness. Moreover, let

$$\mathfrak{R}(\rho) = \bigcap_{s\in\mathbb{F}_q}\left\{\sum_{i=1}^{m}\sum_{j=1}^{k_i}\mathbb{1}\left\{\boldsymbol{\xi}_{ij} = s\right\} = \rho(s)\right\}, \qquad \mathfrak{S} = \left\{\forall i \in [m] : \sum_{j=1}^{k_i}\boldsymbol{\chi}_{ij}\boldsymbol{\xi}_{ij} = 0\right\}. \tag{7.5}$$

In words, $\mathfrak{R}(\rho)$ is the event that the empirical distribution induced by the random vector $\boldsymbol{\xi}_{ij}$, truncated at $i = m$ and $j = k_i$ for every i, works out to be $\rho \in \mathfrak{P}_q$. Furthermore, $\mathfrak{S}$ is the event that all m checks are satisfied if we substitute the independent values $\boldsymbol{\xi}_{ij}$ for the variables.

Crucially, $\mathfrak{S}$ ignores that the various equations share variables, or conversely that variables may appear in several distinct checks. Hence, the *unconditional* event $\mathfrak{S}$ effectively just deals with a linear system whose Tanner graph consists of m checks with degrees $k_1, \ldots, k_m$ and $\sum_{i=1}^m k_i$ variable nodes of degree one each. However, the *conditional* probability $\mathbb{P}_{\mathfrak{A}}[\mathfrak{S} \mid \mathfrak{R}(\rho)]$ equals the probability that a random assignment $\boldsymbol{\sigma}$ lies in the kernel of $\underline{\boldsymbol{A}}$ given that $\rho_{\boldsymbol{\sigma}} = \rho$:

Claim 7.5. *With the previous notation, for any* $\rho \in \mathfrak{P}_q$,

$$\mathbb{P}_{\mathfrak{A}}[\mathfrak{S} \mid \mathfrak{R}(\rho)] = \mathbb{P}_{\mathfrak{A}}[\boldsymbol{\sigma} \in \ker \underline{\boldsymbol{A}} \mid \rho_{\boldsymbol{\sigma}} = \rho]. \tag{7.6}$$

Proof. We relate both probabilities in (7.6) to the same random experiment. For this, let $\rho \in \mathfrak{P}_q$ be an empirical distribution that is compatible with the fixed vertex degrees, and additionally fix non-zero coefficients $(\boldsymbol{\chi}_1, \ldots, \boldsymbol{\chi}_{k_\ell})$ for every equation. Thus, we consider the linear system as fixed.

We first take a look at the left hand side of (7.6): Conditionally on the empirical distribution of the variables $(\boldsymbol{\xi}_{11}, \ldots, \boldsymbol{\xi}_{1k_1}, \boldsymbol{\xi}_{21}, \ldots, \boldsymbol{\xi}_{mk_m})$ being ρ, by exchangeability, every possible assignment

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

of values to the Δ positions in the linear system has the same probability $\binom{\Delta}{\rho}^{-1}$. The left hand side of (7.6) is thus equal to the number of all satisfying assignments with $\rho(s)$ s-entries for each $s \in \mathbb{F}_q$ divided by $\binom{\Delta}{\rho}$.

On the other hand, and turning to the right-hand side of (7.6), in the pairing model, variable clones are matched to check clones in a uniformly random manner. In such a uniform matching, for any fixed assignment σ with empirical distribution ρ, the probability to end up with a specific assignment of values to the Δ positions in the linear system has probability $\binom{\Delta}{\rho}^{-1}$. The right-hand side of (7.6) is thus equal to the number of all satisfying assignments with $\rho(s)$ s-entries for each $s \in \mathbb{F}_q$, to positions in the fixed linear system, divided by $\binom{\Delta}{\rho}$. □

We are going to see momentarily that the unconditional probabilities of $\mathfrak{R}(\rho)$ and $\mathfrak{S}$ are easy to calculate. In addition, we will be able to calculate the conditional probability $\mathbb{P}_{\mathfrak{A}}[\mathfrak{S} \mid \mathfrak{R}(\rho)]$ by way of the local limit theorem for sums of independent random variables. Finally, Lemma 7.2 will follow from these estimates via Bayes' rule.

Claim 7.6. *For any $\varepsilon > 0$, there exists $C = C(\varepsilon) > 0$ such that for all $\rho \in \mathfrak{P}_q(\varepsilon)$, $\mathbb{P}_{\mathfrak{A}}[\mathfrak{S}] \leq q^{m(C\sum_{s\in\mathbb{F}_q}|\Delta^{-1}\rho(s)-1/q|^3-1+\varepsilon^3)}$.*

Proof. For any $\rho \in \mathfrak{P}_q$, $h \geq 3$ and any $\chi_1, \ldots, \chi_h \in \operatorname{supp} \boldsymbol{\chi}$ we aim to calculate

$$P_h = \log \sum_{\sigma \in \mathbb{F}_q^h} \mathbb{1}\left\{\sum_{i=1}^h \chi_i \sigma_i = 0\right\} \prod_{i=1}^h \frac{\rho(\sigma_i)}{\Delta}.$$

With this notation, $\mathbb{P}_{\mathfrak{A}}[\mathfrak{S}] = \prod_{i=1}^m e^{P_{k_i}}$. We regard P_h as a function of the variables $(\rho(s))_{s\in\mathbb{F}_q}$ and will use Taylor's theorem to expand it around the constant vector $\bar{\rho} = q^{-1}\Delta\mathbb{1}$:

$$P_h(\rho) = P_h(\bar{\rho}) + DP_h(\bar{\rho})^T(\rho - \bar{\rho}) + \frac{1}{2}(\rho - \bar{\rho})^T D^2 P_h(\bar{\rho})(\rho - \bar{\rho}) + R_{\bar{\rho},3}(\rho) \tag{7.7}$$

for an appropriate error term

$$R_{\bar{\rho},3}(\rho) = \frac{1}{6} \sum_{s,s',s''\in\mathbb{F}_q} \frac{\partial^3 P_h}{\partial\rho(s)\partial\rho(s')\partial\rho(s'')}(z)\left(\rho(s) - \frac{\Delta}{q}\right)\left(\rho(s') - \frac{\Delta}{q}\right)\left(\rho(s'') - \frac{\Delta}{q}\right),$$

where z is some point z on the segment from $\bar{\rho}$ to ρ. Firstly, $P_h(\bar{\rho}) = -\log q$. The derivatives of P_h work out to be

$$\frac{\partial P_h}{\partial \rho(s)} = \frac{\sum_{j=1}^h \sum_{\sigma\in\mathbb{F}_q^h} \mathbb{1}\left\{\sum_{i=1}^h \chi_i\sigma_i = 0,\ \sigma_j = s\right\} \prod_{i\neq j} \frac{\rho(\sigma_i)}{\Delta}}{\Delta e^{P_h}} \qquad (s \in \mathbb{F}_q),$$

$$\frac{\partial^2 P_h}{\partial\rho(s)\partial\rho(s')} = \frac{\sum_{j\neq j'} \sum_{\sigma\in\mathbb{F}_q^h} \mathbb{1}\left\{\sum_{i=1}^h \chi_i\sigma_i = 0,\ \sigma_j = s, \sigma_{j'} = s'\right\} \prod_{i\neq j,j'} \frac{\rho(\sigma_i)}{\Delta}}{\Delta^2 e^{P_h}} - \frac{\partial P_h}{\partial\rho(s)} \frac{\partial P_h}{\partial\rho(s')} \quad (s, s' \in \mathbb{F}_q).$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Evaluating the derivatives at the equitable $\bar{\rho} = q^{-1}\Delta\mathbb{1}$ we obtain for any $h \geq 3$,

$$\left.\frac{\partial P_h}{\partial \rho(s)}\right|_{\bar{\rho}} = \frac{hq^{-1}}{\Delta q^{-1}} = \frac{h}{\Delta},$$
$$\left.\frac{\partial^2 P_h}{\partial \rho(s)\partial \rho(s')}\right|_{\bar{\rho}} = \frac{h(h-1)q^{-1}}{\Delta^2 q^{-1}} - \frac{h^2}{\Delta^2} = -\frac{h}{\Delta^2}, \qquad (s, s' \in \mathbb{F}_q).$$

Hence, the Jacobi matrix and the Hessian work out to be

$$DP_h(\bar{\rho}) = \frac{h}{\Delta}\mathbb{1}_q, \qquad D^2 P_h(\bar{\rho}) = -\frac{h}{\Delta^2}\mathbb{1}_{q\times q}. \tag{7.8}$$

For all $h \leq h^*$ and $\rho \in \mathfrak{P}_q(\varepsilon)$, the third partial derivatives are clearly uniformly bounded, that is, there is a constant $C(\varepsilon, h^*)$ such that

$$\frac{\partial^3 P_h}{\partial \rho(s)\partial \rho(s')\partial \rho(s'')} \leq C(\varepsilon, h^*)\cdot \Delta^{-3}. \tag{7.9}$$

Finally, for any $\varepsilon > 0$, because of assumptions (P1) and (P3), we can choose h^* large enough such that for n large enough, there are at most $\varepsilon^3 m$ equations with more than h^* variables. For these, we trivially bound $e^{P_h} \leq 1$. For the remaining equations of uniformly bounded degree, we use the previously described approach based on the Taylor expansion: Since $\rho - \bar{\rho} \perp \mathbb{1}_q$, (7.7), (7.8) and (7.9) imply the assertion. □

Claim 7.7. *For any* $\varepsilon > 0$, *there exists* $C = C(\varepsilon) > 0$ *such that for all* $\rho \in \mathfrak{P}_q(\varepsilon)$, $\mathbb{P}_{\mathfrak{A}}[\mathfrak{R}(\rho)] \geq C(\varepsilon)\cdot n^{(1-q)/2}$.

Proof. Since the $\boldsymbol{\xi}_{ij}$ are mutually independent, the probability of $\mathfrak{R}(\rho)$ given $\mathfrak{A}$ is nothing but

$$\mathbb{P}_{\mathfrak{A}}[\mathfrak{R}(\rho)] = \binom{\Delta}{(\rho(s))_{s\in\mathbb{F}_q}} \prod_{s\in\mathbb{F}_q}\left(\frac{\rho(s)}{\Delta}\right)^{\rho(s)}.$$

The claim therefore follows from Stirling's formula, together with assumption (P1). □

Claim 7.8. *For all* $\varepsilon > 0$ *small enough and for all* $\rho \in \mathfrak{P}_q(\varepsilon)$, $\mathbb{P}_{\mathfrak{A}}[\mathfrak{R}(\rho) \mid \mathfrak{S}] = O(n^{(1-q)/2})$.

Proof. The claim follows from the local limit theorem for sums of independent random variables (e.g. [18]). To elaborate, even once we condition on the event $\mathfrak{S}$ the random *vectors* $(\boldsymbol{\xi}_{ij})_{j\in[k_i]}$, $1 \leq i \leq m$, remain independent for different $i \in [m]$ due to the independence of the $(\boldsymbol{\xi}_{ij})_{i,j}$. Indeed, $\mathfrak{S}$ only asks that each check be satisfied separately, without inducing dependencies among different checks. Thus, the vector

$$\left(\sum_{i=1}^{m}\sum_{j=1}^{k_i}\mathbb{1}\left\{\boldsymbol{\xi}_{ij} = s\right\}\right)_{s\in\mathbb{F}_q} \qquad \text{given } \mathfrak{S}$$

is a sum of m independent random vectors. We first argue that $(\sum_{i=1}^{m}\sum_{j=1}^{k_i}\mathbb{1}\{\boldsymbol{\xi}_{ij} = s\})_{s\in\mathbb{F}_q^*}$ given $\mathfrak{S}$ satisfies a central limit theorem (note that we removed one coordinate from each vector). For this, let $\mathscr{C} \in \mathbb{R}^{\mathbb{F}_q^*\times\mathbb{F}_q^*}$ be defined by setting $\mathscr{C}(s, s') = \mathbb{1}\{s = s'\}\frac{1}{q} - \frac{1}{q^2}$. Then, thanks to the conditioning, for $\rho \in \mathfrak{P}_q(\varepsilon)$, for all n, i, all entries of $C_i = \mathrm{Cov}((\sum_{j=1}^{k_i}\mathbb{1}\{\boldsymbol{\xi}_{ij} = s\})_{s\in\mathbb{F}_q^*}|\mathfrak{S})$ will have distance at most δ from the corresponding entries of the matrix $k_i\cdot\mathscr{C}$, where $\delta = \delta(\varepsilon)$ and can be made arbitrarily small by choosing ε smaller. In particular, for ε small enough, all covariance matrices C_i are positive definite.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

By the Lindeberg-Feller CLT, the standardised sequence $((\sum_{i=1}^m C_i)^{-1/2} \sum_{i=1}^m \sum_{j=1}^{k_i} (\mathbb{1}\{\boldsymbol{\xi}_{ij} = s\} - \rho(s)/\Delta))_{s\in\mathbb{F}_q^*}$ given $\mathfrak{S}$ converges in distribution towards a multivariate standard Gaussian random variable if for every $\delta > 0$,

$$\sum_{i=1}^m \mathbb{E}\left[\left\|\left(\sum_{i=1}^m C_i\right)^{-1/2}\left(\sum_{j=1}^{k_i} \mathbb{1}\left\{\boldsymbol{\xi}_{ij} = s\right\} - \frac{k_i\rho(s)}{\Delta}\right)_{s\in\mathbb{F}_q^*}\right\|_2^2 \mathbb{1}\left\{\left\|\left(\sum_{i=1}^m C_i\right)^{-1/2}\left(\sum_{j=1}^{k_i} \mathbb{1}\left\{\boldsymbol{\xi}_{ij} = s\right\} - \frac{k_i\rho(s)}{\Delta}\right)_{s\in\mathbb{F}_q^*}\right\|_2 > \delta\right\} \middle| \mathfrak{S}\right] \longrightarrow 0. \tag{7.10}$$

To show (7.10), it is sufficient to show that for every $\delta' > 0$,

$$\frac{1}{m}\sum_{i=1}^m k_i^2 \mathbb{1}\left\{k_i > \delta'\sqrt{m}\right\} = \mathbb{E}\left[k_n^2 \mathbb{1}\left\{k_n > \delta'\sqrt{m}\right\}\right] \longrightarrow 0.$$

However, since $m = \Theta(n)$, this follows from the dominated convergence theorem via assumption (P3). Thus, the Lindeberg-Feller CLT applies. Moreover, since $\rho \in \mathfrak{P}_q(\varepsilon)$, $\mathbb{P}\left(\boldsymbol{\xi}_{ij} = s\right) \in (1/q - \varepsilon, 1/q + \varepsilon)$ for all $s \in \mathbb{F}_q$, so also the second condition of [18, Theorem 2.1] is satisfied: The local limit theorem therefore implies that the probability of the most likely outcome of this random vector is of order $n^{(1-q)/2}$; in symbols,

$$\max_{r\in\mathfrak{P}_q(\varepsilon)} \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{R}(r) \mid \mathfrak{S}\right] = O(n^{(1-q)/2}). \tag{7.11}$$

The assertion is an immediate consequence of (7.11). □

Proof of Lemma 7.2. Fix $\rho \in \mathfrak{P}_q(\varepsilon)$ such that $\omega\sqrt{\Delta} \le \sum_{s\in\mathbb{F}_q} |\rho(s) - \Delta/q| < \varepsilon\Delta$. Combining Claims 7.6–7.8 with Bayes' rule, we conclude that

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{S} \mid \mathfrak{R}(\rho)\right] = \frac{\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{S}\right]\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{R}(\rho) \mid \mathfrak{S}\right]}{\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{R}(\rho)\right]} = O(\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{S}\right]) = q^{m(O(\sum_{s\in\mathbb{F}_q} |\rho(s)/\Delta - 1/q|^3) + \varepsilon^3 - 1) + O(1)}. \tag{7.12}$$

Consequently, (7.6) and (7.12) imply that

$$\mathbb{P}_{\mathfrak{A}}\left[\boldsymbol{\sigma} \in \ker \underline{\boldsymbol{A}} \mid \rho_{\boldsymbol{\sigma}} = \rho\right] = \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{S} \mid \mathfrak{R}(\rho)\right] = q^{m(O(\sum_{s\in\mathbb{F}_q} |\rho(s)/\Delta - 1/q|^3) + \varepsilon^3 - 1) + O(1)}. \tag{7.13}$$

Hence, combining Claim 7.4 with (7.13) and Lemma 7.17 and using the bound $\sum_{s\in\mathbb{F}_q} |\rho(s) - \Delta/q| < \varepsilon\Delta$, we obtain

$$\begin{aligned}\mathbb{P}_{\mathfrak{A}}\left[\boldsymbol{\sigma} \in \ker \underline{\boldsymbol{A}},\ \rho_{\boldsymbol{\sigma}} = \rho\right] &= q^{m(\varepsilon^3 + O(\sum_{s\in\mathbb{F}_q} |\rho(s)/\Delta - 1/q|^3) - (\Omega(\sum_{s\in\mathbb{F}_q} |\rho(s)/\Delta - 1/q|^2) - 1) + O(1)} \\ &= q^{m(-1 - \Omega(\sum_{s\in\mathbb{F}_q} |\rho(s)/\Delta - 1/q|^2) + O(1)}.\end{aligned} \tag{7.14}$$

Multiplying (7.14) with q^n and summing on $\rho \in \mathfrak{P}_q(\varepsilon)$ such that $\omega n^{-1/2}\Delta \le \sum_{s\in\mathbb{F}_q} |\rho(s) - \Delta/q|$, we finally obtain

$$\begin{aligned}&\mathbb{E}_{\mathfrak{A}}\left[\mathscr{Z}_{\omega n^{-1/2},\varepsilon}\right] \\ &= q^{n-m+O(1)} \sum_{\substack{\rho\in\mathfrak{P}_q \\ \omega n^{-1/2}\Delta \le \sum_{s\in\mathbb{F}_q} |\rho(s) - \Delta/q| < \varepsilon\Delta}} \exp\left(-\Omega\left(n\sum_{s\in\mathbb{F}_q} |\rho(s)/\Delta - 1/q|^2\right)\right) < \varepsilon q^{n-m},\end{aligned}$$

provided $\omega = \omega(\varepsilon) > 0$ is chosen large enough. □

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

7.3 Proof of Lemma 7.3

By comparison to the proof of Lemma 7.2, the main difference here is that we need to be more precise. Specifically, while in Claims 7.7 and 7.8 we got away with disregarding constant factors, here we need to be accurate up to a multiplicative $1+o(1)$. Working out the probability term turns out to be delicate. As in Section 7.2, we introduce auxiliary $\mathbb{F}_q$-valued random variables $\boldsymbol{\xi}=(\boldsymbol{\xi}_{ij})_{i,j\geq 1}$. These random variables are mutually independent as well as independent of all other randomness. But this time all $\boldsymbol{\xi}_{ij}$ are *uniform* on $\mathbb{F}_q$. Let $\mathfrak{R}(\rho)$ and $\mathfrak{S}$ be the events from (7.5).

Similarly as in Section 7.2 we will ultimately apply Bayes' rule to compute the probability of $\mathfrak{S}$ given $\mathfrak{R}(\rho)$ and hence the conditional mean of $\mathscr{Z}_\rho$. The individual probability of $\mathfrak{R}(\rho)$ is easy to compute.

Claim 7.9. *For any $\rho\in\mathfrak{P}_q$ we have $\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{R}(\rho)\right]=\binom{\Delta}{\rho}q^{-\Delta}$.*

Proof. This is similar to the proof of Claim 7.7. As the $\boldsymbol{\xi}_{ij}$ are uniformly distributed and independent, we obtain

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{R}(\rho)\right]=\binom{\Delta}{(\rho(s))_{s\in\mathbb{F}_q}}\prod_{s\in\mathbb{F}_q}q^{-\rho(s)}=\binom{\Delta}{\rho}q^{-\Delta},$$

as claimed. □

As a next step we calculate the conditional probability of $\mathfrak{S}$ given $\mathfrak{R}(\rho)$. Similar to (7.1), for $s\in\mathbb{F}_q$ define the empirical frequency

$$\boldsymbol{\rho}(s)=\sum_{i=1}^{m}\sum_{j=1}^{k_i}\mathbb{1}\left\{\boldsymbol{\xi}_{ij}=s\right\} \tag{7.15}$$

and let $\boldsymbol{\rho}=(\boldsymbol{\rho}(s))_{s\in\mathbb{F}_q}$ as well as $\hat{\boldsymbol{\rho}}=(\boldsymbol{\rho}(s))_{s\in\mathbb{F}_q^*}$. Of course, Proposition 2.5 implies that for some $\rho\in\mathfrak{P}_q$ the event $\mathfrak{S}$ may be impossible given $\mathfrak{R}(\rho)$. Hence, to characterise the distributions ρ for which $\mathfrak{S}$ can occur at all, we let

$$\mathfrak{L}=\left\{r\in\mathbb{Z}^{\mathbb{F}_q^*}:\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}=r\right]>0\text{ and }\left\|r-q^{-1}\Delta\mathbb{1}\right\|_1\leq\omega n^{-1/2}\Delta\right\}, \tag{7.16}$$

$$\mathfrak{L}_0=\left\{r\in\mathfrak{L}:\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}=r\mid\mathfrak{S}\right]>0\right\}, \tag{7.17}$$

$$\mathfrak{L}_*=\left\{r\in\mathfrak{L}:\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}}=r\right]>0\right\}. \tag{7.18}$$

Thus, $\mathfrak{L}$ contains all conceivable outcomes of truncated frequency vectors. Moreover, $\mathfrak{L}_0$ comprises those frequency vectors that can occur given $\mathfrak{S}$, and $\mathfrak{L}_*$ those that can result from random assignments $\boldsymbol{\sigma}$ to the variables. Hence, $\mathfrak{L}_0$ is a finite subset of the $\mathbb{Z}$-module generated by those sets $\mathscr{S}_q(\chi_1,\ldots,\chi_\ell)$ from (2.15) with $\boldsymbol{m}(\chi_1,\ldots,\chi_\ell)>0$. The following lemma shows that actually the conditional probability $\mathfrak{S}$ given $\mathfrak{R}(\rho)$ is asymptotically the same for all $\rho\in\mathfrak{L}_0$, that is, for all conceivably satisfying ρ that are nearly equitable.

Lemma 7.10. *W.h.p. uniformly for all $r\in\mathfrak{L}_0$ we have $\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{S}\mid\hat{\boldsymbol{\rho}}=r\right]\sim q^{\mathbb{1}\{|\mathrm{supp}\,\boldsymbol{\chi}|=1\}-m}$.*

We complement Lemma 7.10 by the following estimate of the probability that a uniformly random assignment $\boldsymbol{\sigma}\in\mathbb{F}_q^n$ hits the set $\mathfrak{L}_0$ in the first place.

Lemma 7.11. *Assume that $\mathfrak{d}$ and q are coprime. Then w.h.p., $\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}}\in\mathfrak{L}_0\right]\leq(1+o(1))q^{-\mathbb{1}\{|\mathrm{supp}\,\boldsymbol{\chi}|=1\}}$.*

We prove Lemmas 7.10 and 7.11 in Sections 7.4 and 7.5, respectively.

Proof of Lemma 7.3. Formula (7.6) extends to the present auxiliary probability space with uniformly distributed and independent $\boldsymbol{\xi}_{ij}$ (for precisely the same reasons given in Section 7.2).

Hence, (7.6), (7.16) and (7.17) show that

$$\mathbb{E}_{\mathfrak{A}}[\mathscr{Z}_{\omega n^{-1/2}}] \leq \sum_{\sigma \in \mathbb{F}_q^n} \mathbb{1}\left\{\hat{\rho}_\sigma \in \mathfrak{L}\right\} \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{S} \mid \hat{\boldsymbol{\rho}} = \hat{\rho}_\sigma\right] = \sum_{\sigma \in \mathbb{F}_q^n} \mathbb{1}\left\{\hat{\rho}_\sigma \in \mathfrak{L}_0\right\} \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{S} \mid \hat{\boldsymbol{\rho}} = \hat{\rho}_\sigma\right]. \quad (7.19)$$

Finally, combining (7.19) with Lemma 7.10 and Lemma 7.11, we obtain

$$\begin{aligned}\mathbb{E}_{\mathfrak{A}}[\mathscr{Z}_{\omega n^{-1/2}}] &\leq (1+o(1))q^{\mathbb{1}\{|\mathrm{supp}\,\boldsymbol{\chi}|=1\}-m} \sum_{\sigma \in \mathbb{F}_q^n} \mathbb{1}\left\{\hat{\rho}_\sigma \in \mathfrak{L}_0\right\} \\ &= (1+o(1))q^{n-m+\mathbb{1}\{|\mathrm{supp}\,\boldsymbol{\chi}|=1\}} \mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{L}_0\right] \leq (1+o(1))q^{n-m},\end{aligned}$$

as desired. □

7.4 Proof of Lemma 7.10

Given $\omega > 0$ (from (7.16)) we choose $\varepsilon_0 = \varepsilon_0(\omega, q)$ sufficiently small and let $0 < \varepsilon < \varepsilon_0$. Moreover, recall that the degree sequences $(d_1, \ldots, d_n)$ and $(k_1, \ldots, k_m)$ satisfy properties (P1)-(P3). The proof hinges on a careful analysis of the conditional distribution of $\hat{\boldsymbol{\rho}}$ given $\mathfrak{S}$. We begin by observing that the vector $\hat{\boldsymbol{\rho}}$ is asymptotically normal given $\mathfrak{S}$. Let $\boldsymbol{I}_{(q-1)\times(q-1)}$ the $(q-1) \times (q-1)$-identity matrix and let $\boldsymbol{N} \in \mathbb{R}^{\mathbb{F}_q^*}$ be a Gaussian vector with zero mean and covariance matrix

$$\mathscr{C} = q^{-1}\boldsymbol{I}_{(q-1)\times(q-1)} - q^{-2}\mathbb{1}_{(q-1)\times(q-1)}. \quad (7.20)$$

Claim 7.12. *There exists a function $\alpha = \alpha(n, q) = o(1)$ such that for all axis-aligned cubes $U \subseteq \mathbb{R}^{\mathbb{F}_q^*}$ we have*

$$\left|\mathbb{P}_{\mathfrak{A}}\left[\Delta^{-1/2}(\hat{\boldsymbol{\rho}} - q^{-1}\Delta\mathbb{1}) \in U \mid \mathfrak{S}\right] - \mathbb{P}\left[N \in U\right]\right| \leq \alpha.$$

Proof. The conditional mean of $\hat{\boldsymbol{\rho}}$ given $\mathfrak{S}$ is uniform. To see this, consider any $i \in [m]$ and $h \in [k_i]$. We claim that for any vector $(\tau_j)_{j\in[k_i]\setminus\{h\}}$,

$$\mathbb{P}_{\mathfrak{A}}\left[\forall j \in [k_i] \setminus \{h\} : \boldsymbol{\xi}_{ij} = \tau_j \mid \mathfrak{S}\right] = q^{1-k_i}. \quad (7.21)$$

Indeed, for any such vector $(\tau_j)_{j\in[k_i]\setminus\{h\}}$ there is exactly one value $\boldsymbol{\xi}_{ih}$ that will satisfy the equation, namely

$$\boldsymbol{\xi}_{ih} = -\boldsymbol{\chi}_{ih}^{-1} \sum_{j\in[k_i]\setminus\{h\}} \boldsymbol{\chi}_{ij}\tau_j.$$

Hence, given $\mathfrak{S}$ the events $\{\forall j \in [k_i] \setminus \{h\} : \boldsymbol{\xi}_{ij} = \tau_j\}$ are equally likely for all τ, which implies (7.21). Furthermore, together with the definition (7.15) of $\boldsymbol{\rho}$, (7.21) readily implies that $\mathbb{E}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}|\mathfrak{S}\right] = q^{-1}\Delta\mathbb{1}$. Similarly, (7.21) also shows that $\Delta^{-1/2}\hat{\boldsymbol{\rho}}$ has covariance matrix $\mathscr{C}$ given $\mathfrak{S}$.

Finally, we are left to prove the desired uniform convergence to the normal distribution. To this end we employ the multivariate Berry-Esseen theorem (e.g. [45]). Specifically, given a small $\alpha > 0$ choose $K = K(q, \alpha) > 0$ and $m_0 = m_0(K)$, $n_0 = n_0(K, m_0)$ sufficiently large. Assuming

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

$n > n_0$, since $m = \Theta(n)$, we can ensure that $m > m_0$. Also let

$$k_i' = \mathbb{1}\{k_i \le K\}k_i, \qquad k_i'' = k_i - k_i',$$

$$\hat{\boldsymbol{\rho}}'(s) = \sum_{1\le i\le m: k_i \le K} \sum_{j=1}^{k_i} \mathbb{1}\{\boldsymbol{\xi}_{ij} = s\}, \qquad \hat{\boldsymbol{\rho}}''(s) = \sum_{1\le i\le m: k_i > K} \sum_{j=1}^{k_i} \mathbb{1}\{\boldsymbol{\xi}_{ij} = s\},$$

$$\Delta' = \sum_{i=1}^{m} k_i', \qquad \Delta'' = \sum_{i=1}^{m} k_i''.$$

Now again, assumption (P3) implies that the sequence $(\boldsymbol{k}_n)_n$ is uniformly integrable, such that for large enough n,

$$\Delta'' < \alpha^8 \Delta. \tag{7.22}$$

Moreover, by the same reasoning as in the previous paragraph the random vectors $\hat{\boldsymbol{\rho}}'$ and $\hat{\boldsymbol{\rho}}''$ have means $q^{-1}\Delta'$ and $q^{-1}\Delta''$ and covariances $\Delta'\mathscr{C}$ and $\Delta''\mathscr{C}$, respectively. Thus, (7.22) and Chebyshev's inequality show that

$$\mathbb{P}_{\mathfrak{A}}\left[\left\|\frac{\hat{\boldsymbol{\rho}}'' - q^{-1}\Delta''\mathbb{1}}{\sqrt{\Delta}}\right\| > \alpha^2 \middle| \mathfrak{S}\right] < \alpha^2. \tag{7.23}$$

Further, the Berry–Esseen theorem shows that

$$\mathbb{P}_{\mathfrak{A}}\left[\frac{\hat{\boldsymbol{\rho}}' - q^{-1}\Delta'\mathbb{1}}{\sqrt{\Delta'}} \in U \middle| \mathfrak{S}\right] - \mathbb{P}\left[\boldsymbol{N} \in U\right] = O(K \cdot n^{-1/2}) \qquad \text{for all cubes } U. \tag{7.24}$$

Here, $O(\cdot)$ refers to an n- and K-independent factor. Combining (7.24) and (7.23), we see that

$$\left|\mathbb{P}_{\mathfrak{A}}\left[\frac{\hat{\boldsymbol{\rho}} - q^{-1}\Delta\mathbb{1}}{\sqrt{\Delta}} \in U \middle| \mathfrak{S}\right] - \mathbb{P}\left[\boldsymbol{N} \in U\right]\right| \le \alpha. \tag{7.25}$$

The assertion follows from (7.25) by taking $\alpha \to 0$ slowly as $n \to \infty$. For example, it is possible to choose $\alpha = \log^{-1} n$ and $K = \Theta\left(n^{1/4}\right)$ thanks to assumption (P3). □

The following claim states that the normal approximation from Claim 7.12 also holds for the unconditional random vector $\hat{\boldsymbol{\rho}}$.

Claim 7.13. *There exists a function $\alpha = \alpha(n, q) = o(1)$ such that for all convex sets $U \subseteq \mathbb{R}^{\mathbb{F}_q^*}$ we have*

$$\left|\mathbb{P}_{\mathfrak{A}}\left[\Delta^{-1/2}(\hat{\boldsymbol{\rho}} - q^{-1}\Delta\mathbb{1}) \in U\right] - \mathbb{P}\left[N \in U\right]\right| \le \alpha.$$

Proof. This is an immediate consequence of Claim 7.9 and Stirling's formula. □

Let $k_0 = \min \operatorname{supp}(\boldsymbol{k})$. In the case that $|\operatorname{supp}\boldsymbol{\chi}| = 1$ we set $\chi_1 = \cdots = \chi_{k_0}$ to the single element of $\operatorname{supp}\boldsymbol{\chi}$. Moreover, in the case that $|\operatorname{supp}\boldsymbol{\chi}| > 1$ we pick and fix any $\chi_1, \ldots, \chi_{k_0} \in \operatorname{supp}\boldsymbol{\chi}$ such that $|\{\chi_1, \ldots, \chi_{k_0}\}| > 1$. Let $\mathfrak{I}_0$ be the set of all $i \in [m]$ such that $k_i = k_0$ and $\boldsymbol{\chi}_{ij} = \chi_j$ for $j = 1, \ldots, k_0$ and let $\mathfrak{I}_1 = [m] \setminus \mathfrak{I}_0$. Then $|\mathfrak{I}_0| = \Theta(n)$ w.h.p. Further, set

$$\boldsymbol{r}_0(s) = \sum_{i\in\mathfrak{I}_0} \sum_{j\in[k_i]} \mathbb{1}\left\{\boldsymbol{\xi}_{ij} = s\right\}, \qquad \boldsymbol{r}_1(s) = \sum_{i\in\mathfrak{I}_1} \sum_{j\in[k_i]} \mathbb{1}\left\{\boldsymbol{\xi}_{ij} = s\right\} \qquad (s \in \mathbb{F}_q^*).$$

Then $\hat{\boldsymbol{\rho}} = \boldsymbol{r}_0 + \boldsymbol{r}_1$.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Because the vectors $\boldsymbol{\xi}_i = (\boldsymbol{\xi}_{i,1}, \ldots, \boldsymbol{\xi}_{i,k_i})$ are mutually independent, so are $\boldsymbol{r}_0 = (\boldsymbol{r}_0(s))_{s\in\mathbb{F}_q^*}$ and $\boldsymbol{r}_1 = (\boldsymbol{r}_1(s))_{s\in\mathbb{F}_q^*}$. To analyse $\boldsymbol{r}_0$ precisely, let

$$\mathscr{S}_0 = \left\{ \sigma \in \mathbb{F}_q^{k_0} : \sum_{i=1}^{k_0} \chi_i \sigma_i = 0 \right\}.$$

Moreover, for $\sigma \in \mathscr{S}_0$ let $\boldsymbol{R}_\sigma$ be the number of indices $i \in \mathfrak{I}_0$ such that $\boldsymbol{\xi}_i = \sigma$. Then conditionally on $\mathfrak{S}$, we have

$$\boldsymbol{r}_0(s) = \sum_{i\in\mathfrak{I}_0} \sum_{j\in[k_i]} \mathbb{1}\left\{\boldsymbol{\xi}_{ij} = s\right\} = \sum_{\sigma\in\mathscr{S}_0} \sum_{j=1}^{k_0} \mathbb{1}\left\{\sigma_j = s\right\} \boldsymbol{R}_\sigma \qquad \text{given } \mathfrak{S},$$

which reduces our task to the investigation of $\boldsymbol{R} = (\boldsymbol{R}_\sigma)_{\sigma\in\mathscr{S}_0}$.

This is not too difficult because given $\mathfrak{S}$ the random vector $\boldsymbol{R}$ has a multinomial distribution with parameter $|\mathfrak{I}_0|$ and uniform probabilities $|\mathscr{S}_0|^{-1}$. In effect, the individual entries $\boldsymbol{R}(\sigma)$, $\sigma \in \mathfrak{S}_0$, will typically differ by only a few standard deviations, that is, their typical difference will be of order $O(\sqrt{\Delta})$. We require a precise quantitative version of this statement.

Recalling the sets from (7.16) to (7.18), for $r_* \in \mathfrak{L}_0$ and $0 < \varepsilon < \varepsilon_0$ we let

$$\mathfrak{L}_0(r_*, \varepsilon) = \left\{ r \in \mathfrak{L}_0 : \|r - r_*\|_\infty < \varepsilon\sqrt{\Delta} \right\}.$$

Furthermore, we say that $\boldsymbol{R}$ is *t-tame* if $|\boldsymbol{R}_\sigma - |\mathscr{S}_0|^{-1}|\mathfrak{I}_0|| \le t\sqrt{\Delta}$ for all $\sigma \in \mathscr{S}_0$. Let $\mathfrak{T}(t)$ be the event that $\boldsymbol{R}$ is t-tame.

Lemma 7.14. *W.h.p. for every $r_* \in \mathfrak{L}_0$ there exists $r^* \in \mathfrak{L}_0(r_*, \varepsilon)$ such that*

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r^* \mid \mathfrak{S}\right] \ge \frac{1}{2|\mathfrak{L}_0(r_*, \varepsilon)|} \qquad \textit{and} \qquad \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log\varepsilon) \mid \mathfrak{S},\ \hat{\boldsymbol{\rho}} = r^*\right] \ge 1 - \varepsilon^4. \tag{7.26}$$

Proof. Recall that the event $\{\hat{\boldsymbol{\rho}} = r\}$ is the same as $\mathfrak{R}(r')$ with $r'(s) = r(s)$ for $s \in \mathbb{F}_q^*$ and $r'(0) = \Delta - \|r\|_1$. As a first step we observe that $\boldsymbol{R}$ given $\mathfrak{S}$ is reasonably tame with a reasonably high probability. More precisely, since $\boldsymbol{R}$ has a multinomial distribution given $\mathfrak{A}$ and $\mathfrak{S}$, the Chernoff bound shows that w.h.p.

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log\varepsilon) \mid \mathfrak{S}\right] \ge 1 - \exp(-\Omega_\varepsilon(\log^2(\varepsilon))). \tag{7.27}$$

Further, Claim 7.12 implies that $\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} \in \mathfrak{L}_0(r_*, \varepsilon) \mid \mathfrak{S}\right] \ge \Omega_\varepsilon(\varepsilon^{q-1}) \ge \varepsilon^q$ w.h.p., provided $\varepsilon < \varepsilon_0 = \varepsilon_0(\omega)$ is small enough. Combining this estimate with (7.27) and Bayes' formula, we conclude that w.h.p. for every $r_* \in \mathfrak{L}_0$,

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log\varepsilon) \mid \mathfrak{S},\ \hat{\boldsymbol{\rho}} \in \mathfrak{L}_0(r_*, \varepsilon)\right] \ge 1 - \varepsilon^5. \tag{7.28}$$

To complete the proof, assume that there does not exist $r^* \in \mathfrak{L}_0(r_*, \varepsilon)$ that satisfies (7.26). Then for every $r \in \mathfrak{L}_0(r_*, \varepsilon)$ we either have

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}=r \mid \mathfrak{S}\right]<\frac{1}{2|\mathfrak{L}_0(r_*,\varepsilon)|} \qquad \text{or} \tag{7.29}$$

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log\varepsilon) \mid \mathfrak{S},\ \hat{\boldsymbol{\rho}}=r\right]<1-\varepsilon^4. \tag{7.30}$$

Let $\mathfrak{X}_0$ be the set of all $r\in\mathfrak{L}_0(r_*,\varepsilon)$ for which (7.29) holds, and let $\mathfrak{X}_1=\mathfrak{L}_0(r_*,\varepsilon)\setminus\mathfrak{X}_0$. Then (7.29)–(7.30) yield

$$\begin{aligned}
&\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log\varepsilon) \mid \mathfrak{S},\ \hat{\boldsymbol{\rho}}\in\mathfrak{L}_0(r_*,\varepsilon)\right]\\
&\leq \frac{\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}\in\mathfrak{X}_0 \mid \mathfrak{S}\right]+\sum_{r\in\mathfrak{X}_1}\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log\varepsilon) \mid \mathfrak{S},\ \hat{\boldsymbol{\rho}}=r\right]\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}=r \mid \mathfrak{S}\right]}{\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}\in\mathfrak{L}_0(r_*,\varepsilon) \mid \mathfrak{S}\right]}\\
&<\frac{\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}\in\mathfrak{X}_0 \mid \mathfrak{S}\right]+(1-\varepsilon^4)\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}\in\mathfrak{X}_1 \mid \mathfrak{S}\right]}{\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}}\in\mathfrak{L}_0(r_*,\varepsilon) \mid \mathfrak{S}\right]}<1-\varepsilon^4,
\end{aligned}$$

provided that $1-\varepsilon^4>\frac{1}{2}$, in contradiction to (7.28). □

Let $\mathfrak{M}=\mathfrak{M}_q(\chi_1,\ldots,\chi_{k_0})$ and let $\mathfrak{b}_1,\ldots,\mathfrak{b}_{q-1}$ be the basis of $\mathfrak{M}$ supplied by Proposition 2.5. Let us fix vectors $\tau^{(1)},\ldots,\tau^{(q-1)}\in\mathscr{S}_0$ whose frequency vectors as defined in (2.16) coincide with $\mathfrak{b}_1,\ldots,\mathfrak{b}_{q-1}$, that is,

$$\hat{\tau}^{(i)}=\mathfrak{b}_i \qquad \text{for } i=1,\ldots,q-1.$$

Also let $\mathfrak{T}(r,t)$ be the event that $\hat{\boldsymbol{\rho}}=r$ and that $\boldsymbol{R}$ is t-tame. The following lemma summarises the key step of the proof of Lemma 7.10.

Lemma 7.15. *W.h.p. for any $r_*\in\mathfrak{L}_0$, any $1\leq t\leq\log n$ and any $r,r'\in\mathfrak{L}_0(r_*,\varepsilon)$ there exists a one-to-one map $\psi:\mathfrak{T}(r,t)\to\mathfrak{T}(r',t+O_\varepsilon(\varepsilon))$ such that for all $(R,r_1)\in\mathfrak{T}(r,t)$ we have*

$$\log\frac{\mathbb{P}_{\mathfrak{A}}\left[(\boldsymbol{R},\boldsymbol{r}_1)=(R,r_1) \mid \mathfrak{S}\right]}{\mathbb{P}_{\mathfrak{A}}\left[(\boldsymbol{R},\boldsymbol{r}_1)=\psi(R,r_1) \mid \mathfrak{S}\right]}=O_\varepsilon(\varepsilon(\omega+t)). \tag{7.31}$$

Proof. Since $r,r'\in\mathfrak{M}$, we have $r-r'\in\mathfrak{M}$ w.h.p. Indeed, if $\mathrm{supp}\,\boldsymbol{\chi}>1$, then Proposition 2.5 shows that $\mathfrak{M}=\mathbb{Z}^{\mathbb{F}_q^*}$ w.h.p. Moreover, if $\mathrm{supp}\,\boldsymbol{\chi}=1$, then $\mathfrak{M}$ is a proper subset of the integer lattice $\mathbb{Z}^{\mathbb{F}_q^*}$. Nonetheless, Proposition 2.5 shows that the modules

$$\mathfrak{M}_q(\underbrace{1,\ldots,1}_{\ell\text{ times}})$$

coincide for all $\ell\geq 3$, and therefore $\mathfrak{M}$ coincides with the $\mathbb{Z}$-module generated by $\mathfrak{L}_0$. Hence, in either case there is a unique representation

$$r'-r=\sum_{i=1}^{q-1}\lambda_i\mathfrak{b}_i \qquad (\lambda_i\in\mathbb{Z}) \tag{7.32}$$

in terms of the basis vectors. Because $r,r'\in\mathfrak{L}_0(r_*,\varepsilon)$ and

$$\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_{q-1}\end{pmatrix}=\left(\mathfrak{b}_1\ \cdots\ \mathfrak{b}_{q-1}\right)^{-1}(r-r'),$$

the coefficients satisfy

$$|\lambda_i|=O_\varepsilon(\varepsilon\sqrt{\Delta}) \qquad \text{for all } i=1,\ldots,q-1. \tag{7.33}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Now let $\lambda_0 = -\sum_{i=1}^{q-1} \lambda_i$, obtain the vector R' from R by amending the entry R_0' corresponding to the zero solution $0 \in \mathscr{S}_0$ to

$$R'_0 = R_0 + \lambda_0, \qquad R_{\tau^{(i)}}{}' = R_{\tau^{(i)}} + \lambda_i \quad \text{for all } i \in [q-1] \qquad \text{and} \qquad R_\sigma{}' = R_\sigma \quad \text{for all } \sigma \notin \{0, \tau^{(1)}, \ldots, \tau^{(q-1)}\}.$$

Further, define $\psi(R, r_1) = (R', r_1)$. Then $\psi(R, r_1) \in \mathfrak{T}(r', t + O_\varepsilon(\varepsilon))$ due to (7.32) and (7.33). Moreover, Stirling's formula and the mean value theorem show that

$$\begin{aligned}\frac{\mathbb{P}_{\mathfrak{A}}\left[(\boldsymbol{R}, \boldsymbol{r}_1) = (R, r_1) \mid \mathfrak{S}\right]}{\mathbb{P}_{\mathfrak{A}}\left[(\boldsymbol{R}, \boldsymbol{r}_1) = \psi(R, r_1) \mid \mathfrak{S}\right]} &= \binom{|\mathfrak{I}_0|}{R}\binom{|\mathfrak{I}_0|}{R'}^{-1} = \exp\left[\sum_{\sigma \in \mathscr{S}_0} O_\varepsilon\left(R_\sigma \log R_\sigma - R'_\sigma \log R'_\sigma\right)\right] \\ &= \exp\left[O_\varepsilon(|\mathfrak{I}_0|) \sum_{\sigma \in \mathscr{S}_0} \left|\int_{R'_\sigma/|\mathfrak{I}_0|}^{R_\sigma/|\mathfrak{I}_0|} \log z dz\right|\right] \\ &= \exp\left[O_\varepsilon(|\mathfrak{I}_0|) \sum_{\sigma \in \mathscr{S}_0} \left(\frac{R_\sigma}{|\mathfrak{I}_0|} - \frac{R'_\sigma}{|\mathfrak{I}_0|}\right) \log\left(\frac{1}{q} + O_\varepsilon\left(\frac{\omega + t}{\sqrt{\Delta}}\right)\right)\right] \\ &= \exp\left[O_\varepsilon(|\mathfrak{I}_0|) \sum_{\sigma \in \mathscr{S}_0} O_\varepsilon\left(\frac{\omega + t}{\sqrt{\Delta}}\left(\frac{R_\sigma}{|\mathfrak{I}_0|} - \frac{R'_\sigma}{|\mathfrak{I}_0|}\right)\right)\right]. \end{aligned} \tag{7.34}$$

Since $|\mathfrak{I}_0| = \Theta_\varepsilon(\Delta) = \Theta_\varepsilon(n)$ w.h.p., (7.34) implies (7.31). Finally, ψ is one to one because each vector has a unique representation with respect to the basis $(\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1})$. □

Roughly speaking, Lemma 7.15 shows that any two tame $r, r' \in \mathfrak{L}_0(r_*, \varepsilon)$ close to a conceivable $r_* \in \mathfrak{L}_0$ are about equally likely. However, the map ψ produces solutions that are a little less tame than the ones we start from. The following corollary, which combines Lemmas 7.14 and 7.15, remedies this issue.

Corollary 7.16. *W.h.p. for all $r_* \in \mathfrak{L}_0$ and all $r, r' \in \mathfrak{L}_0(r_*, \varepsilon)$ we have*

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -3\log\varepsilon) \mid \mathfrak{S}\right] = (1 + o_\varepsilon(1))\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r', -3\log\varepsilon) \mid \mathfrak{S}\right].$$

Proof. Let r^* be the vector supplied by Lemma 7.14. Applying Lemma 7.15 to r^* and $r \in \mathfrak{L}_0(r_*, \varepsilon)$, we see that w.h.p.

$$\begin{aligned}\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -2\log\varepsilon) \mid \mathfrak{S}\right] &\geq \left(1 + O_\varepsilon\left(\varepsilon \log \varepsilon\right)\right) \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r^*, -\log\varepsilon) \mid \mathfrak{S}\right] \\ &\geq \frac{1}{3|\mathfrak{L}_0(r_*, \varepsilon)|} \qquad \text{for all } r \in \mathfrak{L}_0(r_*, \varepsilon).\end{aligned} \tag{7.35}$$

In addition, we claim that w.h.p.

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -4\log\varepsilon) \setminus \mathfrak{T}(r, -3\log\varepsilon) \mid \mathfrak{S}\right] \leq \varepsilon \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r^*, -\log\varepsilon) \mid \mathfrak{S}\right] \qquad \text{for all } r \in \mathfrak{L}_0(r_*, \varepsilon). \tag{7.36}$$

Indeed, applying Lemma 7.15 twice to r and r^* and invoking (7.26), we see that w.h.p.

$$\begin{aligned}\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -2\log\varepsilon) \mid \mathfrak{S}\right] &\geq \exp(O_\varepsilon(\varepsilon \log \varepsilon))\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r^*, -3\log\varepsilon) \mid \mathfrak{S}\right] \\ &\geq \left(1 - O_\varepsilon(\varepsilon \log \varepsilon)\right) \mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r^* \mid \mathfrak{S}\right],\end{aligned} \tag{7.37}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -4\log\varepsilon) \setminus \mathfrak{T}(r, -3\log\varepsilon) \mid \mathfrak{S}\right]$$
$$\leq \exp(O_\varepsilon(\varepsilon\log\varepsilon))\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r^*, -3\log\varepsilon) \setminus \mathfrak{T}(r^*, -2\log\varepsilon) \mid \mathfrak{S}\right] O_\varepsilon(\varepsilon^4)\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r^* \mid \mathfrak{S}\right]. \tag{7.38}$$

Combining (7.37) and (7.38) yields (7.36).

Finally, (7.26), (7.35) and (7.36) show that w.h.p.

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-3\log\varepsilon) \mid \hat{\boldsymbol{\rho}} = r, \mathfrak{S}\right] \geq 1 - \sqrt{\varepsilon},$$
$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-3\log\varepsilon) \mid \hat{\boldsymbol{\rho}} = r', \mathfrak{S}\right] \geq 1 - \sqrt{\varepsilon} \qquad \text{for all } r, r' \in \mathfrak{L}_0(r_*, \varepsilon), \tag{7.39}$$

and combining (7.39) with Lemma 7.15 completes the proof. □

Proof of Lemma 7.10. We are going to show that the conditional probability $\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r \mid \mathfrak{S}\right]$ of hitting some particular $r \in \mathfrak{L}_0$ coincides with the unconditional probability $\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r\right]$ up to a factor of $(1 + o_\varepsilon(1))q^{\mathbb{1}\{|\mathrm{supp}\,\boldsymbol{\chi}|=1\}}$. Then the assertion follows from Bayes' formula.

The unconditional probability $\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r\right]$ is given precisely by Claim 7.9. Hence, recalling the $(q-1) \times (q-1)$-matrix $\mathscr{C} = q^{-1}\boldsymbol{I}_{(q-1)\times(q-1)} - q^{-2}\mathbb{1}_{(q-1)\times(q-1)}$ from (7.20) and applying Stirling's formula, we obtain

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r\right] \sim \frac{q^{q/2}}{(2\pi\Delta)^{(q-1)/2}} \exp\left[-\frac{(r - q^{-1}\Delta\mathbb{1})^\top \mathscr{C}^{-1}(r - q^{-1}\Delta\mathbb{1})}{2\Delta}\right] \tag{7.40}$$

w.h.p.

Next we will show that the conditional probability $\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r \mid \mathfrak{S}\right]$ works out to be asymptotically the same, up to an additional factor of $q^{\mathbb{1}\{|\mathrm{supp}\,\boldsymbol{\chi}|=1\}}$. Indeed, Claim 7.12 shows that for any $r \in \mathfrak{L}_0$ the conditional probability that $\hat{\boldsymbol{\rho}}$ hits the set $\mathfrak{L}_0(r, \varepsilon)$ is asymptotically equal to the probability of the event $\{\|\boldsymbol{N} - \Delta^{-1/2}(r - q^{-1}\Delta\mathbb{1})\|_\infty < \varepsilon\}$. Moreover, Corollary 7.16 implies that given $\hat{\boldsymbol{\rho}} \in \mathfrak{L}_0(r, \varepsilon)$, $\hat{\boldsymbol{\rho}}$ is within $o_\varepsilon(1)$ of the uniform distribution on $\mathfrak{L}_0(r, \varepsilon)$. Furthermore, Lemma 3.6 and Proposition 2.5 show that the number of points in $\mathfrak{L}_0(r, \varepsilon)$ satisfies

$$\frac{|\mathfrak{L}_0(r, \varepsilon)|}{\left|\left\{z \in \mathbb{Z}^{q-1} : \|z - r\|_\infty \leq \varepsilon\sqrt{\Delta}\right\}\right|} \sim q^{-\mathbb{1}\{|\mathrm{supp}\,\boldsymbol{\chi}|=1\}}.$$

Therefore, w.h.p. for all $r \in \mathfrak{L}_0$ we have

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\boldsymbol{\rho}} = r \mid \mathfrak{S}\right] = (1 + o_\varepsilon(1))q^{\mathbb{1}\{|\mathrm{supp}\,\boldsymbol{\chi}|=1\}} \frac{q^{q/2}}{(2\pi\Delta)^{(q-1)/2}} \exp\left[-\frac{(r - q^{-1}\Delta\mathbb{1})^\top \mathscr{C}^{-1}(r - q^{-1}\Delta\mathbb{1})}{2\Delta}\right]. \tag{7.41}$$

Finally, we observe that

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{S}\right] = q^{-m}. \tag{7.42}$$

Indeed, since the $\boldsymbol{\xi}_{ij}$ are uniform and independent, for each $i \in [m]$ we have $\sum_{j=1}^{k_i} \chi_{i,j}\boldsymbol{\xi}_{ij} = 0$ with probability $1/q$ independently. Combining (7.40)–(7.42) completes the proof. □

7.5 Proof of Lemma 7.11

We continue to denote by $\boldsymbol{\sigma} \in \mathbb{F}_q^n$ a uniformly random assignment and by $\boldsymbol{I}_{(q-1)\times(q-1)}$ the $(q-1) \times (q-1)$-identity matrix. Also recall ρ_σ from (7.1) and for $\rho = (\rho(s))_{s\in\mathbb{F}_q}$ obtain $\hat{\rho} = (\rho(s))_{s\in\mathbb{F}_q^*}$

by dropping the 0-entry. The following claim, which we prove via the local limit theorem for sums of independent random variables, determines the distribution of $\rho_{\boldsymbol{\sigma}}$. Let $\bar{\rho} = q^{-1}\Delta\mathbb{1}_{q-1}$.

Claim 7.17. *Let $\mathscr{C}$ be the $(q-1)\times(q-1)$-matrix from (7.20) and $\Delta_2 = \sum_{i=1}^n d_i^2$. Then w.h.p. for all $\rho \in \mathfrak{P}_q$ we have*

$$\mathbb{P}_{\mathfrak{A}}\left[\rho_{\boldsymbol{\sigma}} = \rho\right] = \frac{q^{q/2}\mathfrak{d}^{q-1}}{(2\pi\Delta_2)^{(q-1)/2}}\exp\left(-\frac{(\hat{\rho}-\bar{\rho})^\top\mathscr{C}^{-1}(\hat{\rho}-\bar{\rho})}{2\Delta_2}\right) + o(n^{(1-q)/2})$$

The proof of Claim 7.17 is based on local limit theorem techniques similar to but simpler than the ones from Section 7.4. In fact, the proof strategy is somewhat reminiscent of that of the well-known local limit theorem for sums of independent random vectors from [18]. However, the local theorem from that paper does not imply Claim 7.17 directly because a key assumption (that increments of vectors in each direction can be realised) is not satisfied here. We therefore carry the details out in the appendix.

Claim 7.17 demonstrates that $\rho_{\boldsymbol{\sigma}}$ satisfies a local limit theorem. Hence, let $\mathbf{N}' \in \mathbb{R}^{q-1}$ be a mean-zero Gaussian vector with covariance matrix $\mathscr{C}$. Moreover, fix $\varepsilon > 0$ and let $U = v + [-\varepsilon, \varepsilon]^{q-1} \subseteq \mathbb{R}^{q-1}$ be a box of side length 2ε. Then w.h.p. we have

$$\mathbb{P}_{\mathfrak{A}}\left[\Delta_2^{-1/2}(\hat{\rho}_{\boldsymbol{\sigma}} - q^{-1}\Delta\mathbb{1}) \in U\right] = \mathbb{P}_{\mathfrak{A}}\left[\mathbf{N}' \in U\right] + o(1), \tag{7.43}$$

where Δ_2 is as in Claim 7.17. This can be seen as in the proof of Lemma 7.8. Indeed, Claim 7.17 implies that $\hat{\rho}_{\boldsymbol{\sigma}}$ is asymptotically uniformly distributed on the lattice points of the box $\Delta_2(U + q^{-1}\Delta\mathbb{1})$ whose coordinates are divisible by $\mathfrak{d}$ w.h.p. Thus, w.h.p. for any $z, z' \in \Delta_2(U + q^{-1}\Delta\mathbb{1}) \cap \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}$ we have

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} = z\right] = (1 + o_\varepsilon(1))\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} = z'\right]. \tag{7.44}$$

Let $\tilde{U} = \Delta_2(U + q^{-1}\Delta\mathbb{1})$. Moreover, we claim that

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{L}_0 \mid \hat{\rho}_{\boldsymbol{\sigma}} \in \tilde{U}\right] \sim \frac{\left|\tilde{U} \cap \mathfrak{L}_0 \cap \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}\right|}{\left|\tilde{U} \cap \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}\right|} \le \frac{\left|\tilde{U} \cap \mathfrak{M} \cap \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}\right|}{\left|\tilde{U} \cap \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}\right|} \le (1 + o(1))q^{-\mathbb{1}\{|\mathrm{supp}\boldsymbol{\chi}| = 1\}}. \tag{7.45}$$

Indeed, if $|\mathrm{supp}\boldsymbol{\chi}| > 1$, then (7.45) is satisfied w.h.p. for the trivial reason that the r.h.s. equals $1 + o(1)$. Hence, suppose that $|\mathrm{supp}\boldsymbol{\chi}| = 1$, let $\mathfrak{M} \supset \mathfrak{L}_0$ be the module from Proposition 2.5 and let $\mathfrak{b}_1, \ldots, \mathfrak{b}_{q-1}$ be its assorted basis. Clearly, $\mathfrak{M} \cap \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*} \supseteq \mathfrak{d}\mathfrak{M}$. Conversely, Cramer's rule shows that any $y \in \mathfrak{M} \cap \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}$ can be expressed as

$$(\mathfrak{b}_1 \;\cdots\; \mathfrak{b}_{q-1})z, \qquad \text{with } z_i = \frac{\det(\mathfrak{b}_1 \cdots \mathfrak{b}_{i-1}\, y\, \mathfrak{b}_{i+1} \cdots \mathfrak{b}_{q-1})}{q}.$$

In particular, all coordinates z_i are divisible by $\mathfrak{d}$ because $y \in \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}$. Hence, $y \in \mathfrak{d}\mathfrak{M}$ because $\mathfrak{d}$ and q are coprime. Lemma 3.6 therefore implies (7.45). Finally, the assertion follows from (7.43)–(7.45).

8. Proof of Proposition 4.1

We prove Proposition 4.1 by way of a coupling argument inspired by the Aizenman-Sims-Starr scheme from spin glass theory [5]. The proof is a close adaptation of the coupling argument used in [10] to prove the approximate rank formula (1.4). We will therefore be able to reuse some of the technical steps from that paper. The main difference is that we need to accommodate the extra ternary equations t_i. Their presence gives rise to the second parameter β in (4.5).

8.1 Overview

The basic idea behind the Aizenman-Sims-Starr scheme is to compute the expected difference $\mathbb{E}[\mathrm{nul}\boldsymbol{A}[n+1,\varepsilon,\delta,\Theta]] - \mathbb{E}[\mathrm{nul}\boldsymbol{A}[n,\varepsilon,\delta,\Theta]]$ of the nullity upon increasing the size of the matrix. We then obtain (4.5) by writing a telescoping sum. In order to estimate the expected change of the nullity, we set up a coupling of $\boldsymbol{A}[n,\varepsilon,\delta,\Theta]$ and $\boldsymbol{A}[n+1,\varepsilon,\delta,\Theta]$.

To this end it is helpful to work with a description of the random matrix model that is different from the earlier definition of the model in Section 4, which is closer to the original matrix model. The present modification is owed to the fact that it will turn out beneficial to actually order the check variables according to their degree: Specifically, let $\boldsymbol{M} = (\boldsymbol{M}_j)_{j\geq 1}$, $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_j)_{j\geq 1}$, $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ be Poisson variables with means

$$\mathbb{E}[\boldsymbol{M}_j] = (1-\varepsilon)\mathbb{P}\left[\boldsymbol{k}=j\right] dn/k, \qquad \mathbb{E}[\boldsymbol{\Delta}_j] = (1-\varepsilon)\mathbb{P}\left[\boldsymbol{k}=j\right] d/k, \qquad \mathbb{E}\left[\boldsymbol{\lambda}\right] = \delta n,$$
$$\mathbb{E}\left[\boldsymbol{\eta}\right] = \delta. \tag{8.1}$$

All these random variables are mutually independent and independent of $\boldsymbol{\theta}$ and the $(\boldsymbol{d}_i)_{i\geq 1}$. Further, let

$$\boldsymbol{M}_j^+ = \boldsymbol{M}_j + \boldsymbol{\Delta}_j, \qquad \boldsymbol{m}_{\varepsilon,n} = \sum_{j\geq 1} \boldsymbol{M}_j, \qquad \boldsymbol{m}_{\varepsilon,n}^+ = \sum_{j\geq 1} \boldsymbol{M}_j^+, \qquad \boldsymbol{\lambda}^+ = \boldsymbol{\lambda} + \boldsymbol{\eta}. \tag{8.2}$$

Since $\sum_{j\geq 1} \boldsymbol{M}_j \sim \mathrm{Po}((1-\varepsilon)dn/k)$, (8.2) is consistent with (4.1).

We define a random Tanner (multi-)graph $\mathbf{G}\left[n,\boldsymbol{M},\boldsymbol{\lambda}\right]$ with variable nodes $x_1,\ldots,x_n$ and check nodes $a_{i,j}$, $i\geq 1, j\in[\boldsymbol{M}_i]$, $t_1,\ldots,t_{\boldsymbol{\lambda}}$ and $p_1,\ldots,p_{\boldsymbol{\theta}}$. Here, the first index of each check variable $a_{i,j}$ will indicate its degree. The edges between variables and the check nodes $a_{i,j}$ are induced by a random maximal matching $\boldsymbol{\Gamma}\left[n,\boldsymbol{M}\right]$ of the complete bipartite graph with vertex classes

$$\bigcup_{h=1}^{n} \{x_h\} \times [\boldsymbol{d}_h] \quad \text{and} \quad \bigcup_{i\geq 1}\bigcup_{j=1}^{\boldsymbol{M}_i} \{a_{i,j}\} \times [i].$$

Moreover, for each $j\in[\boldsymbol{\lambda}]$ we choose $\mathbf{i}_{j,1},\mathbf{i}_{j,2},\mathbf{i}_{j,3}$ uniformly and independently from $[n]$ and add edges between $x_{\mathbf{i}_{j,1}}, x_{\mathbf{i}_{j,2}}, x_{\mathbf{i}_{j,3}}$ and t_j. In addition, we insert an edge between p_i and x_i for every $i\in[\boldsymbol{\theta}]$.

To define the random matrix $\boldsymbol{A}\left[n,\boldsymbol{M},\boldsymbol{\lambda}\right]$ to go with $\mathbf{G}\left[n,\boldsymbol{M},\boldsymbol{\lambda}\right]$, let

$$\boldsymbol{A}\left[n,\boldsymbol{M},\boldsymbol{\lambda}\right]_{p_i,x_h} = \mathbb{1}\left\{i=h\right\} \qquad (i\in[\boldsymbol{\theta}], h\in[n]), \tag{8.3}$$

$$\boldsymbol{A}\left[n,\boldsymbol{M},\boldsymbol{\lambda}\right]_{a_{i,j},x_h} = \boldsymbol{\chi}_{i,h} \sum_{\ell=1}^{i}\sum_{s=1}^{\boldsymbol{d}_h} \mathbb{1}\{(x_h,s),(a_{i,j},\ell)\}\in\boldsymbol{\Gamma}_{n,\boldsymbol{M}}\} \qquad (i\geq 1, j\in[\boldsymbol{M}_i], h\in[n]), \tag{8.4}$$

$$\boldsymbol{A}\left[n,\boldsymbol{M},\boldsymbol{\lambda}\right]_{t_i,x_h} = \boldsymbol{\chi}_{\boldsymbol{m}_{\varepsilon,n}+i,h} \sum_{\ell=1}^{3} \mathbb{1}\{\mathbf{i}_{i,\ell}=h\} \qquad (i\in[\boldsymbol{\lambda}], h\in[n]). \tag{8.5}$$

The Tanner graph $\mathbf{G}\left[n+1,\boldsymbol{M}^+,\boldsymbol{\lambda}^+\right]$ and its associated random matrix $\boldsymbol{A}\left[n+1,\boldsymbol{M}^+,\boldsymbol{\lambda}^+\right]$ are defined analogously using $n+1$ variable nodes instead of n, $\boldsymbol{M}^+$ instead of $\boldsymbol{M}$ and $\boldsymbol{\lambda}^+$ instead of $\boldsymbol{\lambda}$.

Fact 8.1. *For any $\varepsilon,\delta>0$ we have*

$$\mathbb{E}[\mathrm{nul}\boldsymbol{A}\left[n,\varepsilon,\delta,\Theta\right]] = \mathbb{E}[\mathrm{nul}\boldsymbol{A}\left[n,\boldsymbol{M},\boldsymbol{\lambda}\right]], \quad \mathbb{E}[\mathrm{nul}\boldsymbol{A}\left[n+1,\varepsilon,\delta,\Theta\right]] = \mathbb{E}\left[\mathrm{nul}\boldsymbol{A}\left[n+1,\boldsymbol{M}^+,\boldsymbol{\lambda}^+\right]\right].$$

Proof. Because the check degrees $\boldsymbol{k}_i$ of the random factor graph $\mathbf{G}\left[n,\varepsilon,\delta,\Theta\right]$ are drawn independently, the only difference between $\mathbf{G}\left[n,\varepsilon,\delta,\Theta\right]$ and $\mathbf{G}\left[n,\boldsymbol{M},\boldsymbol{\lambda}\right]$ is the bookkeeping of the number of checks of each degree. The same is true of $\mathbf{G}\left[n+1,\varepsilon,\delta,\Theta\right]$ and $\mathbf{G}\left[n+1,\boldsymbol{M},\boldsymbol{\lambda}\right]$. $\square$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

To construct a coupling of $\boldsymbol{A}[n, \boldsymbol{M}, \boldsymbol{\lambda}]$ and $\boldsymbol{A}\left[n+1, \boldsymbol{M}^+, \boldsymbol{\lambda}^+\right]$ we introduce a third, intermediate random matrix. Hence, let $\boldsymbol{\gamma}_i \geq 0$ be the number of checks $a_{i,j}$, $j \in [\boldsymbol{M}_i^+]$, adjacent to the last variable node x_{n+1} in $\boldsymbol{G}\left[n+1, \boldsymbol{M}^+, \boldsymbol{\lambda}^+\right]$. Set $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_i)_{i \geq 3}$. Also let

$$\lambda^- = \delta(n+1) \cdot \left(\frac{n}{n+1}\right)^3 \tag{8.6}$$

be the expected number of extra ternary checks of $\boldsymbol{G}\left[n+1, \boldsymbol{M}^+, \boldsymbol{\lambda}^+\right]$ in which x_{n+1} does not appear (recall that each of the $\mathrm{Po}(\delta(n+1))$ ternary checks chooses its variables independently and uniformly at random from all $(n+1)^3$ possibilities). Let

$$\boldsymbol{M}_i^- = (\boldsymbol{M}_i - \boldsymbol{\gamma}_i) \vee 0, \qquad \text{as well as} \qquad \boldsymbol{\lambda}^- \sim \mathrm{Po}(\lambda^-). \tag{8.7}$$

Consider the random Tanner graph $\boldsymbol{G}' = \boldsymbol{G}\left[n, \boldsymbol{M}^-, \boldsymbol{\lambda}^-\right]$ induced by a random maximal matching $\boldsymbol{\Gamma}' = \boldsymbol{\Gamma}\left[n, \boldsymbol{M}^-\right]$ of the complete bipartite graph with vertex classes

$$\bigcup_{h=1}^{n} \{x_h\} \times [\boldsymbol{d}_h] \quad \text{and} \quad \bigcup_{i \geq 1} \bigcup_{j=1}^{\boldsymbol{M}_i^-} \{a_{i,j}\} \times [i]. \tag{8.8}$$

Each matching edge $\{(x_h, s), (a_{i,j}, \ell)\} \in \boldsymbol{\Gamma}\left[n, \boldsymbol{M}^-\right]$ induces an edge between x_h and $a_{i,j}$ in the Tanner graph. For each $j \in [\boldsymbol{\lambda}^-]$ and $\mathbf{i}_{j,1}^-, \mathbf{i}_{j,2}^-, \mathbf{i}_{j,3}^-$ uniform and independent in $[n]$, we add the edges between $x_{\mathbf{i}_{j,1}^-}, x_{\mathbf{i}_{j,2}^-}, x_{\mathbf{i}_{j,3}^-}$ and t_j. In addition, there is an edge between p_i and x_i for every $i \in [\boldsymbol{\theta}]$. Let $\boldsymbol{A}'$ denote the corresponding random matrix.

For each variable x_i, $i = 1, \ldots, n$, let $\mathscr{C}$ be the set of clones from $\bigcup_{i \in [n]} \{x_i\} \times [\boldsymbol{d}_i]$ that $\boldsymbol{\Gamma}\left[n, \boldsymbol{M}^-\right]$ leaves unmatched. We call the elements of $\mathscr{C}$ *cavities.*

From $\boldsymbol{G}'$, we finally construct two further Tanner graphs. Obtain the Tanner graph $\boldsymbol{G}''$ from $\boldsymbol{G}'$ by adding new check nodes $a_{i,j}''$ for each $i \geq 3$, $j \in [\boldsymbol{M}_i - \boldsymbol{M}_i^-]$ and ternary check nodes $t_i''$for $i \in [\boldsymbol{\lambda}'']$, where

$$\boldsymbol{\lambda}'' \sim \mathrm{Po}(\delta n - \lambda^-) = \mathrm{Po}\left(\delta n \left(1 - \left(\frac{n}{n+1}\right)^2\right)\right). \tag{8.9}$$

The new checks $a_{i,j}''$ are joined by a random maximal matching $\boldsymbol{\Gamma}''$ of the complete bipartite graph on

$$\mathscr{C} \qquad \text{and} \qquad \bigcup_{i \geq 1} \bigcup_{j \in [M_i - M_i^-]} \{a_{i,j}''\} \times [i].$$

Moreover, for each $j \in [\boldsymbol{\lambda}'']$ we choose $\mathbf{i}_{j,1}'', \mathbf{i}_{j,2}'', \mathbf{i}_{j,3}'' \in [n]$ uniformly and independently of everything else and add the edges between $x_{\mathbf{i}_{j,1}}'', x_{\mathbf{i}_{j,2}}'', x_{\mathbf{i}_{j,3}}''$ and t_j''. Let $\boldsymbol{A}''$ denote the corresponding random matrix, where as before, each new edge is represented by an independent copy of $\boldsymbol{\chi}$.

Finally, let

$$\boldsymbol{\lambda}''' \sim \mathrm{Po}(\delta(n+1) - \lambda^-) = \mathrm{Po}\left(\delta(n+1)\left(1 - \left(\frac{n}{n+1}\right)^3\right)\right). \tag{8.10}$$

We analogously obtain $\boldsymbol{G}'''$ by adding one variable node x_{n+1} as well as check nodes $a_{i,j}'''$, $i \geq 1$, $j \in [\boldsymbol{\gamma}_i]$, $b_{i,j}'''$, $i \geq 1, j \in [\boldsymbol{M}_i^+ - \boldsymbol{M}_i^- - \boldsymbol{\gamma}_i]$, t_i''', $i \in [\boldsymbol{\lambda}''']$. The new checks $a_{i,j}'''$ and $b_{i,j}'''$ are connected to $\boldsymbol{G}'$ via a random maximal matching $\boldsymbol{\Gamma}'''$ of the complete bipartite graph on

$$\mathscr{C} \qquad \text{and} \qquad \bigcup_{i \geq 1} \left(\bigcup_{j \in [\boldsymbol{\gamma}_i]} \{a_{i,j}'''\} \times [i-1] \cup \bigcup_{j \in [M_i^+ - M_i^- - \boldsymbol{\gamma}_i]} \{b_{i,j}'''\} \times [i] \right).$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

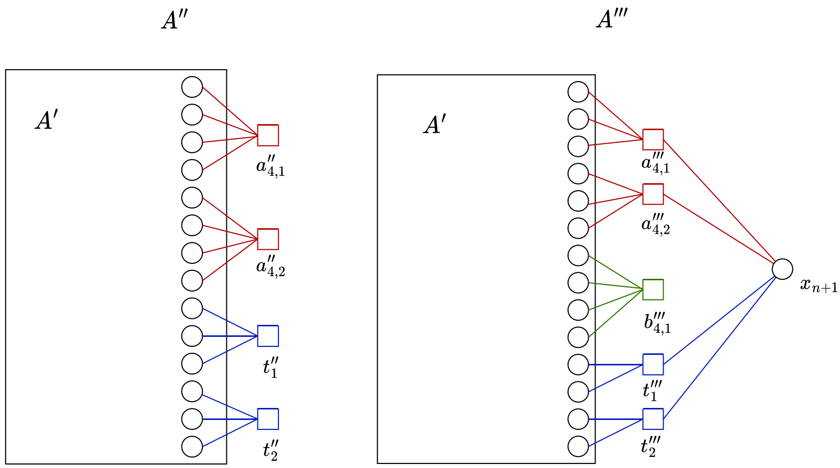

Figure 8. Visualisation of the construction of the auxiliary matrices $\boldsymbol{A}''$ and $\boldsymbol{A}'''$ from $\boldsymbol{A}'$. The matrices are identified with their Tanner graph in the graphical representation.

For each matching edge we insert the corresponding variable-check edge and in addition each of the check nodes $a'''_{i,j}$ gets connected to x_{n+1} by exactly one edge. Then we connect each t'''_i to $x'''_{\mathbf{i}_{i,1}}, x'''_{\mathbf{i}_{i,2}}$ and x_{n+1}, with $\boldsymbol{i}'''_{i,1}, \boldsymbol{i}'''_{i,2} \in [n+1]$ chosen uniformly and independently. Once again each edge is represented by an independent copy of $\boldsymbol{\chi}$. Let $\boldsymbol{A}'''$ denote the resulting random matrix.

The following lemma connects $\boldsymbol{A}'', \boldsymbol{A}'''$ with the random matrices $\boldsymbol{A}[n, \boldsymbol{M}, \boldsymbol{\lambda}]$, $\boldsymbol{A}\left[n+1, \boldsymbol{M}^+, \boldsymbol{\lambda}^+\right]$ and thus, in light of Fact 8.1, with $\boldsymbol{A}[n, \varepsilon, \delta]$ and $\boldsymbol{A}[n+1, \varepsilon, \delta]$ (See Fig. 8).

Lemma 8.2. *We have* $\mathbb{E}[\mathrm{nul}(\boldsymbol{A}'')] = \mathbb{E}[\mathrm{nul}(\boldsymbol{A}[n, \boldsymbol{M}, \boldsymbol{\lambda}])] + o(1)$ *and* $\mathbb{E}[\mathrm{nul}(\boldsymbol{A}''')] = \mathbb{E}[\mathrm{nul}(\boldsymbol{A}[n+1, \boldsymbol{M}^+, \boldsymbol{\lambda}^+])] + o(1)$.

We defer the simple proof of Lemma 8.2 to Section 8.5.

The core of the proof of Proposition 4.1 is to estimate the difference of the nullities of $\boldsymbol{A}'''$ and $\boldsymbol{A}'$ and of $\boldsymbol{A}''$ and $\boldsymbol{A}'$. The following two lemmas express these differences in terms of two random variables $\boldsymbol{\alpha}, \boldsymbol{\beta}$. Specifically, let $\boldsymbol{\alpha}$ be the fraction of frozen cavities of $\boldsymbol{A}'$ and let $\boldsymbol{\beta}$ be the fraction of frozen variables of $\boldsymbol{A}'$.

Lemma 8.3. *For large enough* $\Theta(\varepsilon) > 0$ *and small enough* $0 < \delta < \delta_0$ *we have*

$$\mathbb{E}[\mathrm{nul}(\boldsymbol{A}''') - \mathrm{nul}(\boldsymbol{A}')] = \mathbb{E}\left[\exp\left(-3\delta\boldsymbol{\beta}^2\right) D(1 - K'(\boldsymbol{\alpha})/k)\right] + \frac{d}{k}\mathbb{E}\left[K'(\boldsymbol{\alpha}) + K(\boldsymbol{\alpha})\right]$$
$$- \frac{d(k+1)}{k} - 3\delta\mathbb{E}\left[1 - \boldsymbol{\beta}^2\right] + o_\varepsilon(1).$$

Lemma 8.4. *For large enough* $\Theta(\varepsilon) > 0$ *and small enough* $0 < \delta < \delta_0$ *we have*

$$\mathbb{E}[nul(\boldsymbol{A}'') - nul(\boldsymbol{A}')] = -d + \frac{d}{k}\mathbb{E}\left[\boldsymbol{\alpha}K'(\boldsymbol{\alpha})\right] - 2\delta\mathbb{E}\left[1 - \boldsymbol{\beta}^3\right] + o_\varepsilon(1).$$

After some preparations in Section 8.2 we will prove Lemmas 8.3 and 8.4 in Sections 8.3 and 8.4.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Proof of Proposition 4.1. For any $\varepsilon, \delta > 0$, by Fact 8.1 and Lemma 8.2, we have

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}[n,\varepsilon,\delta,\Theta])\right] \le \limsup_{n\to\infty} \mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}[n+1,\boldsymbol{M}^{+},\boldsymbol{\lambda}^{+},\Theta])\right] - \mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}[n,\boldsymbol{M},\boldsymbol{\lambda},\Theta])\right]$$
$$\le \limsup_{n\to\infty} \mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}''') - \mathrm{nul}(\boldsymbol{A}')\right] - \mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}')\right].$$

For large enough $\Theta(\varepsilon) > 0$ and small enough $0 < \delta < \delta_0$, we can further upper bound the last expression via Lemma 8.3 and Lemma 8.4. Taking the maximum over all possible realisations of the random variables $\boldsymbol{\alpha}, \boldsymbol{\beta}$ finishes the proof of Proposition 4.1. □

8.2 Preparations

To facilitate the proofs of Lemmas 8.3 and 8.4 we establish a few basic statements about the coupling. Some of these are immediate consequences of statements from [10], where a similar coupling was used. Let us begin with the following lower bound on the likely number of cavities.

Lemma 8.5. *W.h.p. we have* $|\mathscr{C}| \ge \varepsilon dn/2$.

Proof. Apart from the extra ternary check nodes $t_1, \ldots t_{\boldsymbol{\lambda}'}$, the construction of $\boldsymbol{G}'$ coincides with that of the Tanner graph from [10]. Because the presence of $t_1, \ldots t_{\boldsymbol{\lambda}'}$ does not affect the number of cavities, the assertion therefore follows from [10, Lemma 5.5]. □

As a next step we show that w.h.p. the random matrix $\boldsymbol{A}'$ does not have very many short linear relations. Specifically, if we choose a bounded number of variables and a bounded number of cavities randomly, then it is quite unlikely that the chosen coordinates form a proper relation. Formally, let $\mathscr{R}(\ell_1, \ell_2)$ be the set of all sequences $(i_1, \ldots, i_{\ell_1}) \in [n]^{\ell_1}$, $(u_1, j_1), \ldots, (u_{\ell_2}, j_{\ell_2}) \in \mathscr{C}$ such that $(i_1, \ldots, i_{\ell_1}, u_1, \ldots, u_{\ell_2})$ is a proper relation of $\boldsymbol{A}'$. Furthermore, let $\mathfrak{R}(\zeta, \ell)$ be the event that $|\mathscr{R}(\ell_1, \ell_2)| \le \zeta n^{\ell_1}|\mathscr{C}|^{\ell_2}$ for all $0 \le \ell_1, \ell_2 \le \ell$.

Lemma 8.6. *For any* $\zeta > 0$, $\ell > 0$ *exist* $\Theta_0 = \Theta_0(\varepsilon, \zeta, \ell) > 0$ *and* $n_0 > 0$ *such that for all* $n \ge n_0$, $\Theta \ge \Theta_0$ *we have* $\mathbb{P}\left[\mathfrak{R}(\zeta, \ell)\right] > 1 - \zeta$.

Proof. Fix any $\ell_1, \ell_2 \le \ell$ such that $\ell_1 + \ell_2 > 0$ and let $\mathfrak{R}(\zeta, \ell_1, \ell_2)$ be the event that $|\mathscr{R}(\ell_1, \ell_2)| < \zeta n^{\ell_1}|C|^{\ell_2}$. Then it suffices to show that $\mathbb{P}\left[\mathfrak{R}(\zeta, \ell_1, \ell_2)\right] > 1 - \zeta$ as we can just replace ζ by $\zeta/(\ell + 1)^2$ and apply the union bound. To this end we may assume that $\zeta < \zeta_0(\varepsilon, \ell)$ for a small enough $\zeta_0(\varepsilon, \ell) > 0$.

We will actually estimate $|\mathscr{R}(\ell_1, \ell_2)|$ on a certain likely event. Specifically, due to Lemma 8.5 we have $|\mathscr{C}| \ge \varepsilon n/2$ w.h.p. In addition, let $\mathscr{A}$ be the event that $\boldsymbol{A}'$ is $(\zeta^4/L^{\ell}, \ell)$-free. Then Proposition 3.4 shows that $\mathbb{P}\left[\mathscr{A}\right] > 1 - \zeta/3$, provided that $n \ge n_0$ for a large enough $n_0 = n_0(\zeta, \ell)$. To see this, consider the matrix $\boldsymbol{B}$ obtained from $\boldsymbol{A}'$ by deleting the rows representing the unary checks p_i. Then Proposition 3.4 shows that the matrix $\boldsymbol{B}[\boldsymbol{\theta}]$ obtained from $\boldsymbol{B}$ via the pinning operation is (ζ^4, L^{ℓ})-free with probability $1 - \zeta/3$, provided that Θ is chosen sufficiently large. The only difference between $\boldsymbol{B}[\boldsymbol{\theta}]$ and $\boldsymbol{A}'$ is that in the former random matrix we apply the pinning operation to $\boldsymbol{\theta}$ random coordinates, while in $\boldsymbol{A}'$ the unary checks p_i pin the first $\boldsymbol{\theta}$ coordinates. However, the distribution of $\boldsymbol{A}'$ is actually invariant under permutations of the columns. Therefore, the matrices $\boldsymbol{A}'$ and $\boldsymbol{B}[\boldsymbol{\theta}]$ are (ζ^4, L^{ℓ})-free with precisely the same probability. Hence, Proposition 3.4 implies that $\mathbb{P}\left[\mathscr{A}\right] > 1 - \zeta/3$.

Further, Markov's inequality shows that for any $L > 0$,

$$\mathbb{P}\left[\sum_{i=1}^{n} \mathbf{d}_i \mathbb{1}\{\mathbf{d}_i > L\} \ge \frac{\varepsilon\zeta^2 n}{16\ell}\right] \le \frac{16\ell\mathbb{E}\left[\mathbf{d}\mathbb{1}\{\mathbf{d} > L\}\right]}{\varepsilon\zeta^2}.$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Therefore, since $\mathbb{E}[\mathbf{d}] = O_\varepsilon(1)$ we can choose $L = L(\varepsilon, \zeta, \ell) > 0$ big enough such that the event

$$\mathscr{L} = \left\{ \sum_{i=1}^{n} \mathbf{d}_i \mathbb{1}\{\mathbf{d}_i > L\} < \frac{\varepsilon\zeta^2 n}{16\ell} \right\}$$

has probability at least $1 - \zeta/3$. Thus, the event $\mathscr{E} = \mathscr{A} \cap \mathscr{L} \cap \{|\mathscr{C}| \geq \varepsilon n/2\}$ satisfies $\mathbb{P}[\mathscr{E}] > 1 - \zeta$. Hence, it suffices to show that

$$|\mathscr{R}(\ell_1, \ell_2)| < \zeta n^{\ell_1} |\mathscr{C}|^{\ell_2} \qquad \text{if the event } \mathscr{E} \text{ occurs.} \tag{8.11}$$

To bound $\mathscr{R}(\ell_1, \ell_2)$ on $\mathscr{E}$ we need to take into consideration that the cavities are degree weighted. Hence, let $\mathscr{R}'(\ell_1, \ell_2)$ be the set of all sequences $(i_1, \ldots, i_{\ell_1}, (u_1, j_1), \ldots, (u_{\ell_2}, j_{\ell_2})) \in \mathscr{R}(\ell_1, \ell_2)$ such that the degree of some variable node u_i exceeds L. Assuming $\ell_2 > 0$, on $\mathscr{E}$ we have

$$|\mathscr{R}'(\ell_1, \ell_2)| \leq n^{\ell_1} \cdot \ell_2 \cdot |\mathscr{C}|^{\ell_2 - 1} \sum_{i=1}^{n} \mathbf{d}_i \mathbb{1}\{\mathbf{d}_i > L\} \leq n^{\ell_1} |\mathscr{C}|^{\ell_2} \cdot \frac{2}{\varepsilon n} \cdot \ell_2 \cdot \frac{\varepsilon\zeta^2 n}{16\ell} < \frac{\zeta}{2} n^{\ell_1} |\mathscr{C}|^{\ell_2}, \tag{8.12}$$

provided that $\zeta > 0$ is small enough. Here, we have used that on $\mathscr{E}$, $|\mathscr{C}| \geq \varepsilon n/2$ and thus $n^{\ell_1} |\mathscr{C}|^{\ell_2 - 1} \leq n^{\ell_1} |\mathscr{C}|^{\ell_2} \cdot \frac{2}{\varepsilon n}$.

Finally, we bound the size of $\mathscr{R}''(\ell_1, \ell_2) = \mathscr{R}(\ell_1, \ell_2) \setminus \mathscr{R}'(\ell_1, \ell_2)$. Since for any $(i_1, \ldots, i_{\ell_1}, (u_1, j_1), \ldots, (u_{\ell_2}, j_{\ell_2})) \in \mathscr{R}''(\ell_1, \ell_2)$ the sequence $(i_1, \ldots, i_{\ell_1}, u_1, \ldots, u_{\ell_2})$ is a proper relation and since there are no more than L^{ℓ_2} ways of choosing the indices $j_1, \ldots, j_{\ell_2}$, on the event $\mathscr{E}$ we have

$$\begin{aligned} |\mathscr{R}''(\ell_1, \ell_2)| &\leq \frac{\zeta^4}{L^\ell} \cdot L^{\ell_2} n^\ell && [\text{because } \boldsymbol{A}' \text{ is } \zeta^4/L^\ell, \ell)\text{-free}] \\ &\leq \zeta^4 \left(\frac{2}{\varepsilon}\right)^{\ell_2} \cdot n^{\ell_1} |\mathscr{C}|^{\ell_2} && [\text{because } |\mathscr{C}| > \varepsilon n/2] \\ &< \frac{\zeta}{2} n^{\ell_1} |\mathscr{C}|^{\ell_2}, \end{aligned} \tag{8.13}$$

provided that $\zeta < \zeta_0(\varepsilon, \ell)$ is sufficiently small. Thus, (8.11) follows from (8.12) and (8.13). □

Let $(\hat{\boldsymbol{k}}_i)_{i \geq 1}$ be a sequence of copies of $\hat{\boldsymbol{k}}$, mutually independent and independent of everything else. Also let

$$\hat{\boldsymbol{\gamma}}_j = \sum_{i=1}^{\boldsymbol{d}_{n+2}} \mathbb{1}\left\{\hat{\boldsymbol{k}}_i = j\right\}, \qquad \hat{\boldsymbol{\gamma}} = (\hat{\boldsymbol{\gamma}}_j)_{j \geq 1}.$$

Additionally, let $(\hat{\Delta}_j)_{j \geq 3}$ be a family of independent random variables with distribution

$$\hat{\boldsymbol{\Delta}}_j = \mathrm{Po}\left((1 - \varepsilon)\mathbb{P}\left[\boldsymbol{k} = j\right] d/k\right). \tag{8.14}$$

Further, let Σ' be the σ-algebra generated by $\boldsymbol{G}', \boldsymbol{A}', \boldsymbol{\theta}, \boldsymbol{\lambda}^-, \boldsymbol{M}^-, \boldsymbol{\Gamma}_{n, M^-}, (\boldsymbol{\chi}_{i,j,h}{}')_{i,j,h \geq 1}$ and $(\boldsymbol{d}_i)_{i \in [n]}$. In particular, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are Σ'-measurable.

Lemma 8.7. *With probability* $1 - \exp(-\Omega_\varepsilon(1/\varepsilon))$, *we have*

$$d_{TV}\big(\mathbb{P}\left[\{\boldsymbol{\gamma} \in \cdot\,\} \,|\, \Sigma'\right], \hat{\boldsymbol{\gamma}}\big) + d_{TV}\Big(\mathbb{P}\left[\{\Delta \in \cdot\,\} \,|\, \Sigma'\right], \hat{\boldsymbol{\Delta}}\Big) = O_\varepsilon(\sqrt{\varepsilon}).$$

Proof. Because $\boldsymbol{G}'$ is distributed the same as the Tanner graph from [10], apart from the extra ternary checks t_i, which do not affect the random vector $\boldsymbol{\gamma}$, the assertion follows from [10, Lemma 5.8]. □

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Let $\ell_* = \lceil \exp(1/\varepsilon^4) \rceil$ and $\delta_* = \exp(-1/\varepsilon^4)$ and consider the event

$$\mathscr{E} = \mathfrak{R}(\delta_*, \ell_*). \tag{8.15}$$

Further, consider the event

$$\mathscr{E}' = \left\{ |\mathscr{C}| \geq \varepsilon dn/2 \wedge \max_{i \leq n} \boldsymbol{d}_i \leq n^{1/2} \right\}. \tag{8.16}$$

Corollary 8.8. *For sufficiently large $\Theta = \Theta(\varepsilon) > 0$ we have $\mathbb{P}\left[\boldsymbol{A}' \in \mathscr{E} \right] > \exp(-1/\varepsilon^4)$. Moreover, $\mathbb{P}\left[\mathscr{E}' \right] = 1 - o(1)$.*

Proof. The first statement follows from Lemma 8.6. The second statement follows from the choice of the parameters in (8.1), Lemma 3.9 and Lemma 8.5. □

With these preparations in place we are ready to proceed to the proofs of Lemmas 8.3 and 8.4.

8.3 Proof of Lemma 8.3

Let

$$\boldsymbol{X} = \sum_{i \geq 1} \boldsymbol{\Delta}_i, \qquad \boldsymbol{Y} = \sum_{i \geq 1} i\boldsymbol{\Delta}_i, \qquad \boldsymbol{Y}' = \sum_{i \geq 1} i\boldsymbol{\gamma}_i.$$

Then the total number of new non-zero entries upon going from $\boldsymbol{A}'$ to $\boldsymbol{A}'''$ is bounded by $\boldsymbol{Y} + \boldsymbol{Y}' + 3\boldsymbol{\lambda}'''$. Let

$$\mathscr{E}'' = \left\{ X \vee Y \vee Y' \vee \boldsymbol{\lambda}''' \leq 1/\varepsilon \right\}.$$

Claim 8.9. *We have $\mathbb{P}\left[\mathscr{E}'' \right] = 1 - O_\varepsilon(\varepsilon)$.*

Proof. Apart from the additional ternary checks the argument is similar to [10, Proof of Claim 5.9]. The construction (8.1) ensures that $\mathbb{E}[\boldsymbol{X}], \mathbb{E}[\boldsymbol{Y}] = O_\varepsilon(1)$. Therefore, $\mathbb{P}\left[\boldsymbol{X} > 1/\varepsilon\right] = O_\varepsilon(\varepsilon)$, $\mathbb{P}\left[\boldsymbol{Y} > 1/\varepsilon\right] = O_\varepsilon(\varepsilon)$ by Markov's inequality. Further, a given check node of degree i is adjacent to x_{n+1} with probability at most $i\boldsymbol{d}_{n+1} / \sum_{i=1}^n \boldsymbol{d}_i \geq n \leq i\boldsymbol{d}_{n+1}/n$. Consequently,

$$\mathbb{E}\left[\boldsymbol{Y}' \right] = \mathbb{E} \sum_{i \geq 1} i\boldsymbol{\gamma}_i \leq \mathbb{E} \sum_{i \in [\boldsymbol{m}_{\varepsilon,n}^+]} \boldsymbol{k}_i^2 \boldsymbol{d}_{n+1}/n = O_\varepsilon(1).$$

Moreover, (8.10) shows that $\mathbb{E}[\boldsymbol{\lambda}'''] = O_\varepsilon(1)$. Thus, the assertion follows from Markov's inequality. □

We obtain $\boldsymbol{G}'''$ from $\boldsymbol{G}'$ by adding checks $a_{i,j}''', i \geq 1, j \in [\boldsymbol{\gamma}_i]$, $b_{i,j}''', i \geq 1, j \in [\boldsymbol{M}_i^+ - \boldsymbol{M}_i^- - \boldsymbol{\gamma}_i]$ and $t_i''', i \in [\boldsymbol{\lambda}''']$. Let

$$\mathscr{X}''' = \left(\bigcup_{i \geq 1} \bigcup_{j=1}^{\boldsymbol{\gamma}_i} \partial a_{i,j}''' \setminus \{x_{n+1}\} \right) \cup \left(\bigcup_{i \geq 1} \bigcup_{j \in [M_i^+ - M_i^- - \boldsymbol{\gamma}_i]} \partial b_{i,j}''' \right) \cup \bigcup_{i=1}^{\boldsymbol{\lambda}'''} \partial t_i''' \setminus \{x_{n+1}\}$$

be the set of variable neighbours of these new checks among $x_1, \ldots, x_n$. Further, let

$$\mathscr{E}''' = \left\{ |\mathscr{X}'''| = Y + \sum_{i \geq 1} (i-1)\boldsymbol{\gamma}_i + \boldsymbol{\lambda}''' \right\}$$

be the event that the variables of $\boldsymbol{G}'$ where the new checks connect are pairwise distinct.

Claim 8.10. *We have $\mathbb{P}\left[\mathscr{E}''' \mid \mathscr{E}' \cap \mathscr{E}'' \right] = 1 - o(1)$.*

Proof. By the same token as in [10, proof of Claim 5.10], given that $\mathscr{E}'$ occurs the total number of cavities comes to $\Omega(n)$. At the same time, the maximum variable node degree is of order

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

$O(\sqrt{n})$. Moreover, given the event $\mathscr{E}''$ no more than $\boldsymbol{Y}+\boldsymbol{Y}'=O_\varepsilon(1/\varepsilon)$ random cavities are chosen as neighbours of the new checks $a'''_{i,j}, b'''_{i,j}$. Thus, by the birthday paradox the probability that the checks $a'''_{i,j}, b'''_{i,j}$ occupy more than one cavity of any variable node is $o(1)$. Furthermore, the additional ternary nodes t'''_i choose their two neighbours among $x_1,\ldots,x_n$ mutually independently and independently of the $a'''_{i,j}, b'''_{i,j}$. Since $\boldsymbol{\lambda}'''$ is bounded given $1/\varepsilon$, the overall probability of choosing the same variable twice is $o(1)$. □

The following claim shows that the unlikely event that $\mathscr{E}\cap\mathscr{E}'\cap\mathscr{E}''\cap\mathscr{E}'''$ does not occur does not contributed significantly to the expected change in nullity.

Claim 8.11. *We have* $\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}''')-\mathrm{nul}(\boldsymbol{A}')\right|(1-\mathbb{1}\mathscr{E}\cap\mathscr{E}'\cap\mathscr{E}''\cap\mathscr{E}''')\right]=o_\varepsilon(1)$.

Proof. We modify the proof of [10, Claim 5.11] to accommodate the extra ternary nodes. Since $\boldsymbol{A}'''$ results from $\boldsymbol{A}'$ by adding one column and no more than $\boldsymbol{X}+\boldsymbol{d}_{n+1}+\boldsymbol{\lambda}'''$ rows, we have $\left|\mathrm{nul}(\boldsymbol{A}''')-\mathrm{nul}(\boldsymbol{A}')\right|\le X+\boldsymbol{d}_{n+1}+\boldsymbol{\lambda}'''+1$. Because $\boldsymbol{X},\boldsymbol{d}_{n+1}^2,\boldsymbol{\lambda}'''$ have bounded second moments, the Cauchy-Schwarz inequality therefore yields the estimate

$$\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}''')-\mathrm{nul}(\boldsymbol{A}')\right|(1-\mathbb{1}\mathscr{E}'')\right]\le\mathbb{E}\left[(X+\boldsymbol{d}_{n+1}+\boldsymbol{\lambda}'''+1)^2\right]^{1/2}\left(1-\mathbb{P}\left[\mathscr{E}''\right]\right)^{1/2}=o_\varepsilon(1). \tag{8.17}$$

Moreover, combining Corollary 8.8 and Claims 8.9–8.10, we obtain

$$\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}''')-\mathrm{nul}(\boldsymbol{A}')\right|\mathbb{1}\mathscr{E}''\setminus\mathscr{E}\right]\le O_\varepsilon(\varepsilon^{-1})\exp(-1/\varepsilon^4)=o_\varepsilon(1), \tag{8.18}$$

$$\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}''')-\mathrm{nul}(\boldsymbol{A}')\right|\mathbb{1}\mathscr{E}''\setminus\mathscr{E}'\right],\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}''')-\mathrm{nul}(\boldsymbol{A}')\right|\mathbb{1}\mathscr{E}''\cap\mathscr{E}'\setminus\mathscr{E}'''\right]=o(1). \tag{8.19}$$

The assertion follows from (8.17) to (8.19). □

Recall that $\boldsymbol{\alpha}$ denotes the fraction of frozen cavities and $\boldsymbol{\beta}$ the fraction of frozen variables of $\boldsymbol{A}'$. Further, let $\Sigma''\supset\Sigma'$ be the σ-algebra generated by $\boldsymbol{\theta}$, $\boldsymbol{G}'$, $\boldsymbol{A}'$, $\boldsymbol{M}_-$, $(\boldsymbol{d}_i)_{i\in[n+1]}$, $\boldsymbol{\gamma}$, $\boldsymbol{M}$, $\boldsymbol{\Delta}$, $\boldsymbol{\lambda}^-$, $\boldsymbol{\lambda}'''$. Then $\boldsymbol{\alpha},\boldsymbol{\beta}$ as well as $\mathscr{E},\mathscr{E}',\mathscr{E}''$ are Σ''-measurable but $\mathscr{E}'''$ is not.

Claim 8.12. *On the event* $\mathscr{E}\cap\mathscr{E}'\cap\mathscr{E}''$ *we have*

$$\mathbb{E}\left[\left(\mathrm{nul}(\boldsymbol{A}''')-\mathrm{nul}(\boldsymbol{A}')\right)\mathbb{1}\mathscr{E}'''\mid\Sigma''\right]=o_\varepsilon(1)+(1-\boldsymbol{\beta}^2)^{\boldsymbol{\lambda}'''}\prod_{i\ge1}(1-\boldsymbol{\alpha}^{i-1})^{\boldsymbol{\gamma}_i}-\sum_{i\ge1}(1-\boldsymbol{\alpha}^{i-1})\boldsymbol{\gamma}_i$$
$$-\boldsymbol{\lambda}'''(1-\boldsymbol{\beta}^2)-\sum_{i\ge1}(1-\boldsymbol{\alpha}^i)(\boldsymbol{M}_i^+-\boldsymbol{M}_i^--\boldsymbol{\gamma}_i).$$

Proof. We modify the proof of [10, Claim 5.12] by taking the additional ternary checks into consideration. Let

$$\mathscr{A}=\left\{a'''_{i,j}:i\ge1,\ j\in[\boldsymbol{\gamma}_i]\right\},\quad\mathscr{B}=\left\{b'''_{i,j}:i\ge1,\ j\in[\boldsymbol{M}_i^+-\boldsymbol{M}_i^--\boldsymbol{\gamma}_i]\right\},\quad\mathscr{T}=\left\{t_i:i\in[\boldsymbol{\lambda}''']\right\}.$$

We set up a random matrix $\boldsymbol{B}$ with rows indexed by $\mathscr{A}\cup\mathscr{B}\cup\mathscr{T}$ and columns indexed by $V_n=\{x_1,\ldots,x_n\}$. For a check $a\in\mathscr{A}\cup\mathscr{B}\cup\mathscr{T}$ and a variable $x\in V_n$ the (a,x)-entry of $\boldsymbol{B}$ equals zero unless $x\in\partial_{\boldsymbol{G}'''}a$. Further, the non-zero entries of $\boldsymbol{B}$ are independent copies of $\boldsymbol{\chi}$. Additionally, obtain $\boldsymbol{B}_*$ from $\boldsymbol{B}$ by zeroing out the x-column for every variable $x\in\mathfrak{F}(\boldsymbol{A}')$. Finally, let $\boldsymbol{C}\in\mathbb{F}^{\mathscr{A}\cup\mathscr{B}\cup\mathscr{T}}$ be a random vector whose entries $\boldsymbol{C}_a$, $a\in\mathscr{A}\cup\mathscr{T}$, are independent copies of $\boldsymbol{\chi}$, while $\boldsymbol{C}_b=0$ for all $b\in\mathscr{B}$.

If $\mathscr{E}'''$ occurs, $\boldsymbol{B}$ has row full rank because there is at most one non-zero entry in every column and at least one non-zero entry in every row. Hence,

$$\mathrm{rk}(\boldsymbol{B})=|\mathscr{A}\cup\mathscr{B}\cup\mathscr{T}|=\sum_{i\ge1}\boldsymbol{M}_i^+-\boldsymbol{M}_i^-+\boldsymbol{\lambda}'''.$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Furthermore, since the rank is invariant under row and column permutations, given $\mathscr{E}\cap\mathscr{E}'\cap\mathscr{E}''\cap\mathscr{E}'''$ we have

$$\mathrm{nul}\boldsymbol{A}'''=\mathrm{nul}\begin{pmatrix}\boldsymbol{A}' & 0\\ \boldsymbol{B} & \boldsymbol{C}\end{pmatrix}.$$

Moreover, given $\mathscr{E}'$ the set $\mathscr{X}'''$ of all non-zero columns of $\boldsymbol{B}$ satisfies $|\mathscr{X}'''|\leq \boldsymbol{Y}+\boldsymbol{Y}'+\boldsymbol{\lambda}'''\leq 3/\varepsilon$ while $|\mathscr{C}|\geq\varepsilon dn/2$. Therefore, the set of cavities that $\boldsymbol{\Gamma}'''$ occupies is within total variation distance $o(1)$ of a commensurate number of cavities drawn independently, that is, with replacement. Furthermore, the variables where the checks from $\mathscr{T}$ attach are chosen uniformly at random from $x_1,\ldots,x_n$. Listing the neighbours of $\mathscr{T}$ first and then the cavities chosen as neighbours of checks in $\mathscr{A}\cup\mathscr{B}$, the conditional probability that $\mathscr{X}'''$ forms a proper relation of $\boldsymbol{A}'$ can be upper bounded by the number of such choices that yield proper relations, divided by the total number of choices of variables and cavities. Given $\mathscr{E}'$, we had observed that $|\mathscr{X}'''|\leq 3/\varepsilon$. Moreover, on (8.15), for all $0\leq\ell_1,\ell_2\leq\ell_*=\lceil\exp(1/\varepsilon^4)\rceil$, the proportion of proper relations among all choices of ℓ_1 variables and ℓ_2 cavities is at most $\delta_*=\exp(-1/\varepsilon^4)$. Therefore, on $\mathscr{E}\cap\mathscr{E}'\cap\mathscr{E}''$ the conditional probability given $\mathscr{E}'''$ that $\mathscr{X}'''$ forms a proper relation is bounded by $O_\varepsilon(\exp(-1/\varepsilon^4))$. Consequently, Lemma 3.2 implies that on the event $\mathscr{E}\cap\mathscr{E}'\cap\mathscr{E}''$,

$$\mathbb{E}\left[\left(\mathrm{nul}(\boldsymbol{A}''')-\mathrm{nul}(\boldsymbol{A}')\right)\mathbb{1}\mathscr{E}'''\mid\Sigma''\right]=1-\mathbb{E}\left[\mathrm{rk}\,(\boldsymbol{B}_*\ \boldsymbol{C})\mid\Sigma''\right]+o_\varepsilon(1).\tag{8.20}$$

We are thus left to calculate the rank of $\boldsymbol{Q}=(\boldsymbol{B}_*\ \boldsymbol{C})$. Given $\mathscr{E}'''$ this block matrix decomposes into the $\mathscr{A}\cup\mathscr{T}$-rows $\boldsymbol{Q}_{\mathscr{A}\cup\mathscr{T}}$ and the $\mathscr{B}$-rows $\boldsymbol{Q}_{\mathscr{B}}$ such that $\mathrm{rk}(\boldsymbol{Q})=\mathrm{rk}(\boldsymbol{Q}_{\mathscr{A}\cup\mathscr{T}})+\mathrm{rk}(\boldsymbol{Q}_{\mathscr{B}})$. Therefore, it suffices to prove that

$$\mathbb{E}\left[\mathrm{rk}\,(\boldsymbol{Q}_{\mathscr{B}})\mid\Sigma''\right]=\sum_{i\geq1}\left(1-\boldsymbol{\alpha}^i\right)(\boldsymbol{M}_i^+-\boldsymbol{M}_i^--\boldsymbol{\gamma}_i)+o(1),\tag{8.21}$$

$$\mathbb{E}\left[\mathrm{rk}(\boldsymbol{Q}_{\mathscr{A}\cup\mathscr{T}})\mid\Sigma''\right]=\lambda'''(1-\boldsymbol{\beta}^2)+\sum_{i\geq1}\left(1-\boldsymbol{\alpha}^{i-1}\right)\boldsymbol{\gamma}_i$$
$$+1-(1-\boldsymbol{\beta}^2)^{\boldsymbol{\lambda}'''}\prod_{i\geq1}\left(1-\boldsymbol{\alpha}^{i-1}\right)^{\boldsymbol{\gamma}_i}+o(1).\tag{8.22}$$

Towards (8.21) consider a check $b\in\mathscr{B}$ whose corresponding row sports i non-zero entries. Recall that the fraction $\boldsymbol{\alpha}$ of frozen cavities of $\boldsymbol{A}'$ is Σ''-measurable and can thus be regarded as constant. Moreover, we may pretend (up to $o(1)$ in total variation) that these i entries are drawn uniformly and independently from the set of cavities, so that the probability that these i independent and uniform draws all hit frozen cavities comes to $\boldsymbol{\alpha}^i+o(1)$. We emphasise that this calculation only requires the draws to be independent and uniform, but makes no assumption on the underlying dependencies between cavities. Since there are $\boldsymbol{M}_i^+-\boldsymbol{M}_i^--\boldsymbol{\gamma}_i$ such checks $b\in\mathscr{B}$, we obtain (8.21).

Moving on to (8.22), consider $a\in\mathscr{A}$ whose corresponding row has $i-1$ non-zero entries, and recall that $V_n=\{x_1,\ldots,x_n\}$. By the same token as in the previous paragraph, the probability that all entries in the a-row correspond to frozen cavities comes to $\boldsymbol{\alpha}^{i-1}+o(1)$. Hence, the expected rank of the $\mathscr{A}\times V_n$-minor works out to be $\sum_{i\geq1}\left(1-\boldsymbol{\alpha}^{i-1}\right)\boldsymbol{\gamma}_i+o(1)$, which is the second summand in (8.22). Similarly, a $t\in\mathscr{T}$-row adds to the rank unless both the variables in the corresponding check are frozen. The latter event occurs with probability $\boldsymbol{\beta}^2$. Hence the first summand. Finally, the $\boldsymbol{C}$-column adds to the rank if none of the $\mathscr{A}\cup\mathscr{T}$-rows become all zero, which occurs with probability $(1-\boldsymbol{\beta}^2)^{\boldsymbol{\lambda}'''}\prod_{i\geq1}\left(1-\boldsymbol{\alpha}^{i-1}\right)^{\boldsymbol{\gamma}_i}+o(1)$. □

Proof of Lemma 8.3. Letting $\mathfrak{E} = \mathscr{E} \cap \mathscr{E}' \cap \mathscr{E}'' \cap \mathscr{E}'''$ and combining Claims 8.9–8.12, we obtain

$$\mathbb{E}\left|\mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}''') - \mathrm{nul}(\boldsymbol{A}') \mid \Sigma''\right] - \Big((1-\boldsymbol{\beta}^2)^{\boldsymbol{\lambda}'''}\prod_{i\geq 1}(1-\boldsymbol{\alpha}^{i-1})^{\boldsymbol{\gamma}_i} - \sum_{i\geq 1}(1-\boldsymbol{\alpha}^{i-1})\boldsymbol{\gamma}_i - \sum_{i\geq 1}(1-\boldsymbol{\alpha}^{i})(\boldsymbol{M}_i^+ - \boldsymbol{M}_i^- - \boldsymbol{\gamma}_i) - \boldsymbol{\lambda}'''(1-\boldsymbol{\beta}^2)\Big)\mathbb{1}\mathfrak{E}\right| = o_\varepsilon(1). \tag{8.23}$$

On $\mathfrak{E}$ all i with $\boldsymbol{M}_i^+ - \boldsymbol{M}_i^- - \boldsymbol{\gamma}_i > 0$ are bounded. Moreover, w.h.p. we have $\boldsymbol{M}_i \sim \mathbb{E}[\boldsymbol{M}_i] = \Omega(n)$ for all bounded i by Chebyshev's inequality. Hence, (8.7) implies that $\boldsymbol{M}_i^- = \boldsymbol{M}_i - \boldsymbol{\gamma}_i$ w.h.p. Consequently, (8.23) becomes

$$\mathbb{E}\left|\mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}''') - \mathrm{nul}(\boldsymbol{A}') \mid \Sigma''\right] - \Big((1-\boldsymbol{\beta}^2)^{\boldsymbol{\lambda}'''}\prod_{i\geq 1}(1-\boldsymbol{\alpha}^{i-1})^{\boldsymbol{\gamma}_i} - \sum_{i\geq 1}(1-\boldsymbol{\alpha}^{i-1})\boldsymbol{\gamma}_i - \sum_{i\geq 1}(1-\boldsymbol{\alpha}^{i})\Delta_i - \boldsymbol{\lambda}'''(1-\boldsymbol{\beta}^2)\Big)\mathbb{1}\mathfrak{E}\right| = o_\varepsilon(1). \tag{8.24}$$

We proceed to estimate the various terms on the r.h.s. of (8.24) separately. Since $\mathbb{P}\left[\mathfrak{E}\right] = 1 - o_\varepsilon(1)$ by Corollary 8.8 and Claims 8.9 and 8.10, Lemma 8.7 yield

$$\begin{aligned}\mathbb{E}&\left[\mathbb{1}\mathfrak{E}\cdot(1-\boldsymbol{\beta}^2)^{\boldsymbol{\lambda}'''}\prod_{i\geq 1}(1-\boldsymbol{\alpha}^{i-1})^{\boldsymbol{\gamma}_i} \mid \Sigma''\right]\\ &= \mathbb{E}\left[(1-\boldsymbol{\beta}^2)^{\boldsymbol{\lambda}'''}\prod_{i\geq 1}(1-\boldsymbol{\alpha}^{i-1})^{\hat{\boldsymbol{\gamma}}_i} \mid \Sigma''\right] + o_\varepsilon(1)\\ &= \exp(-3\delta\boldsymbol{\beta}^2)D(1-K'(\boldsymbol{\alpha})/k) \qquad \text{[by (3.2) and (8.10)]}.\end{aligned} \tag{8.25}$$

Moreover, since $\sum_{i\geq 1}\boldsymbol{\gamma}_i \leq \boldsymbol{d}_{n+1}$ and $\boldsymbol{d}_{n+1}$ has a bounded second moment, Lemma 8.7 implies that

$$\mathbb{E}\left[\mathbb{1}\mathfrak{E}\cdot\sum_{i\geq 1}(1-\boldsymbol{\alpha}^{i-1})\boldsymbol{\gamma}_i \mid \Sigma''\right] = \mathbb{E}\left[\sum_{i\geq 1}(1-\boldsymbol{\alpha}^{i-1})\hat{\boldsymbol{\gamma}}_i \mid \Sigma''\right] + o_\varepsilon(1) = d - \frac{d}{k}K'(\boldsymbol{\alpha}) + o_\varepsilon(1). \tag{8.26}$$

Further, by Claim 8.9, Lemma 8.7 and (8.14),

$$\mathbb{E}\left[\mathbb{1}\mathfrak{E}\cdot\sum_{i\geq 1}(1-\boldsymbol{\alpha}^{i})\boldsymbol{\Delta}_i \mid \Sigma''\right] = \mathbb{E}\left[\sum_{i\geq 1}(1-\boldsymbol{\alpha}^{i})\boldsymbol{\Delta}_i \mid \Sigma''\right] + o_\varepsilon(1) = o_\varepsilon(1) + \frac{d}{k} - \frac{d}{k}\mathbb{E}[K(\boldsymbol{\alpha})]. \tag{8.27}$$

Finally, (8.10) yields

$$\mathbb{E}\left[\mathbb{1}\mathfrak{E}\cdot\boldsymbol{\lambda}'''(1-\boldsymbol{\beta}^2) \mid \Sigma''\right] = 3\delta(1-\boldsymbol{\beta}^2) + o_\varepsilon(1). \tag{8.28}$$

Thus, the assertion follows from (8.24)–(8.28). □

8.4 Proof of Lemma 8.4

We proceed similarly as in the proof of Lemma 8.3; actually matters are a bit simpler because we only add checks, while in the proof of Lemma 8.3 we also had to deal with the extra variable node

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

x_{n+1}. Let $\mathcal{E}, \mathcal{E}'$ be the events from (8.15) and (8.16) and let $\mathcal{E}'' = \left\{\boldsymbol{d}_{n+1} + \boldsymbol{\lambda}'' \leq 1/\varepsilon\right\}$. As a direct consequence of the assumption $\mathbb{E}[\boldsymbol{d}_{n+1}^2] = O_{\varepsilon,n}(1)$ and of (8.9), we obtain the following.

Fact 8.13. *We have* $\mathbb{P}\left[\mathcal{E}''\right] = 1 - O_\varepsilon(\varepsilon^2)$.

Let

$$\mathcal{X}'' = \bigcup_{i\geq 1} \bigcup_{j\in[\boldsymbol{M}_i - \boldsymbol{M}_i^-]} \partial_{\boldsymbol{G}''} a''_{i,j} \cup \bigcup_{i=1}^{\boldsymbol{\lambda}''} \partial t''_i$$

be the set of variable nodes where the new checks that we add upon going from $\boldsymbol{A}'$ to $\boldsymbol{A}''$ attach. Let $\mathcal{E}'''$ be the event that in $\boldsymbol{G}''$ no variable from $\mathcal{X}''$ is connected with the checks $\{a''_{i,j} : i \geq 1, j \in [\boldsymbol{M}_i - \boldsymbol{M}_i^-]\} \cup \{t''_i : i \in [\boldsymbol{\lambda}'']\}$ by more than one edge.

Claim 8.14. *We have* $\mathbb{P}\left[\mathcal{E}''' \mid \mathcal{E}' \cap \mathcal{E}''\right] = 1 - o(1)$.

Proof. This follows from the 'birthday paradox' (see the proof of Claim 8.10). □

Claim 8.15. *We have* $\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}')\right| (1 - \mathbb{1}\mathcal{E} \cap \mathcal{E}' \cap \mathcal{E}'' \cap \mathcal{E}''')\right] = o_\varepsilon(1)$.

Proof. We have $\left|\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}')\right| \leq \boldsymbol{d}_{n+1} + \boldsymbol{\lambda}''$ as we add at most $\boldsymbol{d}_{n+1} + \boldsymbol{\lambda}''$ rows. Because $\mathbb{E}[(\boldsymbol{d}_{n+1} + \boldsymbol{\lambda}'')^2] = O_\varepsilon(1)$ by (8.9), Claim 8.13 and the Cauchy-Schwarz inequality yield

$$\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}')\right| (1 - \mathbb{1}\mathcal{E}'')\right] \leq \mathbb{E}\left[(\boldsymbol{d}_{n+1} + \boldsymbol{\lambda}'')^2\right]^{1/2} (1 - \mathbb{P}\left[\mathcal{E}\right])^{1/2} = o_\varepsilon(1). \tag{8.29}$$

Moreover, Corollary 8.8 and Claim 8.14 show that

$$\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}')\right| \mathbb{1}\mathcal{E}'' \setminus \mathcal{E}\right], \mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}')\right| \mathbb{1}\mathcal{E}'' \setminus \mathcal{E}'\right],$$
$$\mathbb{E}\left[\left|\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}')\right| \mathbb{1}\mathcal{E}'' \setminus \mathcal{E}'''\right] = o_\varepsilon(1). \tag{8.30}$$

The assertion follows from (8.29) and (8.30). □

The matrix $\boldsymbol{A}''$ results from $\boldsymbol{A}'$ by adding checks $a''_{i,j}$, $i \geq 1, j \in [\boldsymbol{M}_i - \boldsymbol{M}_i^-]$ that are connected to random cavities of $\boldsymbol{A}'$.

Moreover, as before let $\Sigma'' \supset \Sigma'$ be the σ-algebra generated by $\boldsymbol{\theta}$, $\boldsymbol{G}'$, $\boldsymbol{A}'$, $\boldsymbol{M}_-$, $(\boldsymbol{d}_i)_{i\in[n+1]}$, $\boldsymbol{\gamma}$, $\boldsymbol{M}$, $\boldsymbol{\Delta}$, $\boldsymbol{\lambda}^-$, $\boldsymbol{\lambda}'''$. Then $\mathcal{E}, \mathcal{E}', \mathcal{E}''$ are Σ''-measurable, but $\mathcal{E}'''$ is not.

Claim 8.16. *On* $\mathcal{E} \cap \mathcal{E}' \cap \mathcal{E}''$ *we have*

$$\mathbb{E}\left[(\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}'))\mathbb{1}\mathcal{E}''' \mid \Sigma''\right] = o_\varepsilon(1) - \sum_{i\geq 1} (1 - \boldsymbol{\alpha}^i)(M_i - M_i^-) - \boldsymbol{\lambda}''(1 - \boldsymbol{\beta}^3).$$

Proof. Let $\mathcal{A} = \{a''_{i,j} : i \geq 1,\ j \in [\boldsymbol{M}_i - \boldsymbol{M}_i^-]\}$. Moreover, let $\mathcal{T}$ be the set of new ternary checks t''_i, $i \in [\boldsymbol{\lambda}'']$. Let $\boldsymbol{B}$ be the $\mathbb{F}_q$-matrix whose rows are indexed by $\mathcal{A} \cup \mathcal{T}$ and whose columns are indexed by $V_n = \{x_1, \ldots, x_n\}$. The (a, x)-entry of $\boldsymbol{B}$ is zero unless a, x are adjacent in $\boldsymbol{G}''$, in which case the entry is an independent copy of $\boldsymbol{\chi}$. Given $\mathcal{E}'''$ the matrix $\boldsymbol{B}$ has full row rank $\mathrm{rk}(\boldsymbol{B}) = |\mathcal{A}| = \boldsymbol{\lambda}'' + \sum_{i\geq 1} \boldsymbol{M}_i^+ - \boldsymbol{M}_i$, because no column contains two non-zero entries and each row has at least one non-zero entry. Further, obtain $\boldsymbol{B}_*$ from $\boldsymbol{B}$ by zeroing out the x-column of every $x \in \mathfrak{F}(\boldsymbol{A}')$.

On $\mathcal{E} \cap \mathcal{E}' \cap \mathcal{E}'' \cap \mathcal{E}'''$ we see that

$$\mathrm{nul}\boldsymbol{A}'' = \mathrm{nul}\begin{pmatrix} \boldsymbol{A}' \\ \boldsymbol{B} \end{pmatrix}. \tag{8.31}$$

Moreover, let $\mathcal{I}$ be the set of non-zero columns of $\boldsymbol{B}$. Then on $\mathcal{E}' \cap \mathcal{E}''$ we have $|\mathcal{I}| \leq \boldsymbol{d}_{n+1} + \boldsymbol{\lambda}'' \leq 1/\varepsilon$. Hence, on $\mathcal{E} \cap \mathcal{E}' \cap \mathcal{E}'' \cap \mathcal{E}'''$ the probability that $\mathcal{I}$ forms a proper relation is bounded by $\exp(-1/\varepsilon^4)$. Hence, Lemma 3.2 shows

$$\mathbb{E}\left[\left(\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}')\right)\mathbb{1}\mathscr{E}''' \mid \Sigma''\right] = o_\varepsilon(1) - \mathbb{E}\left[\mathrm{rk}\left(\boldsymbol{B}_*\right) \mid \Sigma''\right]. \tag{8.32}$$

We are thus left to calculate the rank of $\boldsymbol{B}_*$. Recalling that $\boldsymbol{\alpha}$ stands for the fraction of frozen cavities, we see that for $a \in \mathscr{A}$ of degree i the a-row is all zero in $\boldsymbol{B}_*$ with probability $\boldsymbol{\alpha}^i + o(1)$. Similarly, for $a \in \mathscr{T}$ the a-row of $\boldsymbol{B}$ gets zeroed out with probability $\boldsymbol{\beta}^3$. Hence, we conclude that

$$\mathbb{E}\left[\mathrm{rk}\left(\boldsymbol{B}_*\right) \mid \Sigma''\right] = o_\varepsilon(1) + \boldsymbol{\lambda}''(1-\boldsymbol{\beta}^3) + \sum_{i\geq 1}\left(1-\boldsymbol{\alpha}^i\right)(\boldsymbol{M}_i - \boldsymbol{M}_i^-). \tag{8.33}$$

Combining (8.32) and (8.33) completes the proof. □

Proof of Lemma 8.4. Let $\mathfrak{E} = \mathscr{E} \cap \mathscr{E}' \cap \mathscr{E}'' \cap \mathscr{E}'''$. Combining Claims 8.15–8.16, we see that

$$\mathbb{E}\left|\mathbb{E}[\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}') \mid \Sigma''] + \left(\boldsymbol{\lambda}''(1-\boldsymbol{\beta}^3) + \sum_{i\geq 1}(1-\boldsymbol{\alpha}^i)(\boldsymbol{M}_i - \boldsymbol{M}_i^-)\right)\mathbb{1}\mathfrak{E}\right| = o_\varepsilon(1). \tag{8.34}$$

On $\mathfrak{E}$ all degrees i with $\boldsymbol{M}_i^+ - \boldsymbol{M}_i^- > 0$ are bounded. Moreover, $\boldsymbol{M}_i^- = \Omega(n)$ w.h.p. for every bounded i by Chebyshev's inequality. Therefore, (8.7) shows that $\boldsymbol{M}_i - \boldsymbol{M}_i^- = \boldsymbol{\gamma}_i$ for all i with $\boldsymbol{M}_i^+ - \boldsymbol{M}_i^- > 0$ w.h.p. Hence, (8.34) turns into

$$\mathbb{E}\left|\mathbb{E}[\mathrm{nul}(\boldsymbol{A}'') - \mathrm{nul}(\boldsymbol{A}') \mid \Sigma''] + \left(\boldsymbol{\lambda}''(1-\boldsymbol{\beta}^3) + \sum_{i\geq 1}(1-\boldsymbol{\alpha}^i)\boldsymbol{\gamma}_i\right)\mathbb{1}\mathfrak{E}\right| = o_\varepsilon(1). \tag{8.35}$$

We now estimate the two parts of the last expression separately. Since $\mathbb{P}[\mathfrak{E}] = 1 - o_\varepsilon(1)$ by Corollary 8.8, Fact 8.13 and Claim 8.14, the definition (8.9) of $\boldsymbol{\lambda}''$ yields

$$\mathbb{E}\left|\boldsymbol{\lambda}''(1-\boldsymbol{\beta}^3)\mathbb{1}\mathfrak{E}\right| = 2\delta(1 - \mathbb{E}[\boldsymbol{\beta}^3]) + o_\varepsilon(1). \tag{8.36}$$

Moreover, because $\sum_{i\geq 1}\boldsymbol{\gamma}_i \leq \boldsymbol{d}_{n+1}$, $\mathbb{E}[\boldsymbol{d}_{n+1}] = O_\varepsilon(1)$,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i\geq 1}(1-\boldsymbol{\alpha}^i)\boldsymbol{\gamma}_i\mathbb{1}\mathfrak{E}\right] \\
&= \mathbb{E}\left[\sum_{i\geq 1}(1-\boldsymbol{\alpha}^i)\hat{\boldsymbol{\gamma}}_i\mathbb{1}\left\{\boldsymbol{\lambda}'' + \sum_{i\geq 1}\hat{\boldsymbol{\gamma}}_i \leq \varepsilon^{-1/4}\right\}\right] + o_\varepsilon(1) \quad \text{[by Lemma 8.7 and Claim 8.13]} \\
&= d\mathbb{E}[1-\boldsymbol{\alpha}^{\hat{\boldsymbol{k}}}] + o_\varepsilon(1) = d - d\mathbb{E}[\boldsymbol{\alpha}K'(\boldsymbol{\alpha})]/k + o_\varepsilon(1) \qquad \text{[by (3.2)]}.
\end{aligned} \tag{8.37}$$

Combining (8.36) and (8.37) completes the proof. □

8.5 *Proof of Lemma 8.2*

The proof is relatively straightforward, not least because once again we can reuse some technical statements from [10]. Let us deal with $\boldsymbol{A}''$ and $\boldsymbol{A}'''$ separately.

Claim 8.17. *We have* $\mathbb{E}[\mathrm{nul}(\boldsymbol{A}'')] = \mathbb{E}[\mathrm{nul}(\boldsymbol{A}[n, \boldsymbol{M}, \boldsymbol{\lambda}])] + o(1)$.

Proof. The matrix models $\mathbb{E}[\mathrm{nul}(\boldsymbol{A}[n, \boldsymbol{M}, \boldsymbol{\lambda}])]$ and $\boldsymbol{A}''$ coincide with the corresponding models from [10, Claim 5.17], except that here we add extra ternary checks. Because these extra checks are added independently, the coupling from [10, Claim 5.17] directly induces a coupling of the enhanced models by attaching the same number $\boldsymbol{\lambda}''$ of ternary equations to the same neighbours. □

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Claim 8.18. *We have* $\mathbb{E}[\text{nul}(\boldsymbol{A}''')] = \mathbb{E}[\text{nul}(\boldsymbol{A}[n+1, \boldsymbol{M}^+, \boldsymbol{\lambda}^+])] + o(1)$.

Proof. The matrix models $\mathbb{E}[\text{nul}(\boldsymbol{A}[n+1, \boldsymbol{M}^+, \boldsymbol{\lambda}^+])]$ and $\boldsymbol{A}'''$ coincide with the corresponding models from [10, Section 5.5] plus the extra independent ternary equations. Hence, the coupling from [10, Claim 5.17] yields a coupling of the enhanced models just as in Claim 8.17. □

Proof of Lemma 8.2. The lemma is an immediate consequence of Claims 8.17 and 8.18. □

References

[1] Achlioptas, D. and McSherry, F. (2001) Fast computation of low-rank matrix approximations. *J. ACM* **54** 2007–2009.
[2] Achlioptas, D. and Molloy, M. (2015) The solution space geometry of random linear equations. *Random Struct. Algor.* **46**(2) 197–231.
[3] Achlioptas, D. and Moore, C. (2006) Random k-SAT: two moments suffice to cross a sharp threshold. *SIAM J. Comput.* **36** 740–762.
[4] Achlioptas, D., Naor, A., Peres, Y. (2005) Rigorous location of phase transitions in hard optimization problems. *Nature* **435**(7043) 759–764.
[5] Aizenman, M., Sims, R. and Starr, S. (2003) An extended variational principle for the SK spin-glass model. *Phys. Rev. B* **68**(21) 214403.
[6] Ayre, P., Coja-Oghlan, A., Gao, P. and Müller, N. (2020) The satisfiability threshold for random linear equations. *Combinatorica* **40**(2) 179–235.
[7] Balakin, G. (1968) The distribution of random matrices over a finite field. *Theory Probab. Appl.* **13**(4) 631–641.
[8] Blömer, J., Karp, R. and E., Welzl (1997) The rank of sparse random matrices over finite fields. *Random Struct. Algor.* **10** 407–419.
[9] Bordenave, C., Lelarge, M. and Salez, J. (2011) The rank of diluted random graphs. *Ann. Probab.* **39**(3) 1097–1121.
[10] Coja-Oghlan, A., Ergür, A., Gao, P., Hetterich, S. and Rolvien, M. (2023) The rank of sparse random matrices. *Random Struct. Algor.* **62**(1) 68–130.
[11] Coja-Oghlan, A., Krzakala, F., Perkins, W. and Zdeborova, L. (2018) Information-theoretic thresholds from the cavity method. *Adv. Math.* **333** 694–795.
[12] Coja-Oghlan, A. and K, 2016) Panagiotou: The asymptotic k-SAT threshold. *Adv. Math.* **288** 985–1068.
[13] Coja-Oghlan, A. and K. (2012) Panagiotou: Catching the k-NAESAT threshold, 899–908, *Proc. 44th STOC.*
[14] Coja-Oghlan, A., Müller, N. and Ravelomanana, J. (2022) Belief Propagation on the random k-SAT model. *Ann. Appl. Probab.* **32**(5) 3718–3796.
[15] Cooper, C., Frieze, A. and Pegden, W. (2019) On the rank of a random binary matrix. *Electron. J. Comb.* **26** P4.12.
[16] Costello, K. and V (2008) Vu: The rank of random graphs. *Random Struct. Algor.* **33**(3) 269–285.
[17] Costello, K. and Vu, V. (2010) On the rank of random sparse matrices. *Comb. Probab. Comput.* **19**(3) 321–342.
[18] Davis, B. and McDonald, D. (1995) An elementary proof of the local limit theorem. *J. Theor. Probab.* **8**(3) 693–701.
[19] Dietzfelbinger, M., Goerdt, A., Mitzenmacher, M., Montanari, A., Pagh, R. and Rink, M. (2010) Tight thresholds for cuckoo hashing via XORSAT, 213–225, *Proc. 37th ICALP*
[20] Ding, J., Sly, A. and N. (2015) Sun: Proof of the satisfiability conjecture for large k, 59–68, *Proc. 47th STOC*
[21] Dubois, O. and Mandler, J. (2002) The 3-XORSAT threshold, 769–778, *Proc. 43rd FOCS*
[22] Ferber, A., Kwan, M., Sah, A. and Sawhney, M. (2023) Singularity of the k-core of a random graph, arXiv: 2106.05719.
[23] Goerdt, A. and Falke, L. (2012) Satisfiability thresholds beyond k-XORSAT, 148–159, *Proc. 7th International Computer Science Symposium in Russia*
[24] Glasgow, M., Kwan, M., Sah, A. and Sawhney, M. The Exact Rank of Sparse Random Graphs, arXiv: 2303.05435.
[25] van der Hofstad, R. (2016) *Random graphs and complex networks. Vol. 43.* Cambridge University Press.
[26] van der Hofstad, R., Müller, N. and Zhu, H. The rank of sparse symmetric matrices over arbitrary fields, arXiv: 2301.12978.
[27] Huang, J. (2021) Invertibility of adjacency matrices for random d-regular graphs, arXiv: 1807.06465.
[28] Ibrahimi, M., Kanoria, Y., Kraning, M. and Montanari, A. (2015) The set of solutions of random XORSAT formulae. *Ann Appl Probab* **25**(5) 2743–2808.
[29] Janson, S. (2009) The probability that a random multigraph is simple. *Comb. Prob. Comp.* **18**(1-2) 205–225.
[30] Kahn, J., Komloś, J. and Szemerédi, E. (1995) On the probability that a random ± 1-matrix is singular. *J. AMS* **8** 223–240.
[31] Kolchin, V. (1995) Random graphs and systems of linear equations in finite fields. *Random Struct. Algor.* **5** 425–436.
[32] Kolchin, V. and Khokhlov, V. (1992) On the number of cycles in a random non-equiprobable graph. *Discrete Math. Appl.* **2**(1) 109–118.
[33] Kolchin, V. (1993) Consistency of a system of random congruences. *Discrete Math. Appl.* **3**(1) 103–113.
[34] Komlós, J. (1967) On the determinant of (0,1) matrices. *Studia Sci. Math. Hungar.* **2** 7–21.
[35] Kovalenko, I. (1967) On the limit distribution of the number of solutions of a random system of linear equations in the class of Boolean functions. *Theory Probab. Appl.* **12**(1) 51–61.
[36] Kovalenko, I., Levitskaya, A. and Savchuk, M. (1986) Selected problems of probabilistic combinatorics. Naukova Dumka, Kiev.

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

[37] Lenstra, H. (2008) Lattices. In *Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography* (Buhler, J. and Stevenhagen, P., eds), Cambridge University Press, pp. 127–181.
[38] Levitskaya, A. (1982) Theorems on invariance for the systems of random linear equations over an arbitrary finite ring. *Soviet Math. Dokl.* **263** 289–291.
[39] Levitskaya, A. (1985) The probability of consistency of a system of random linear equations over a finite ring. *Theory Probab. Appl.* **30** 339–350.
[40] Mézard, M. and A (2009) *Montanari: Information, Physics and Computation*. Oxford University Press.
[41] Mézard, M., Ricci-Tersenghi, F. and R, 2003) Zecchina: Two solutions to diluted p-spin models and XORSAT problems. *J. Stat. Phys.* **111** 505–533.
[42] Miller, G. and Cohen, G. (2003) The rate of regular LDPC codes. *IEEE Trans. Inform. Theory* **49**(11) 2989–2992.
[43] Montanari, A. (2008) Estimating random variables from random sparse observations. *Eur. Trans. Telecommun.* **19**(4) 385–403.
[44] Pittel, B. and G, 2016) Sorkin: The satisfiability threshold for k-XORSAT. *Comb. Prob. Comput.* **25** 236–268.
[45] Raič, M. (2019) A multivariate Berry-Esseen theorem with explicit constants. *Bernoulli* **25**(4A) 2824–2853.
[46] Raghavendra, P. and N. (2012) Tan: Approximating CSPs with global cardinality constraints using SDP hierarchies, pp. 373–387, *Proc. 23rd SODA*
[47] Richardson, T. and R (2008) *Urbanke: Modern Coding Theory*. Cambridge University Press.
[48] Tao, T. and Vu, V. (2007) On the singularity probability of random Bernoulli matrices. *J. AMS* **20**(3) 603–628.
[49] Tikhomirov, K. (2020) Singularity of random Bernoulli matrices. *Ann. Math.* **191**(2) 593–634.
[50] Vu, V. (2014) Combinatorial problems in random matrix theory, pp. 489–508, *Proc. International Congress of Mathematicians IV*.
[51] Vu, V. (2005) Recent progress in combinatorial random matrix theory, arXiv: 2005.02797.
[52] Wainwright, M., Maneva, E. and Martinian, E. (2010) Lossy source compression using low-density generator matrix codes: analysis and algorithms. *IEEE Trans. Inf. Theory* **56**(3) 1351–1368.

A. Appendix

In this appendix we give a self-contained proof of Claim 7.17, the local limit theorem for sums of independent vectors. We employ a simplified version of the strategy of the proof of Lemma 7.10. Recall that the degree sequences $(d_1, \ldots, d_n)$ and $(k_1, \ldots, k_m)$ satisfy (P1)-(P7) and the notation $\Delta_2 = \sum_{i=1}^n d_i^2$. Finally, we set

$$s_n := \sqrt{\Delta_2}. \tag{A.1}$$

As in the proof of Lemma 7.17, given $\omega > 0$, we choose $\varepsilon_0 = \varepsilon_0(\omega, q)$ sufficiently small and let $0 < \varepsilon < \varepsilon_0$. With these parameters, we set

$$\mathfrak{L}_0 = \left\{ r \in \mathbb{Z}^{\mathbb{F}_q^*} : \mathbb{P}_{\mathfrak{A}}\left(\hat{\rho}_\sigma = r\right) > 0 \text{ and } \left\| r - q^{-1}\Delta\mathbb{1} \right\|_1 < \omega n^{-1/2}\Delta \right\} \quad \text{and}$$

$$\mathfrak{L}_0(r_*, \varepsilon) = \{ r \in \mathfrak{L}_0 : \| r - r_* \|_\infty < \varepsilon s_n \}.$$

Then

$$\mathfrak{L}_0 \subseteq \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}.$$

We begin by observing that the vector $\hat{\rho}_\sigma$ is asymptotically normal given $\mathfrak{A}$. As before we let $\boldsymbol{I}_{(q-1)\times(q-1)}$ the $(q-1)\times(q-1)$-identity matrix and let $\boldsymbol{N} \in \mathbb{R}^{\mathbb{F}_q^*}$ be a Gaussian vector with zero mean and covariance matrix

$$\mathscr{C} = q^{-1}\boldsymbol{I}_{(q-1)\times(q-1)} - q^{-2}\mathbb{1}_{(q-1)\times(q-1)}. \tag{A.2}$$

Claim A.1. *There exists a function $\iota = \iota(n, q) = o(1)$ such that for all axis-aligned cubes $U \subseteq \mathbb{R}^{\mathbb{F}_q^*}$ we have*

$$\left| \mathbb{P}_{\mathfrak{A}}\left[\frac{\hat{\rho}_\sigma - q^{-1}\Delta\mathbb{1}}{s_n} \in U \right] - \mathbb{P}\left[N \in U \right] \right| \le \iota.$$

Proof. Given $\mathfrak{A}$, the mean of $\hat{\rho}_{\boldsymbol{\sigma}}(\tau)$ clearly equals Δ/q for every $\tau \in \mathbb{F}_q^*$. Concerning the covariance matrix, for distinct $s \neq t$ we obtain

$$\mathbb{E}_{\mathfrak{A}}[\hat{\rho}_{\boldsymbol{\sigma}}^2(s)] = \sum_{i,j\in[n]:i\neq j} \frac{d_i d_j}{q^2} + \sum_{i=1}^{n} \frac{d_i^2}{q} = \sum_{i,j=1}^{n} \frac{d_i d_j}{q^2} + \sum_{i=1}^{n} \frac{d_i^2}{q}\left(1 - \frac{1}{q}\right),$$

$$\mathbb{E}_{\mathfrak{A}}[\hat{\rho}_{\boldsymbol{\sigma}}(s)\hat{\rho}_{\boldsymbol{\sigma}}(t)] = \sum_{i,j\in[n]:i\neq j} \frac{d_i d_j}{q^2} = \sum_{i,j=1}^{n} \frac{d_i d_j}{q^2} - \sum_{i=1}^{n} \frac{d_i^2}{q^2}.$$

Hence, the means and covariances of $(\hat{\rho}_{\boldsymbol{\sigma}} - q^{-1}\Delta\mathbb{1})/s_n$ and $\boldsymbol{N}$ match.

We are thus left to prove that $(\hat{\rho}_{\boldsymbol{\sigma}} - q^{-1}\Delta\mathbb{1})/s_n$ is asymptotically normal, with the required uniformity. Thus, given a small $\iota > 0$ we pick $D_1 = D_1(q, \iota) > 0$ and $n_0 = n_0(D_1)$ sufficiently large. Suppose $n > n_0$ and let

$$d_i' = \mathbb{1}\{d_i \leq D_1\}d_i, \qquad d_i'' = d_i - d_i',$$

$$\hat{\rho}_{\boldsymbol{\sigma}}'(s) = \sum_{i=1}^{n} \mathbb{1}\{\boldsymbol{\sigma}_i = s\}d_i', \qquad \hat{\rho}_{\boldsymbol{\sigma}}''(s) = \sum_{i=1}^{n} \mathbb{1}\{\boldsymbol{\sigma}_i = s\}d_i'',$$

$$s_n'^2 = \sum_{i=1}^{n} d_i'^2, \qquad s_n''^2 = \sum_{i=1}^{n} d_i''^2,$$

$$\Delta' = \sum_{i=1}^{n} d_i', \qquad \Delta'' = \sum_{i=1}^{n} d_i''.$$

By construction, we have $\Delta = \Delta' + \Delta''$, $s_n^2 = s_n'^2 + s_n''^2$ as well as $s_n'^2 < D_1^2 n$. Moreover, by (P3) and (P4), both the sequences $(\boldsymbol{d}_n)_n$ and $(\boldsymbol{d}_n^2)_n$ are uniformly integrable, such that for n large enough,

$$\Delta'' < \iota^8 n, \qquad s_n''^2 < \iota^8 n, \tag{A.3}$$

also provided that D_1 is large enough. Hence, the multivariate Berry–Esseen theorem (e.g. [45]) shows that w.h.p. for all U,

$$\mathbb{P}_{\mathfrak{A}}\left[\frac{\hat{\rho}_{\boldsymbol{\sigma}}' - q^{-1}\Delta'\mathbb{1}}{s_n'} \in U\right] - \mathbb{P}\left[\boldsymbol{N} \in U\right] = O(n^{-1/2}). \tag{A.4}$$

Furthermore, combining (A.3) with Chebyshev's inequality, we see that w.h.p.

$$\mathbb{P}_{\mathfrak{A}}\left[\left|\frac{\hat{\rho}_{\boldsymbol{\sigma}}'' - q^{-1}\Delta''\mathbb{1}}{s_n}\right| > \iota^2\right] < \iota^2. \tag{A.5}$$

Thus, combining (A.4) and (A.5), we conclude that w.h.p.

$$\left|\mathbb{P}_{\mathfrak{A}}\left[\frac{\hat{\rho}_{\boldsymbol{\sigma}} - q^{-1}\Delta\mathbb{1}}{s_n} \in U\right] - \mathbb{P}\left[\boldsymbol{N} \in U\right]\right| \leq \iota. \tag{A.6}$$

Finally, the assertion follows from (A.6) by taking the limit $\iota \to 0$ slowly enough as $n \to \infty$. □

Let $\mathfrak{d} = \gcd(\mathrm{supp}(\boldsymbol{d}))$, where $\boldsymbol{d}$ is the weak limit of $(\boldsymbol{d}_n)_n$. Then there exist $g \in \mathbb{N}$, $a_1, \ldots, a_g \in \mathbb{Z}$ and $\delta_1, \ldots, \delta_g$ in the support of $\boldsymbol{d}$ such that the greatest common divisor of the support can be linearly combined as

$$\mathfrak{d} = \sum_{i=1}^{g} a_i \delta_i. \tag{A.7}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

We next count how many variables there are with degree δ_i. For $i \in [g]$, let $\mathfrak{I}_i$ denote the set of all $j \in [n]$ with $d_j = \delta_i$ (the set of all variables that appear in δ_i equations). Set $\mathfrak{I}_0 = [n] \setminus (\mathfrak{I}_1 \cup \ldots \cup \mathfrak{I}_g)$. Then

$$\mathfrak{I}_0 \cup \ldots \cup \mathfrak{I}_g = [n]$$

and $|\mathfrak{I}_1|, \ldots, |\mathfrak{I}_g| = \Theta(n)$ because of assumption (P1). We further count how many entries of value $s \in \mathbb{F}_q^*$ all variables of degree δ_i generate under the assignment $\boldsymbol{\sigma}$, and the contribution from the rest, yielding

$$\boldsymbol{r}_0(s) = \sum_{j \in \mathfrak{I}_0} d_j \mathbb{1}\left\{\boldsymbol{\sigma}_j = s\right\}, \qquad \boldsymbol{r}_i(s) = \sum_{j \in \mathfrak{I}_i} d_j \mathbb{1}\left\{\boldsymbol{\sigma}_j = s\right\}. \qquad (i \in [g], s \in \mathbb{F}_q^*)$$

Then summing the contributions, we get back $\hat{\rho}_{\boldsymbol{\sigma}} = \boldsymbol{r}_0 + \sum_{i=1}^{g} \delta_i \boldsymbol{r}_i$, where $\boldsymbol{r}_i = (\boldsymbol{r}_i(s))_{s \in \mathbb{F}_q^*}$.

Because $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_n$ are mutually independent given $\mathfrak{A}$, so are $\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_g$. Moreover, given $\mathfrak{A}$, for $i \in [g]$, $\boldsymbol{r}_i$ has a multinomial distribution with parameter $|\mathfrak{I}_i|$ and uniform probabilities q^{-1}. In effect, the individual entries $\boldsymbol{r}_i(s)$, $s \in \mathbb{F}_q^*$, will typically differ by only a few standard deviations, that is, their typical difference will be of order $O(\sqrt{|\mathfrak{I}_i|})$. We require a precise quantitative version of this statement.

Furthermore, we say that $\boldsymbol{r}_i$ is *t-tame* if $|\boldsymbol{r}_i(s) - q^{-1}|\mathfrak{I}_i|| \leq t\sqrt{|\mathfrak{I}_i|}$ for all $s \in \mathbb{F}_q^*$. Let $\mathfrak{T}(t)$ be the event that $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_g$ are t-tame.

Lemma A.2. *W.h.p. for every* $r_* \in \mathfrak{L}_0$ *there exists* $r^* \in \mathfrak{L}_0(r_*, \varepsilon)$ *such that*

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} = r^*\right] \geq \frac{1}{2|\mathfrak{L}_0(r_*, \varepsilon)|} \qquad \textit{and} \qquad \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log \varepsilon) \mid \hat{\rho}_{\boldsymbol{\sigma}} = r^*\right] \geq 1 - \varepsilon^4. \tag{A.8}$$

Proof. Since $\boldsymbol{r}_i$ has a multinomial distribution given $\mathfrak{A}$ the Chernoff bound shows that for a large enough $c = c(q)$ w.h.p.

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log \varepsilon)\right] \geq 1 - \exp(-\Omega_\varepsilon(\log^2(\varepsilon))). \tag{A.9}$$

Further, Claim A.1 implies that w.h.p. $\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{L}_0(r_*, \varepsilon)\right] \geq \Omega_\varepsilon(\varepsilon^{q-1}) \geq \varepsilon^q$, provided $\varepsilon < \varepsilon_0 = \varepsilon_0(\omega)$ is small enough. Combining this estimate with (A.9) and Bayes' formula, we conclude that w.h.p. for every $r_* \in \mathfrak{L}_0$,

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log \varepsilon),\ \hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{L}_0(r_*, \varepsilon)\right] \geq 1 - \varepsilon^5. \tag{A.10}$$

To complete the proof, assume that there does not exist $r^* \in \mathfrak{L}_0(r_*, \varepsilon)$ that satisfies (A.8). Then for every $r \in \mathfrak{L}_0(r_*, \varepsilon)$ we either have

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} = r\right] < \frac{1}{2|\mathfrak{L}_0(r_*, \varepsilon)|} \qquad \text{or} \tag{A.11}$$

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log \varepsilon) | \hat{\rho}_{\boldsymbol{\sigma}} = r\right] < 1 - \varepsilon^4. \tag{A.12}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

Let $\mathfrak{X}_0$ be the set of all $r \in \mathfrak{L}_0(r_*, \varepsilon)$ for which (A.11) holds, and let $\mathfrak{X}_1 = \mathfrak{L}_0(r_*, \varepsilon) \setminus \mathfrak{X}_0$. Then (A.11)–(A.12) yield

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log\varepsilon) \mid \hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{L}_0(r_*, \varepsilon)\right] \le \frac{\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{X}_0\right] + \sum_{r \in \mathfrak{X}_1} \mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-\log\varepsilon) | \hat{\rho}_{\boldsymbol{\sigma}} = r\right] \mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} = r\right]}{\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{L}_0(r_*, \varepsilon)\right]}$$

$$\le \frac{\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{X}_0\right] + (1-\varepsilon^4)\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{X}_1\right]}{\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}} \in \mathfrak{L}_0(r_*, \varepsilon)\right]} < 1 - \varepsilon^4,$$

provided that $1 - \varepsilon^4 > \frac{1}{2}$, in contradiction to (A.10). □

Also let $\mathfrak{T}(r, t)$ be the event that $\hat{\rho}_{\boldsymbol{\sigma}} = r$ and that $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_g$ are t-tame. We write $(r_0, \ldots, r_g) \in \mathfrak{T}(r, t)$ if $r_0 + \sum_{i=1}^{g} \delta_i r_i = r$ and $|r_i(s) - q^{-1}|\mathfrak{I}_i|| \le t\sqrt{|\mathfrak{I}_i|}$ for all $s \in \mathbb{F}_q^*$. The following lemma summarises the key step of the proof of Lemma 7.10.

Lemma A.3. *W.h.p. for any $r_* \in \mathfrak{L}_0$, any $1 \le t \le \log n$ and any $r, r' \in \mathfrak{L}_0(r_*, \varepsilon)$ there exists a one-to-one map $\psi : \mathfrak{T}(r, t) \to \mathfrak{T}(r', t + O_\varepsilon(\varepsilon))$ such that for all $(r_0, \ldots, r_g) \in \mathfrak{T}(r, t)$ we have*

$$\log \frac{\mathbb{P}_{\mathfrak{A}}\left[(\boldsymbol{r}_0, \ldots, \boldsymbol{r}_g) = (r_0, \ldots, r_g)\right]}{\mathbb{P}_{\mathfrak{A}}\left[(\boldsymbol{r}_0, \ldots, \boldsymbol{r}_g) = \psi(r_0, \ldots, r_g)\right]} = O_\varepsilon(\varepsilon(\omega + t)). \tag{A.13}$$

Proof. Since $r, r' \in \mathfrak{L}_0(r_*, \varepsilon)$, thanks to assumption (P7), we have $r - r' \in \mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}$. Hence, with $e_1, \ldots, e_{q-1}$ denoting the standard basis of $\mathbb{R}^{\mathbb{F}_q^*}$, there is a unique representation

$$r' - r = \sum_{i=1}^{q-1} \lambda_i \mathfrak{d} e_i \tag{A.14}$$

with $\lambda_1, \ldots, \lambda_{q-1} \in \mathbb{Z}$. Because $r, r' \in \mathfrak{L}_0(r_*, \varepsilon)$ and

$$\lambda := \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{q-1} \end{pmatrix} = \mathfrak{d}^{-1}(r' - r),$$

the coefficients satisfy

$$|\lambda_i| = O_\varepsilon\left(\varepsilon s_n\right) \qquad \text{for all } i = 1, \ldots, q-1. \tag{A.15}$$

Now recall $g \in \mathbb{N}$, $a_1, \ldots, a_g \in \mathbb{Z}$ and $\delta_1, \ldots, \delta_g$ in the support of $\boldsymbol{d}$ with

$$\mathfrak{d} = \sum_{i=1}^{g} a_i \delta_i.$$

Fir $i \in [g]$, we set

$$r_i' = r_i + \frac{a_i}{\mathfrak{d}} \lambda$$

as well as $r_0{}' = r_0$. Further, define $\psi(r_0, \ldots, r_g) = (r_0{}', \ldots, r_g')$. Then clearly

$$r_0 + \sum_{i=1}^{g} \delta_i r_i' = r + \sum_{i=1}^{g} \frac{a_i \delta_i}{\mathfrak{d}} \lambda = r + r' - r = r'. \tag{A.16}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

and due to (A.15), we have $\psi(r_0, \ldots, r_g) \in \mathfrak{T}(r', t + O_\varepsilon(\varepsilon))$. Finally, for $i \in [g]$ set

$$r_i(0) = |\mathfrak{I}_i| - \sum_{s \in \mathbb{F}_q^*} r_i(s), \qquad r_i'(0) = |\mathfrak{I}_i| - \sum_{s \in \mathbb{F}_q^*} r_i'(s).$$

Moreover, Stirling's formula and the mean value theorem show that

$$\begin{aligned}
&\frac{\mathbb{P}_{\mathfrak{A}}\left[(\boldsymbol{r}_0, \ldots, \boldsymbol{r}_g) = (r_0, \ldots, r_g)\right]}{\mathbb{P}_{\mathfrak{A}}\left[(\boldsymbol{r}_0, \ldots, \boldsymbol{r}_g) = \psi(r_0, \ldots, r_g)\right]} \\
&= \frac{\binom{|\mathfrak{I}_1|}{(r_1(0), r_1)} \cdot \ldots \cdot \binom{|\mathfrak{I}_g|}{(r_g(0), r_g)}}{\binom{|\mathfrak{I}_1|}{(r_1'(0), r_1')} \cdot \ldots \cdot \binom{|\mathfrak{I}_g|}{(r_g'(0), r_g')}} = \exp\left[\sum_{i=1}^{g} \sum_{s \in \mathbb{F}_q} O_\varepsilon\left(r_i'(s) \log r_i'(s) - r_i(s) \log r_i(s)\right)\right] \\
&= \exp\left[\sum_{i=1}^{g} O_\varepsilon(|\mathfrak{I}_i|) \sum_{s \in \mathbb{F}_q} \left| \int_{r_i(s)/|\mathfrak{I}_i|}^{r_i'(s)/|\mathfrak{I}_i|} \log z dz \right|\right] \\
&= \exp\left[\sum_{i=1}^{g} O_\varepsilon(|\mathfrak{I}_i|) \sum_{s \in \mathbb{F}_q} \left(\frac{r_i'(s)}{|\mathfrak{I}_i|} - \frac{r_i(s)}{|\mathfrak{I}_i|}\right) \log\left(\frac{1}{q} + O_\varepsilon\left(\frac{(\omega + t)s_n}{|\mathfrak{I}_i|}\right)\right)\right] \\
&= \exp\left[\sum_{i=1}^{g} O_\varepsilon(|\mathfrak{I}_i|) \sum_{s \in \mathbb{F}_q} O_\varepsilon\left(\frac{(\omega + t)s_n}{|\mathfrak{I}_i|}\left(\frac{r_i'(s)}{|\mathfrak{I}_i|} - \frac{r_i(s)}{|\mathfrak{I}_i|}\right)\right)\right].
\end{aligned} \tag{A.17}$$

Since $|\mathfrak{I}_1|, \ldots, |\mathfrak{I}_g| = \Theta_\varepsilon(n)$, (A.17) implies (A.13). Finally, ψ is one to one because each vector has a unique representation with respect to the basis $(e_1, \ldots, e_{q-1})$. □

Roughly speaking, LemmaA.3 shows that any two tame $r, r' \in \mathfrak{L}_0(r_*, \varepsilon)$ close to a conceivable $r_* \in \mathfrak{L}_0$ are about equally likely. However, the map ψ produces solutions that are a little less tame than the ones we start from. The following corollary, which combines Lemmas 7.14 and 7.15, remedies this issue.

Corollary A.4. *W.h.p. for all $r_* \in \mathfrak{L}_0$ and all $r, r' \in \mathfrak{L}_0(r_*, \varepsilon)$ we have*

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -3\log\varepsilon)\right] = (1 + o_\varepsilon(1))\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r', -3\log\varepsilon)\right].$$

Proof. Let r^* be the vector supplied by Lemma A.2. Applying Lemma A.3 to r^* and $r \in \mathfrak{L}_0(r_*, \varepsilon)$, we see that w.h.p.

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -2\log\varepsilon)\right] \geq (1 + O_\varepsilon(\varepsilon\log\varepsilon))\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r^*, -\log\varepsilon)\right] \geq \frac{1}{3|\mathfrak{L}_0(r_*, \varepsilon)|} \qquad \text{for all } r \in \mathfrak{L}_0(r_*, \varepsilon). \tag{A.18}$$

In addition, we claim that w.h.p.

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -4\log\varepsilon) \setminus \mathfrak{T}(r, -3\log\varepsilon)\right] \leq \varepsilon\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r^*, -\log\varepsilon)\right] \qquad \text{for all } r \in \mathfrak{L}_0(r_*, \varepsilon). \tag{A.19}$$

Indeed, applying Lemma 7.15 twice to r and r^* and invoking (7.26), we see that w.h.p.

$$\begin{aligned}
\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r, -2\log\varepsilon)\right] &\geq \exp(O_\varepsilon(\varepsilon\log\varepsilon))\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r^*, -3\log\varepsilon)\right] \\
&\geq \left(1 - O_\varepsilon(\varepsilon\log\varepsilon)\right)\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_\sigma = r^*\right],
\end{aligned} \tag{A.20}$$

https://doi.org/10.1017/S096354832400021X Published online by Cambridge University Press

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r,-4\log\varepsilon)\setminus\mathfrak{T}(r,-3\log\varepsilon)\right]\leq\exp(O_{\varepsilon}(\varepsilon\log\varepsilon))\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(r^{*},-3\log\varepsilon)\setminus\mathfrak{T}(r^{*},-2\log\varepsilon)\right]$$
$$O_{\varepsilon}(\varepsilon^{4})\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}}=r^{*}\right]. \tag{A.21}$$

Combining (A.20) and (A.21) yields (A.19).

Finally, (7.26), (A.18) and (A.19) show that w.h.p.

$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-3\log\varepsilon)\mid\hat{\rho}_{\boldsymbol{\sigma}}=r\right]\geq 1-\sqrt{\varepsilon},$$
$$\mathbb{P}_{\mathfrak{A}}\left[\mathfrak{T}(-3\log\varepsilon)\mid\hat{\rho}_{\boldsymbol{\sigma}}=r'\right]\geq 1-\sqrt{\varepsilon}\qquad\text{for all } r,r'\in\mathfrak{L}_0(r_*,\varepsilon), \tag{A.22}$$

and combining (A.22) with Lemma A.3 completes the proof. □

Proof of Claim 7.17. Claim A.1 shows that for any $r\in\mathfrak{L}_0$ and $\mathbf{N}\sim\mathcal{N}(0,\mathscr{C})$

$$\mathbb{P}_{\mathfrak{A}}\left(\hat{\rho}_{\boldsymbol{\sigma}}\in\mathfrak{L}_0(r,\varepsilon)\right)=\mathbb{P}_{\mathfrak{A}}\left(\left\|\boldsymbol{N}-\frac{r-\Delta\mathbb{1}/q}{s_n}\right\|_{\infty}<\varepsilon\right)+o(1).$$

Moreover, Corollary A.4 implies that given $\hat{\rho}_{\boldsymbol{\sigma}}\in\mathfrak{L}_0(r,\varepsilon)$, $\hat{\rho}_{\boldsymbol{\sigma}}$ is within $o_{\varepsilon}(1)$ of the uniform distribution on $\mathfrak{L}_0(r,\varepsilon)$. Furthermore, Lemma 3.6 applied to the module $\mathfrak{M}=\mathfrak{d}\mathbb{Z}^{\mathbb{F}_q^*}$ with basis $\{\mathfrak{d}e_1,\ldots,\mathfrak{d}e_{q-1}\}$ shows that the number of points in $\mathfrak{L}_0(r,\varepsilon)$ satisfies

$$\frac{|\mathfrak{L}_0(r,\varepsilon)|}{\left|\left\{z\in\mathbb{Z}^{q-1}:\|z-r\|_{\infty}\leq\varepsilon s_n\right\}\right|}\sim\mathfrak{d}^{1-q}.$$

Finally, the eigenvalues of the matrix $\mathscr{C}$ are q^{-2} (once) and q^{-1} ($(q-2)$ times). Hence, $\det\mathscr{C}=q^{-q}$. Therefore, w.h.p. for all $r\in\mathfrak{L}_0$ we have

$$\mathbb{P}_{\mathfrak{A}}\left[\hat{\rho}_{\boldsymbol{\sigma}}=r\right]=(1+o_{\varepsilon}(1))\frac{q^{q/2}\mathfrak{d}^{q-1}}{(2\pi\Delta_2)^{(q-1)/2}}\exp\left[-\frac{(r-q^{-1}\Delta\mathbb{1})^{T}\mathscr{C}^{-1}(r-q^{-1}\Delta\mathbb{1})}{2\Delta_2}\right]. \tag{A.23}$$

□

Cite this article: Coja-Oghlan A, Gao P, Hahn-Klimroth M, Lee J, Müller N, and Rolvien M (2024). The full rank condition for sparse random matrices. *Combinatorics, Probability and Computing* **33**, 643–707. https://doi.org/10.1017/S096354832400021X