The k-XORSAT Threshold Revisited

Amin Coja-Oghlan[a] Mihyun Kang[b] Lena Krieg[c]
Maurice Rolvien[d]

Submitted: Jan 23, 2023; Accepted: Feb 2, 2024; Published: Apr 19, 2024

Abstract

We provide a simplified proof of the random k-XORSAT satisfiability threshold theorem. As an extension we also determine the full rank threshold for sparse random matrices over finite fields with precisely k non-zero entries per row. This complements a result from [Ayre, Coja-Oghlan, Gao, Müller, *Combinatorica*, 2020]. The proof combines physics-inspired message passing arguments with a surgical moment computation.

Mathematics Subject Classifications: 05C88, 05C89

1 Introduction

The random 3-XORSAT problem was one of the first random constraint satisfaction problems whose satisfiability threshold could be pinpointed precisely. A random 3-XORSAT instance consists of a conjunction of XOR-clauses, rather than OR-clauses as in the common k-SAT problem. The goal is to find the maximum number of random XOR-clauses such that the formula remains satisfiable with high probability. The seminal article of Dubois and Mandler [14] that first solved this problem introduced an influential technique, namely the second moment method applied to a pruned problem instance. In their very last sentence Dubois and Mandler asserted that their proof extends to k-XORSAT for any $k \geqslant 3$. However, because of the analytic difficulties associated with estimating the second moment for $k > 3$, this generalisation turned out to be far from straightforward. The first complete proof, covering over 30 pages and involving an (avoidable) bit of computer assistance, was published by Pittel and Sorkin [26] more than a decade

[a] TU Dortmund, Faculty of Computer Science and Faculty of Mathematics, 12 Otto-Hahn-St, Dortmund 44227, Germany (amin.coja-oghlan@tu-dortmund.de).
[b] TU Graz, Institute of Discrete Mathematics, Steyrergasse 30, 8010 Graz, Austria (kang@math.tu-graz.at).
[c] TU Dortmund, Faculty of Computer Science, 12 Otto-Hahn-St, Dortmund 44227, Germany (lena.krieg@tu-dortmund.de).
[d] TU Dortmund, Faculty of Computer Science, 12 Otto-Hahn-St, Dortmund 44227, Germany (maurice.rolvien@tu-dortmund.de).

later. Subsequently a different but still fairly complicated proof that relies on coupling arguments rather than moment calculations was suggested by Ayre, Coja-Oghlan, Gao and Müller [4]. That result covers not only k-XORSAT but also an extension to random matrices over finite fields.

The present contribution develops a relatively short, self-contained derivation of the k-XORSAT threshold as well as said extensions to random matrices via a novel approach that differs significantly from both [4, 26]. The new proof is partly inspired by statistical physics ideas and by recent work on a vaguely related random matrix problem [5, 21]. To elaborate, we first derive a quantitative characterisation of a typical solution to a random k-XORSAT formula by means of what physicists would call a 'quenched' argument. The quenched argument employs Warning Propagation ('WP'), a physics-inspired message passing technique. Then we follow up with a surgical moment computation confined to scenarios that match the precise characteristics predicted by WP. In physics jargon this second bit amounts to an 'annealed' computation. Usually annealed estimates fail to be tight due to large deviations effects. They also tend to be painfully intricate. But because the present specimen carefully homes in on solutions with the correct 'quenched' properties, the calculations are tight as well as elegant.

Let $\boldsymbol{F} = \boldsymbol{F}_k(n, m)$ be a random k-XORSAT instance with n Boolean variables and m random XOR-clauses of length k. To be precise, the clauses are drawn uniformly and independently (with replacement) from the set of all possible $2^k\binom{n}{k}$ XOR-clauses on the variable set $x_1, \ldots, x_n$. The following theorem, first established in [14] for $k = 3$ and in [26] for $k > 3$, provides the k-XORSAT threshold.

Theorem 1. *For $k \geqslant 3$ and $d > 0$ let*

$$\Phi_{d,k}(\alpha) = \exp\left(-d\alpha^{k-1}\right) + d\alpha^{k-1} - \frac{d(k-1)}{k}\alpha^k - \frac{d}{k} \text{ and} \tag{1.1}$$
$$d_k = \sup\left\{d > 0 : \max_{\alpha\in[0,1]} \Phi_{d,k}(\alpha) = 1 - d/k\right\}.$$

For any $\varepsilon > 0$ w.h.p. (i.e., with probability tending to one as $n \to \infty$ with k, ε held fixed) the random k-XORSAT formula $\boldsymbol{F}$ is

(i) satisfiable if $m \leqslant (1-\varepsilon)d_k n/k$,

(ii) unsatisfiable if $m \geqslant (1+\varepsilon)d_k n/k$.

In a nutshell, the k-XORSAT satisfiability threshold equals d_k/k. The threshold admits an explicit combinatorial interpretation, an observation that was vital to the original derivations [14, 26]. To elaborate, we rephrase the k-XORSAT formula $\boldsymbol{F}$ as a linear system over $\mathbb{F}_2$ as follows. Set up a random $m \times n$-matrix $\boldsymbol{A}$ whose i-th row has one-entries in precisely the k columns j such that variable x_j appears in the i-th clause of $\boldsymbol{F}$. Thus, each row of $\boldsymbol{A}$ represents a clause. Further, define $\boldsymbol{y}_i = 1$ iff the number of negations in the i-th clause is even. Then every solution $\sigma \in \mathbb{F}_2^n$ to the linear system $\boldsymbol{A}\sigma = \boldsymbol{y}$ renders a XOR-satisfying assignment of $\boldsymbol{F}$, and vice versa. Because the signs of the literals are

independent of the identities of the underlying variables, the vector $\boldsymbol{y}$ is independent of $\boldsymbol{A}$. Therefore, the random XOR-formula $\boldsymbol{F}$ is satisfiable w.h.p. iff $\boldsymbol{A}$ has full row rank m w.h.p.

Now consider the following process that prunes $\boldsymbol{A}$ down to a minor $\boldsymbol{A}^{(2)}$:

> while there exists a column with at most a single non-zero entry, remove that column along with the row where its non-zero entry appears (if there is one).

This is just the random hypergraph 2-core peeling process phrased in terms of the matrix $\boldsymbol{A}$. Therefore, it is possible (albeit non-trivial) to track the pruning process so as to determine the likely size of $\boldsymbol{A}^{(2)}$ [23]. This analysis evinces that $d_k n/k$ marks the threshold beyond which $\boldsymbol{A}^{(2)}$ has more rows than columns w.h.p. In effect, for $m \geqslant (1+\varepsilon)d_k n/k$ the minor $\boldsymbol{A}^{(2)}$ cannot have full row rank anymore, nor can the original matrix $\boldsymbol{A}$. Consequently, $\boldsymbol{F}$ is unsatisfiable for $m \geqslant (1+\varepsilon)d_k n/k$.

Although for $m \leqslant (1-\varepsilon)d_k n/k$ the minor $\boldsymbol{A}^{(2)}$ has fewer rows than columns, it is by no means a foregone conclusion that $\boldsymbol{A}^{(2)}$ also has full row rank w.h.p. Indeed, in [14, 26] the main technical difficulty lies in demonstrating this fact via the second moment method. The necessary calculations turn out to be delicate because they operate with the outcome $\boldsymbol{A}^{(2)}$ of the pruning process, a matrix whose rows are stochastically dependent. The second moment therefore involves subtle large deviations trade-offs. Luckily, the proof strategy that we propose here requires neither an explicit analysis of the pruning process, nor complicated large deviations arguments.

Theorem 1 admits a natural generalisation to matrices over finite fields beyond $\mathbb{F}_2$. Let $q \geqslant 2$ be a prime power and let $\mathfrak{A} = (\mathfrak{A}_{ij})_{i,j\geqslant 1}$ be an infinite matrix with entries $\mathfrak{A}_{ij} \in \mathbb{F}_q \setminus \{0\}$. Further, given integers $m, n > 0$ and $k \geqslant 3$ let $(\boldsymbol{e}_i)_{i\geqslant 1}$ be a family of independent uniformly random subsets of $[n]$ of size $|e_i| = k$ and define a random $m \times n$-matrix $\boldsymbol{A} = \boldsymbol{A}(k, m, n, q, \mathfrak{A})$ over $\mathbb{F}_q$ by letting

$$\boldsymbol{A}_{ij} = \mathfrak{A}_{ij}\mathbf{1}\{j \in \boldsymbol{e}_i\} \qquad (i \in [m],\, j \in [n]). \tag{1.2}$$

Thus, $\boldsymbol{A}$ has precisely k non-zero entries per row. The positions of the non-zero entries are determined by the $\boldsymbol{e}_i$, while the entries themselves are copied from $\mathfrak{A}$. Naturally, in the case $q = 2$ we simply obtain the matrix induced by the k-XORSAT formula $\boldsymbol{F}$. Therefore, the following theorem encompasses Theorem 1 as a special case.

Theorem 2. *For any $k \geqslant 3$, any prime power $q \geqslant 2$ and any infinite matrix $\mathfrak{A}$ composed of non-zero elements of $\mathbb{F}_q$ the following is true. Let d_k be the threshold from* (1.1)*. Then for any $\varepsilon > 0$,*

(i) if $m \leqslant (1-\varepsilon)d_k n/k$, then w.h.p. $\boldsymbol{A}$ has full row rank.

(ii) if $m \geqslant (1+\varepsilon)d_k n/k$, then w.h.p. $\boldsymbol{A}$ has row rank $m - \Omega(n)$.

Theorem 2 complements [4, Theorem 1.1], where only random matrices with identically distributed rows were considered. By contrast, in Theorem 2 the matrix $\mathfrak{A}$ may proscribe different non-zero entries for each row. That said, in hindsight the theorem shows that the full rank threshold is independent of both q and $\mathfrak{A}$. We proceed to outline the strategy upon which the proof of Theorem 2 is based.

2 Proof strategy

The main difficulty lies in proving the positive statement Theorem 2(i). Suppose we could argue that for $m < (1-\varepsilon)d_k n/k$ w.h.p. a random vector $\boldsymbol{\sigma} \in \ker \boldsymbol{A}$ is approximately 'balanced' in the sense that every value $s \in \mathbb{F}_q$ appears in $\boldsymbol{\sigma}$ about n/q times. Since a straightforward moment calculation shows that the expected number of balanced $\sigma \in \ker \boldsymbol{A}$ equals $(1+o(1))q^{n-m}$, we could then conclude that $|\ker \boldsymbol{A}| = (1+o(1))q^{n-m}$ w.h.p., and thus that $\boldsymbol{A}$ has full row rank w.h.p.

However, we will not be able to prove directly that a random $\boldsymbol{\sigma} \in \ker \boldsymbol{A}$ is balanced w.h.p. Instead we will work with a matrix $\boldsymbol{A}^\dagger$ obtained from $\boldsymbol{A}$ by a small but consequential perturbation called 'pinning'. The matrix $\boldsymbol{A}^\dagger$ contains $\boldsymbol{A}$ as its top $m \times n$-minor, but $\boldsymbol{A}^\dagger$ has $O(\log n)$ additional rows. Pinning guarantees that $\boldsymbol{A}^\dagger$ has only relatively few 'short linear relations', a property that will pave the way for us to bring the Warning Propagation ('WP') message passing scheme to bear. Ultimately we will argue that random $\boldsymbol{\sigma}^\dagger \in \ker \boldsymbol{A}^\dagger$ are balanced w.h.p. As outlined in the previous paragraph, this will imply that $\boldsymbol{A}^\dagger$ has full row rank w.h.p., whence the same is true of $\boldsymbol{A}$.

The purpose of WP is to show that the vectors in the kernel of $\boldsymbol{A}^\dagger$ have a peculiar structure. Specifically, there are certain coordinates $j \in [n]$ that are 'frozen' in $\boldsymbol{A}^\dagger$, meaning that $\sigma_j = 0$ for all $\sigma \in \ker \boldsymbol{A}^\dagger$. By contrast, the values assigned to the unfrozen coordinates are essentially balanced. Hence, if $\boldsymbol{\alpha} n$ variables are frozen, then in a random $\boldsymbol{\sigma}^\dagger \in \ker \boldsymbol{A}^\dagger$ each non-zero value $s \in \mathbb{F}_q \setminus \{0\}$ appears about $(1-\boldsymbol{\alpha})n/q$ times. Ultimately we will argue that $\boldsymbol{\alpha} = o(1)$ w.h.p., which implies that $\boldsymbol{\sigma}^\dagger$ is balanced w.h.p.

But the proof that $\boldsymbol{\alpha} = o(1)$ w.h.p. requires a few more steps. First, from WP we learn that the probability that $j \in [n]$ is frozen depends on the number of non-zero entries in the j-th column of $\boldsymbol{A}^\dagger$. In fact, WP renders detailed 'local' information about the distribution of the frozen coordinates. In the quenched part of the analysis, we will extract this information carefully to obtain a quantitative picture of the structure of the kernel vectors in terms of the as yet unknown value of $\boldsymbol{\alpha}$. Moreover, we will see that the messages exchanged by WP satisfy a certain fixed point property.

Subsequently, we will develop an 'annealed' (moment computation) argument that allows us to bound the number of WP fixed points associated with any conceivable value of $\boldsymbol{\alpha}$. Moreover, we will compute the expected number of vectors $\sigma \in \ker \boldsymbol{A}^\dagger$ that are consistent with a given WP fixed point. This calculation will reveal that w.h.p. for $m < (1-\varepsilon)d_k n/k$ no WP fixed point with $\Omega(n)$ frozen coordinates gives rise to $q^{n-m-o(n)}$ kernel vectors, the number of vectors that we know the kernel of $\boldsymbol{A}^\dagger$ must contain because its rank and its nullity sum to n. Hence, we deduce that $\boldsymbol{\alpha} = o(1)$ w.h.p., as desired.

In the rest of this section we discuss in more detail the proof of Theorem 2(i). We begin with the pinning operation in Section 2.1, then discuss WP and the quenched and annealed analyses. The proof of the second assertion Theorem 2 (ii) is but an afterthought. Indeed, as mentioned in Section 1 this second assertion could be derived from known results about the size of the minor $\boldsymbol{A}^{(2)}$. Nonetheless, Section 5 contains a self-contained proof based on the interpolation method that avoids the analysis of the pruning process.

2.1 Pinning

Adding a few rows to a matrix, the randomised pinning operation mostly removes 'short linear relations'. The operation, devised in this form in [7], actually works on any matrix, not just on the random matrix $\boldsymbol{A}$. Hence, let A be any $\mathbb{F}_q$-matrix of size $M \times N$. For an integer $t \geqslant 0$ obtain $A[t]$ from A by adding t new rows that each contain a single non-zero entry, namely a one in a random position chosen independently and uniformly from the N columns.

The purpose of this operation is to diminish the number of short relations. To be precise, following [7] we call a set $J \subseteq [N]$ of columns a *relation of* A if there exists a vector $y \in \mathbb{F}_q^M$ such that

$$\operatorname{supp}(y^\top A) = \big\{j \in [N] : (y^\top A)_j \neq 0\big\}$$

is a non-empty subset of J. In other words, the non-zero entries of the linear combination $y^\top A \neq 0$ of the rows of A are confined to J. Further, call $j \in [N]$ *frozen in* A if the singleton $\{j\}$ is a relation of A. Thus, j is frozen iff $\sigma_j = 0$ for every $\sigma \in \ker A$. Let $\mathcal{F}(A)$ be the set of all frozen $j \in [N]$.

In addition, call $J \neq \emptyset$ a *proper relation* of A if $J \setminus \mathcal{F}(A)$ is a relation of A. Finally, we say that A is (δ, ℓ)*-free* if A possesses fewer than $\delta\binom{N}{h}$ proper relations I of size $|I| = h$ for any $2 \leqslant h \leqslant \ell$. This definition is meant to express that A contains few relations of size ℓ that are not 'just' composed of frozen $j \in [N]$.[1]

Lemma 3 ([7, Proposition 2.4])**.** *For any $\delta > 0, \ell > 0$ there exists $T_0 = O(\ell^3/\delta^4) > 0$ such that for any $T \geqslant T_0$ and any matrix A for a random $\boldsymbol{t} \in [T]$ we have $\mathbb{P}\left[A[\boldsymbol{t}] \text{ is } (\delta, \ell)\text{-free}\right] > 1 - \delta$.*

Setting $T = \lceil \log n \rceil$, we let $\boldsymbol{A}^\dagger = \boldsymbol{A}[\boldsymbol{t}]$ for a random $\boldsymbol{t} \in [T]$. Since T_0 in Lemma 3 is independent of the size of A and scales polynomially in ℓ, δ, we obtain the following.

Corollary 4. *Let $\omega = \lceil \log \log n \rceil$. W.h.p. $\boldsymbol{A}^\dagger$ is (ω^{-1}, ω)-free.*

Thanks to the scarcity of short proper relations provided by Corollary 4 we will be able to characterise the frozen set $\mathcal{F}(\boldsymbol{A}^\dagger)$ in terms of the WP message passing scheme, which is the next item on our agenda.

2.2 Warning Propagation

Since we will need to work not just with $\boldsymbol{A}^\dagger$ but also with a few other matrices derived from it, we introduce WP for a general matrix A of size $M \times N$. The matrix A naturally induces a bipartite graph $G(A)$ called the *Tanner graph*. Its vertex set comprises a set $V_N = \{v_1, \ldots, v_N\}$ of *variable nodes* and another set $F_M = \{a_1, \ldots, a_M\}$ of *check nodes*. The former represent the columns of A and the latter the rows. An edge $a_i v_j$ is present

[1] Lemma 3 is the only statement beyond textbook knowledge that we apply without a proof in order to derive Theorem 2. The proof, which relies on a potential function argument and a bit of linear algebra, is neither long nor difficult.

in $G(A)$ iff $A_{ij} \neq 0$. For a vertex $u \in V_N \cup F_M$ let $\partial u = \partial_A u$ denote its set of neighbours. Moreover, for a set $S \subseteq V_N \cup F_M$ let $A \setminus S$ be the minor of A obtained by deleting all rows i such that $a_i \in S$ as well as all columns j such that $v_j \in S$. In defining the WP scheme we follow [5].

The thrust of WP is to characterise the set $\mathcal{F}(A)$ of frozen variables in terms of just the immediate local interactions between variables and their adjacent checks. To this end we associate messages with the edges of the Tanner graph. Specifically, each edge $v_j a_i$ of $G(A)$ comes with one message directed from v_j to a_i and a message in the reverse direction. The messages take the symbolic values $\{\mathtt{u}, \mathtt{f}\}$ to represent 'unfrozen' and 'frozen'. Let

$$\mathfrak{M}(A) = \{\mathfrak{m} = (\mathfrak{m}_{v\to a}, \mathfrak{m}_{a\to v})_{v\in V_N, v\in\partial_A a} : \mathfrak{m}_{v\to a}, \mathfrak{m}_{a\to v} \in \{\mathtt{u}, \mathtt{f}\}\}$$

be the set of all possible collections of messages. Further, define the *standard messages* of A by letting

$$\begin{aligned}
\mathfrak{m}_{v_j\to a_i}(A) &= \begin{cases} \mathtt{f} & \text{if } j \in \mathcal{F}(A \setminus \{a_i\}) \\ \mathtt{u} & \text{otherwise} \end{cases} \\
\mathfrak{m}_{a_i\to v_j}(A) &= \begin{cases} \mathtt{f} & \text{if } v_j \in \mathcal{F}(A \setminus (\partial v_j \setminus \{a_i\})) \\ \mathtt{u} & \text{otherwise} \end{cases} \qquad (i \in [M],\, j \in [N]).
\end{aligned} \tag{2.1}$$

Thus, $\mathfrak{m}_{v_j\to a_i}(A) = \mathtt{f}$ indicates that variable v_j is frozen in the matrix obtained from A by deleting row a_i. Similarly, $\mathfrak{m}_{a_i\to v_j}(A) = \mathtt{f}$ if variable v_j is frozen in the matrix obtained from A by deleting all rows $a_h \in \partial_A v_j$ except for a_i.

If indeed freezing were a perfectly local phenomenon transmitted along the edges of the Tanner graph, then the messages (2.1) should remain invariant under the *Warning Propagation update* $\mathrm{WP}_A : \mathfrak{M}(A) \to \mathfrak{M}(A)$, $\mathfrak{m} = (\mathfrak{m}_{\cdot\to\cdot}) \mapsto \mathrm{WP}(\mathfrak{m}) = (\hat{\mathfrak{m}}_{\cdot\to\cdot})$, which is defined by

$$\begin{aligned}
\hat{\mathfrak{m}}_{v_j\to a_i} &= \begin{cases} \mathtt{f} & \text{if } \exists a_h \in \partial v_j \setminus \{a_i\} : \mathfrak{m}_{a_h\to v_j} = \mathtt{f}, \\ \mathtt{u} & \text{otherwise,} \end{cases} \\
\hat{\mathfrak{m}}_{a_i\to v_j} &= \begin{cases} \mathtt{f} & \text{if } \forall x_h \in \partial a_i \setminus \{v_j\} : \mathfrak{m}_{x_h\to a_i} = \mathtt{f}, \\ \mathtt{u} & \text{otherwise.} \end{cases}
\end{aligned} \tag{2.2}$$

Indeed, the first update rule $\hat{\mathfrak{m}}_{v_j\to a_i}$ expresses that we expect v_j to be frozen in $A \setminus \{a_i\}$ iff some other check a_h 'freezes' v_j. Similarly, one might expect that a_i causes v_j to freeze iff all the other variables x_h adjacent to a_i freeze, thereby leaving no other option to satisfy a_i but to always set x_j to zero as well.

Finally, in order to extract the set of frozen variables from the WP messages, we define

$\{\mathtt{u},\mathtt{s},\mathtt{f}\}$-valued labels to go with the variable and check nodes: for $\mathfrak{m} \in \mathfrak{M}(A)$ let

$$\mathfrak{m}_{v_j} = \begin{cases} \mathtt{f} & \text{if } \mathfrak{m}_{a\to v_j} = \mathtt{f} \text{ for at least two } a \in \partial v_j, \\ \mathtt{s} & \text{if } \mathfrak{m}_{a\to v_j} = \mathtt{f} \text{ for precisely one } a \in \partial v_j, \\ \mathtt{u} & \text{otherwise,} \end{cases} \tag{2.3}$$

$$\mathfrak{m}_{a_i} = \begin{cases} \mathtt{f} & \text{if } \mathfrak{m}_{v\to a_i} = \mathtt{f} \text{ for all } v \in \partial a_i, \\ \mathtt{s} & \text{if } \mathfrak{m}_{v\to a_i} = \mathtt{f} \text{ for all but one } v \in \partial a_i, \\ \mathtt{u} & \text{otherwise.} \end{cases} \tag{2.4}$$

Here the new label $\mathfrak{m}_{v_j}(A) = \mathtt{s}$ ('slush') indicates that v_j is 'barely' frozen as there is only one incoming $\mathtt{f}$-message. At first glance the $\mathtt{s}$-label may seem superfluous as it could just be subsumed by $\mathtt{f}$ in the case of $\mathfrak{m}_{v_j}$, and by $\mathtt{u}$ in $\mathfrak{m}_{a_i}$. However, under (2.2) the $\mathtt{s}$-labeled vertices 'return' different messages than those labeled $\mathtt{f}$ or $\mathtt{u}$. For instance, if $\mathfrak{m}_{v_j} = \mathtt{s}$, then $\hat{\mathfrak{m}}_{v_j\to a_i} = \mathtt{u}$ if $\mathfrak{m}_{a_i\to v_j} = \mathtt{f}$, whereas in the case $\mathfrak{m}_{v_j} = \mathtt{f}$ we have $\hat{\mathfrak{m}}_{v_j\to a_i} = \mathtt{f}$ for all $a_i \in \partial v_j$. Let $\mathfrak{m}_{v_j}(A)$, $\mathfrak{m}_{a_i}(A)$ denote the labels extracted via (2.3)–(2.4) from the standard messages $\mathfrak{m}_{\cdot\to\cdot}(A)$ from (2.1).

It is easily verified that the WP messages (2.1) coincide with the updated messages if $G(A)$ is acyclic, i.e.,

$$\mathfrak{m}_{v_j\to a_i}(A) = \hat{\mathfrak{m}}_{v_j\to a_i}(A), \qquad \mathfrak{m}_{a_i\to v_j}(A) = \hat{\mathfrak{m}}_{a_i\to v_j}(A), \tag{2.5}$$

for all i, j such that $a_i \in \partial_A v_j$. But it is equally easy to come up with cyclic Tanner graphs where (2.5) is violated.

Nonetheless, the following proposition shows that (2.5) is satisfied on the random matrix $\boldsymbol{A}^\dagger$ for all but $o(n)$ adjacent pairs a_i, v_j w.h.p. The proposition also shows that the labels extracted via (2.3) correctly identify the set $\mathcal{F}(\boldsymbol{A}^\dagger)$, up to at most $o(n)$ exceptions. Furthermore, in most kernel vectors the values of the unfrozen variables are about 'balanced'. Let $\boldsymbol{\alpha} = |\mathcal{F}(\boldsymbol{A}^\dagger)|/n$ be the fraction of frozen variables of $\boldsymbol{A}^\dagger$ and let $\boldsymbol{\sigma}^\dagger$ be a uniformly random element of $\ker \boldsymbol{A}^\dagger$. Moreover, let $d_{\boldsymbol{A}^\dagger}(v_j)$ denote the degree of a variable node v_j in $G(\boldsymbol{A}^\dagger)$.

Proposition 5. *Let $d > 0, k \geqslant 3$. W.h.p. we have*

$$\sum_{i=1}^{m} \sum_{v_j \in \partial_{\boldsymbol{A}^\dagger} a_i} \mathbf{1}\left\{\mathfrak{m}_{v_j\to a_i}(\boldsymbol{A}^\dagger) \neq \hat{\mathfrak{m}}_{v_j\to a_i}(\boldsymbol{A}^\dagger)\right\} + \mathbf{1}\left\{\mathfrak{m}_{a_i\to v_j}(\boldsymbol{A}^\dagger) \neq \hat{\mathfrak{m}}_{a_i\to v_j}(\boldsymbol{A}^\dagger)\right\} = o(n), \tag{2.6}$$

$$\left|\left\{j \in [n] : \mathfrak{m}_{v_j}(\boldsymbol{A}^\dagger) \neq \mathtt{u}\right\} \triangle \mathcal{F}(\boldsymbol{A}^\dagger)\right| = o(n), \tag{2.7}$$

$$\sum_{s\in\mathbb{F}_q} \sum_{\ell\geqslant 0} \left| \sum_{j=1}^{n} \mathbf{1}\{d_{\boldsymbol{A}^\dagger}(v_j) = \ell,\, \mathfrak{m}_{v_j}(\boldsymbol{A}^\dagger) = \mathtt{u}\} \left(\mathbf{1}\{\boldsymbol{\sigma}_j^\dagger = s\} - 1/q\right) \right| = o(n). \tag{2.8}$$

Observe that (2.8) posits that the unfrozen variables are not just 'balanced' overall (in the sense that every value $s \in \mathbb{F}_q$ occurs with frequency about $1/q$), but that balance

even holds once we break things down to unfrozen variables of some specific degree $\ell \geqslant 0$. The proof Proposition 5, which we carry out in Section 3, rests on the scarcity of short linear relations provided by Corollary 4.

2.3 Quenched analysis

Recall that our goal is to show that $\boldsymbol{\sigma}^\dagger$ is approximately balanced w.h.p. Proposition 5 reduces this task to showing that $\boldsymbol{\alpha} = o(1)$ w.h.p. To this end we are going to extract some more detailed information about the WP messages that the variable and check nodes exchange. Specifically, we are going to estimate the number of variables/checks with specific labels according to (2.3)–(2.4). In fact, we even need to know the number of variables/checks with specific labels and with specific numbers of incoming/outgoing message pairs. Hence, our next goal is to derive such formulas in terms of the (as yet) unknown random variable $\boldsymbol{\alpha}$.

To account for the numbers of message pairs received/sent by the various nodes let $\mathcal{L}$ be the set of all vectors $\ell = (\ell_{\mathtt{uu}}, \ell_{\mathtt{uf}}, \ell_{\mathtt{fu}}, \ell_{\mathtt{ff}}) \in \mathbb{Z}^4_{\geqslant 0}$. For $\ell \in \mathcal{L}$, a label $z \in \{\mathtt{f}, \mathtt{s}, \mathtt{u}\}$, a matrix A and a collection of messages $\mathfrak{m}_{\cdot\to\cdot} \in \mathfrak{M}(A)$ let

$$\begin{aligned}&\Delta_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}) = \\ &\{v \in V(A) : (\mathfrak{m}_v = z) \wedge (\forall s,t \in \{\mathtt{u},\mathtt{f}\} : |\{a \in \partial v : \mathfrak{m}_{a\to v} = s,\ \mathfrak{m}_{v\to a} = t\}| = \ell_{st})\}\,,\end{aligned} \tag{2.9}$$

$$\begin{aligned}&\Gamma_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}) = \\ &\{a \in F(A) : (\mathfrak{m}_a = z) \wedge (\forall s,t \in \{\mathtt{u},\mathtt{f}\} : |\{v \in \partial a : \mathfrak{m}_{v\to a} = s,\ \mathfrak{m}_{a\to v} = s\}| = \ell_{st})\}\,.\end{aligned} \tag{2.10}$$

Thus, $\Delta_{z,\ell}$ comprises variable nodes labelled z by (2.3) that receive/send out numbers of WP messages as detailed by ℓ. To be precise, the first label s of ℓ_{st} encodes the incoming message, while the second index t specifies the outgoing messages. Similarly, $\Gamma_{z,\ell}$ counts checks with a given label and given message statistics.

We are going to calculate $|\Delta_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger))|, |\Gamma_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger))|$ in terms of the fraction $\boldsymbol{\alpha}$ of frozen variables. As a first step, the following sets comprise the conceivable vectors ℓ to go with the various types of variable/check nodes, in line with (2.3)–(2.4):

$$\mathcal{D}(\mathtt{u}) = \{\ell \in \mathcal{L} : \ell_{\mathtt{fu}} = \ell_{\mathtt{uf}} = \ell_{\mathtt{ff}} = 0\}\,, \tag{2.11}$$
$$\mathcal{G}(\mathtt{u}) = \{\ell \in \mathcal{L} : \ell_{\mathtt{uf}} = \ell_{\mathtt{ff}} = 0, \ell_{\mathtt{uu}} \geqslant 2, \ell_{\mathtt{fu}} = k - \ell_{\mathtt{uu}}\}$$
$$\mathcal{D}(\mathtt{s}) = \{\ell \in \mathcal{L} : \ell_{\mathtt{fu}} = 1, \ell_{\mathtt{ff}} = \ell_{\mathtt{uu}} = 0\}\,, \tag{2.12}$$
$$\mathcal{G}(\mathtt{s}) = \{\ell \in \mathcal{L} : \ell_{\mathtt{uu}} = \ell_{\mathtt{ff}} = 0, \ell_{\mathtt{uf}} = 1, \ell_{\mathtt{fu}} = k - 1\}\,,$$
$$\mathcal{D}(\mathtt{f}) = \{\ell \in \mathcal{L} : \ell_{\mathtt{uu}} = \ell_{\mathtt{fu}} = 0, \ell_{\mathtt{ff}} \geqslant 2\}\,, \tag{2.13}$$
$$\mathcal{G}(\mathtt{f}) = \{\ell \in \mathcal{L} : \ell_{\mathtt{uu}} = \ell_{\mathtt{uf}} = \ell_{\mathtt{fu}} = 0, \ell_{\mathtt{ff}} = k\}\,.$$

Further, we hypothesise that the incoming messages at a check node a_i are essentially independent. This seems plausible as the Tanner graph $G(\boldsymbol{A}^\dagger)$ is a sparse random graph with bounded average degree d on the variable side and constant degree k on the check side. Therefore, typically the neighbouring variable nodes $\partial_{\boldsymbol{A}^\dagger} a_i$ should end up being far

from each other in $G(\boldsymbol{A}^\dagger \setminus \{a_i\})$, and far apart vertices might conceivably decorrelate. By a similar token, we expect that the messages received by a typical variable node v_j ought to be nearly independent. If so, and if we presume that variable-to-check messages take the value $\mathtt{f}$ with some probability $0 \leqslant \alpha \leqslant 1$, then check-to-variable messages should take the value $\mathtt{f}$ with probability α^{k-1}; for according to (2.2) a check-to-variable message should be $\mathtt{f}$ iff all of the check's other $k-1$ incoming messages are $\mathtt{f}$. In light of (2.11)–(2.13) we can thus predict the frequencies for the variable/check nodes of the various types. For instance, if $\boldsymbol{\alpha} = \alpha$, then we expect to see about $\bar{\delta}(\alpha, \mathtt{u}) = \exp(-d\alpha^{k-1})n$ variables v_j with $\mathfrak{m}_{v_j}(\boldsymbol{A}^\dagger) = \mathtt{f}$. This is because by (2.11) such a variable v_j must not receive any $\mathtt{f}$-messages, while the mean number of such incoming messages should be $d\alpha^{k-1}$. Similarly, we arrive at predictions for the frequencies of the other node types:

$$\begin{aligned}
\bar{\delta}(\alpha, \mathtt{u}) &= \exp(-d\alpha^{k-1}), && (2.14)\\
\bar{\delta}(\alpha, \mathtt{s}) &= d\alpha^{k-1}\exp(-d\alpha^{k-1}),\\
\bar{\delta}(\alpha, \mathtt{f}) &= 1 - \exp(-d\alpha^{k-1})(1 + d\alpha^{k-1}),\\
\bar{\gamma}(\alpha, \mathtt{u}) &= 1 - k(1-\alpha)\alpha^{k-1} - \alpha^k, && (2.15)\\
\bar{\gamma}(\alpha, \mathtt{s}) &= k(1-\alpha)\alpha^{k-1},\\
\bar{\gamma}(\alpha, \mathtt{f}) &= \alpha^k.
\end{aligned}$$

Finally, extending the reasoning outlined in the previous paragraph, we can derive predictions as to the frequencies of nodes with various labels and given statistics $\ell \in \mathcal{L}$ of incoming/outgoing messages. With $\mathrm{Po}_{\geqslant 2}(\lambda)$ and $\mathrm{Bin}_{\geqslant 2}(N, p)$ denoting the conditional Poisson/Binomial distributions given an outcome of at least two, we obtain the following expressions:

$$\bar{\Delta}_{\mathtt{u},\ell}(\alpha) = \bar{\delta}(\alpha, \mathtt{u})\mathbf{1}\{\ell \in \mathcal{D}(\mathtt{u})\}\mathbb{P}[\mathrm{Po}(d(1-\alpha^{k-1})) = \ell_{\mathtt{uu}}], \tag{2.16}$$

$$\bar{\Delta}_{\mathtt{s},\ell}(\alpha) = \bar{\delta}(\alpha, \mathtt{s})\mathbf{1}\{\ell \in \mathcal{D}(\mathtt{s})\}\mathbb{P}[\mathrm{Po}(d(1-\alpha^{k-1})) = \ell_{\mathtt{uf}}], \tag{2.17}$$

$$\bar{\Delta}_{\mathtt{f},\ell}(\alpha) = \bar{\delta}(\alpha, \mathtt{f})\mathbf{1}\{\ell \in \mathcal{D}(\mathtt{f})\}\mathbb{P}[\mathrm{Po}_{\geqslant 2}(d\alpha^{k-1}) = \ell_{\mathtt{ff}}]\mathbb{P}[\mathrm{Po}(d(1-\alpha^{k-1})) = \ell_{\mathtt{uf}})], \tag{2.18}$$

$$\bar{\Gamma}_{\mathtt{u},\ell}(\alpha) = \bar{\gamma}(\alpha, \mathtt{u})\mathbf{1}\{\ell \in \mathcal{G}(\mathtt{u})\}\mathbb{P}\left[\mathrm{Bin}_{\geqslant 2}(k, 1-\alpha) = \ell_{\mathtt{uu}}\right], \tag{2.19}$$

$$\bar{\Gamma}_{\mathtt{s},\ell}(\alpha) = \bar{\gamma}(\alpha, \mathtt{s})\mathbf{1}\{\ell \in \mathcal{G}(\mathtt{s})\}, \tag{2.20}$$

$$\bar{\Gamma}_{\mathtt{f},\ell}(\alpha) = \bar{\gamma}(\alpha, \mathtt{f})\mathbf{1}\{\ell \in \mathcal{G}(\mathtt{f})\}. \tag{2.21}$$

The following proposition shows that the aforementioned predictions are accurate w.h.p.

Proposition 6. *Let $d > 0, k \geqslant 3$. Then*

$$\sum_{z \in \{\mathtt{f},\mathtt{s},\mathtt{u}\}} \sum_{\ell \in \mathcal{L}} \mathbb{E}\left|\left|\Delta_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger))\right| - n\bar{\Delta}_{z,\ell}(\boldsymbol{\alpha})\right| + \mathbb{E}\left|\left|\Gamma_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger))\right| - m\bar{\Gamma}_{z,\ell}(\boldsymbol{\alpha})\right| = o(n).$$

Thus, $|\Delta_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger))|$, $|\Gamma_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger))|$ approximately equal $\bar{\Delta}_{z,\ell}(\boldsymbol{\alpha})n$, $\bar{\Gamma}_{z,\ell}(\boldsymbol{\alpha})m$ evaluated at the actual fraction $\boldsymbol{\alpha}$ of frozen variables of $\boldsymbol{A}^\dagger$, which is a random variable. The proof of Proposition 6, which can be found in Section 3.3, is based on coupling arguments. In particular, the proof does not reveal the likely value of $\boldsymbol{\alpha}$.

2.4 Annealed arguments

In light of (2.8) from Proposition 5 our main task is to show that $\boldsymbol{\alpha} = o(1)$ w.h.p. if $d < (1-\varepsilon)d_k/k$. To this end we are going to combine Proposition 6 with a first moment argument that shows that for $d < (1-\varepsilon)d_k/k$ only the scenario $\boldsymbol{\alpha} = o(1)$ w.h.p. can account for the $q^{n-m+o(n)}$ vectors that the kernel of the $(m+o(n)) \times n$-matrix $\boldsymbol{A}^\dagger$ must inevitably contain. In other words, we are going to show that WP fixed points with $\Omega(n)$ frozen variables come with too small a number of kernel vectors.

In this respect the present argument differs significantly from prior proofs of Theorem 1 [14, 26]. Instead of first investigating the likely shape of vectors in the kernel (specifically, that they 'come from' WP fixed points with certain statistics), these analyses directly estimate the expected number of vectors in the kernel with a given Hamming weight; of course, this kind of argument is workable only in the case $q = 2$. The drawback of a blunt moment computation is that even extremely rare events make a contribution. Such large deviations effects tend to lead to intricate and technically demanding analytical optimisation problems.

The present 'annealed' argument (viz. moment computation) consists of two layers. First we estimate the expected number of WP fixed points with the 'correct' statistics as provided by (2.16)–(2.21). To be precise, reminding ourselves of the update rules (2.2), we call $\mathfrak{m}_{\cdot\to\cdot} \in \mathfrak{M}(\boldsymbol{A}^\dagger)$ an α-*WP fixed point* if

$$\sum_{i=1}^{m} \sum_{v_j \in \partial_{\boldsymbol{A}^\dagger} a_i} \mathbf{1}\left\{\mathfrak{m}_{v_j \to a_i} \neq \hat{\mathfrak{m}}_{v_j \to a_i}\right\} + \mathbf{1}\left\{\mathfrak{m}_{a_i \to v_j} \neq \hat{\mathfrak{m}}_{a_i \to v_j}\right\} = o(n) \qquad \text{and} \tag{2.22}$$

$$\sum_{z \in \{\mathtt{f},\mathtt{s},\mathtt{u}\}} \sum_{\ell \in \mathcal{L}} \left| |\Delta_{z,\ell}(\mathfrak{m})| - n\bar{\Delta}_{z,\ell}(\alpha) \right| + \left| |\Gamma_{z,\ell}(\mathfrak{m})| - m\bar{\Gamma}_{z,\ell}(\alpha) \right| = o(n). \tag{2.23}$$

Thus, we ask that most messages be invariant under the update (2.2), and that the counts $|\Delta_{z,\ell}(\mathfrak{m})|, |\Gamma_{z,\ell}(\mathfrak{m})|$ be in line with Proposition 6. Performing relatively simple manipulations of the formulas (2.16)–(2.21), we will ultimately see that the expected number of α-WP fixed points is sub-exponential for any $0 \leqslant \alpha \leqslant 1$.

As a next step, we will estimate the number of kernel vectors σ that come with a particular WP fixed point. To be precise, call $\sigma \in \ker \boldsymbol{A}^\dagger$ an *extension of* $\mathfrak{m}_{\cdot\to\cdot} \in \mathfrak{M}(\boldsymbol{A}^\dagger)$ if

$$\sum_{i=1}^{n} \mathbf{1}\{\mathfrak{m}_{v_i} \neq \mathtt{u},\, \sigma_i \neq 0\} + \sum_{s \in \mathbb{F}_q} \sum_{\ell \geqslant 0} \left| \sum_{i=1}^{n} \mathbf{1}\{d_{\boldsymbol{A}^\dagger}(v_i) = \ell,\, \mathfrak{m}_{v_i} = \mathtt{u}\}(\mathbf{1}\{\sigma_i = s\} - 1/q) \right| \tag{2.24}$$
$$= o(n).$$

Thus, σ is required to (mostly) respect the variables that $\mathfrak{m}_{\cdot\to\cdot}$ deems frozen under (2.3) by actually setting them to zero. Moreover, the variables deemed unfrozen according to $\mathfrak{m}_{\cdot\to\cdot}$ need to be assigned in a balanced manner, even when broken down to specific values ℓ of the variable degree, just like in (2.8). In fact, Propositions 5 and 6 show that a random kernel vector $\boldsymbol{\sigma}^\dagger$ is an $\boldsymbol{\alpha}$-extension of the standard messages $\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger)$. Hence,

letting $\boldsymbol{X}_\alpha$ be the number of pairs $(\mathfrak{m}_{\cdot\to\cdot}, \sigma)$ such that $\mathfrak{m}_{\cdot\to\cdot}$ is an α-WP fixed point of $\boldsymbol{A}^\dagger$ and σ is an extension of α, we see that $|\ker \boldsymbol{A}^\dagger| \sim \boldsymbol{X}_{\boldsymbol{\alpha}}$ w.h.p. By comparison, the following proposition, which we prove in Section 4, provides a first moment upper bound on $\boldsymbol{X}_\alpha$ for any $0 \leqslant \alpha \leqslant 1$ in terms of the function $\Phi_{d,k}$ from (1.1).

Proposition 7. *Let $d > 0, k \geqslant 3$. W.h.p. for all $\alpha \in [0,1]$ we have $\mathbb{E}[\boldsymbol{X}_\alpha \mid \mathfrak{D}] \leqslant q^{n\Phi_{d,k}(\alpha)+o(n)}$.*

Since for $d < d_k$ the function $\Phi_{d,k}$ attains its unique maximum at $\alpha = 0$ and $q^{n\Phi_{d,k}(0)} = q^{n-m}$, it is not very difficult to derive the estimate $\boldsymbol{\alpha} = o(1)$ w.h.p. from Proposition 7. From this, in turn, we can deduce that w.h.p. most vectors in the kernel are 'balanced', i.e., contain every value $s \in \mathbb{F}_q$ with about equal frequency. To be precise, for a vector $\sigma \in \mathbb{F}_q^n$ let $\rho(\sigma) = (\rho_s(\sigma))_{s\in\mathbb{F}_q}$ be the vector with entries $\rho_s(\sigma) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{\sigma_i = s\}$.

Corollary 8. *Let $\varepsilon > 0$. If $m < (1-\varepsilon)d_k n/k$, then w.h.p. we have*

$$\mathbb{E}\left[\|\rho(\boldsymbol{\sigma}^\dagger) - q^{-1}\mathbf{1}\|_2 \mid \boldsymbol{A}^\dagger\right] = o(1). \tag{2.25}$$

It is quite easy to calculate the expected number of vectors $\sigma \in \ker \boldsymbol{A}^\dagger$ with $\|\rho(\sigma) - q^{-1}\mathbf{1}\|_2 = o(1)$. Recall that we obtained $\boldsymbol{A}^\dagger$ from $\boldsymbol{A}$ by adding $\boldsymbol{t}$ extra rows with a single non-zero entry each. In Section 4.4 we prove the following.

Lemma 9. *For any $d > 0, k \geqslant 3$ there is $\eta > 0$ such that*
$\mathbb{E}\left|\{\sigma \in \ker \boldsymbol{A}^\dagger : \|\rho(\sigma) - q^{-1}\mathbf{1}\|_2 < \eta\} \mid \boldsymbol{t}\right| \leqslant (1+o(1))q^{n-m-\boldsymbol{t}}$.

Proof of Theorem 2 (i). Since the top m rows of $\boldsymbol{A}^\dagger$ are equal to $\boldsymbol{A}$, it suffices to prove that $\boldsymbol{A}^\dagger$ has full row rank w.h.p. Hence, let $\boldsymbol{y}^\dagger \in \mathbb{F}_q^{m+\boldsymbol{t}}$ be a uniformly random vector that is conditionally independent of $\boldsymbol{A}^\dagger$ given $\boldsymbol{t}$. In order to conclude that $\boldsymbol{A}^\dagger$ has full row rank w.h.p., we just need to show that

$$\mathbb{P}\left[\exists \sigma \in \mathbb{F}_q^n : \boldsymbol{A}^\dagger\sigma = \boldsymbol{y}^\dagger\right] \sim 1. \tag{2.26}$$

Let $\boldsymbol{Z}$ be the number of solutions σ to $\boldsymbol{A}^\dagger\sigma = \boldsymbol{y}^\dagger$. Because $\boldsymbol{y}^\dagger$ is independent of $\boldsymbol{A}^\dagger$ given $\boldsymbol{t}$, we have

$$\mathbb{E}\left[\boldsymbol{Z} \mid \boldsymbol{t}\right] = q^{n-m-\boldsymbol{t}}. \tag{2.27}$$

Further, let $\mathcal{B}$ be the event that $\boldsymbol{A}^\dagger$ enjoys the property (2.25). By Corollary 8 the mean of $\boldsymbol{Z}$ on $\mathcal{B}$ comes to

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{Z}\cdot\mathbf{1}\mathcal{B} \mid \boldsymbol{t}\right] &= \\ &\sum_{\substack{A^\dagger\in\mathbb{F}_q^{(m+\boldsymbol{t})\times n}\\ y^\dagger\in\mathbb{F}_q^{m+\boldsymbol{t}},\sigma\in\mathbb{F}_q^n}} \mathbf{1}\left\{A^\dagger\sigma = y^\dagger,\, A^\dagger\in\mathcal{B}\right\}\mathbb{P}\left[\boldsymbol{A}^\dagger = A^\dagger, \boldsymbol{y}^\dagger = y^\dagger \mid \boldsymbol{t}\right] = q^{n-m-\boldsymbol{t}}\mathbb{P}\left[\boldsymbol{A}^\dagger\in\mathcal{B}\right]\\ &\sim \mathbb{E}[\boldsymbol{Z}\mid\boldsymbol{t}].\end{aligned} \tag{2.28}$$

Similarly, Lemma 9 yields

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{Z}^2\cdot\mathbf{1}\mathcal{B}\mid\boldsymbol{t}\right]&=\mathbb{E}\left[\boldsymbol{Z}|\ker\boldsymbol{A}^\dagger|\cdot\mathbf{1}\mathcal{B}\mid\boldsymbol{t}\right]\\&=\sum_{A^\dagger,y^\dagger,\sigma}\mathbf{1}\left\{A^\dagger\sigma=y^\dagger,\,A^\dagger\in\mathcal{B}\right\}|\ker A^\dagger|\mathbb{P}\left[\boldsymbol{A}^\dagger=A^\dagger,\boldsymbol{y}^\dagger=y^\dagger\mid\boldsymbol{t}\right]\\&=q^{n-m-\boldsymbol{t}}\mathbb{E}\left[|\ker\boldsymbol{A}^\dagger|\cdot\mathbf{1}\mathcal{B}\mid\boldsymbol{t}\right]\leqslant(1+o(1))q^{2(n-m-\boldsymbol{t})}.\end{aligned}\tag{2.29}$$

Combining (2.27)–(2.29) with Chebyshev's inequality, we see that $\boldsymbol{Z}\sim q^{n-m-\boldsymbol{t}}>0$ w.h.p., which implies (2.26). □

2.5 Discussion

Preceding the seminal contribution of Dubois and Mandler [14] that determined the precise 3-XORSAT threshold, Creignou, Daudé and Dubois [11] obtained upper and lower bounds by means of the first and the second moment methods. These methods went on to become a mainstay of the theory of random constraint satisfaction problems, with numerous important additions [2, 13]. Independently of [26], a rigorous derivation of the k-XORSAT threshold for general k was outlined in [12], where the threshold was needed for an application to cuckoo hashing. The k-XORSAT threshold was further investigated from the viewpoint of the physicists' replica and cavity methods [22]. Moreover, the contributions [1, 17] conduct a detailed study of the geometry of the solution space of random k-XORSAT formulas.

Various different analyses of the pruning process have been put forward [8, 9, 15, 18, 20, 27, 29]. The methods employed in these works range from differential equations to branching processes to enumerative arguments. Since none of the proofs are particularly simple, we consider the fact that, in contrast to [14, 26], the present derivation of the k-XORSAT threshold gets by without an explicit investigation of the pruning process a significant plus.

The derivation of the full rank threshold [4] also avoided an analysis of the pruning process and instead relied on the Aizenman-Sims-Starr coupling argument from mathematical physics [3]. The main result of [4] is a variant of Theorem 2 with identically distributed rows. Specifically, the non-zero entries in the rows are drawn independently from a given distribution on $(\mathbb{F}_q\setminus\{0\})^k$. The present proof method can be easily adapted to cover this scenario, but also allows for the non-zero entries to be copied from a given infinite matrix $\mathfrak{A}$, in which case the rows need not be identically distributed anymore. Prior to [4], which still covers over 50 pages, an extension of the k-XORSAT threshold result to random matrices over $\mathbb{F}_3$ was obtained [16] by a generalisation of the moment method from [14, 26]. The article of over 80 pages requires computer assistance.

The techniques developed in [4] were extended to more general random matrix models with identically distributed rows [6]; the main result of that paper also implies the k-XORSAT threshold, but the proof is rather complicated. Additionally, for a still more general model of random matrices over general (not necessarily finite) fields an asymptotic formula for the normalised rank was obtain via the Aizenman-Sims-Starr scheme [7]. Furthermore, an independent result yields the asymptotic rank of the random matrix $\boldsymbol{A}$

over $\mathbb{F}_2$, albeit without obtaining the precise full rank threshold [10]. Here we employ the pinning technique from [7] (Lemma 3), which is an adaptation of the more general pinning method for discrete probability distributions developed in [24, 28].

Finally, a recent article [5] studies sparse square random matrices over $\mathbb{F}_2$ with independent entries. The main results, pertaining to the structure of the kernel of such a random matrices, evince a somewhat remarkable bifurcation that contrasts with the zero-one behaviour otherwise characteristic of probabilistic combinatorics. In the present paper we employ the mathematical formalisation of the WP message passing scheme developed in [5]. Furthermore, the article [5] also employed a moment computation similar to the one that we use to prove Proposition 7, but for a substantially different matrix model and towards a somewhat different overall result (an analysis of the kernel geometry rather than a proof that the matrix has full rank).

2.6 Preliminaries

We need to reflect on the function $\Phi_{d,k}$ and its maxima. Let

$$\phi_{d,k}(\alpha) = 1 - \exp(-d\alpha^{k-1}). \tag{2.30}$$

A tiny bit of calculus reveals that the functions $\phi_{d,k}$ from (2.30) and $\Phi_{d,k}$ from (1.1) are closely related as

$$\Phi'_{d,k}(\alpha) = d(k-1)\alpha^{k-2}\left(\phi_{d,k}(\alpha) - \alpha\right), \tag{2.31}$$
$$\Phi''_{d,k}(\alpha) = d(k-1)(k-2)\alpha^{k-3}\left(\phi_{d,k}(\alpha) - \alpha\right) - d(k-1)\alpha^{k-2}\left(1 - \phi'_{d,k}(\alpha)\right). \tag{2.32}$$

Thus, the fixed points $\alpha \in [0,1]$ of $\phi_{d,k}$ coincide with the stationary points of $\Phi_{d,k}$. In fact, the stable fixed points of $\phi_{d,k}$ are precisely the local maxima of $\Phi_{d,k}$. Moreover, a few lines of calculus reveal the following.

Fact 10. *Let $d > 0, k \geqslant 3$. The function $\phi_{d,k}$ has at most three distinct fixed points in the unit interval, which we denote by $\alpha_{\mathtt{u}}(d,k) \leqslant \alpha_{\mathtt{s}}(d,k) \leqslant \alpha_{\mathtt{f}}(d,k)$. There exists a critical value $0 < d^*_k < d_k$ such that*

- *for $d < d^*_k$ we have $\alpha_{\mathtt{u}}(d,k) \leqslant \alpha_{\mathtt{s}}(d,k) \leqslant \alpha_{\mathtt{f}}(d,k) = 0$,*
- *for $d = d^*_k$ we have $0 = \alpha_{\mathtt{u}}(d,k) < \alpha_{\mathtt{s}}(d,k) = \alpha_{\mathtt{f}}(d,k) < 1$,*
- *for $d > d^*_k$ we have $0 = \alpha_{\mathtt{u}}(d,k) < \alpha_{\mathtt{s}}(d,k) < \alpha_{\mathtt{f}}(d,k) < 1$.*

For $d < d_k$ the function $\Phi_{d,k}$ attains its unique maximum at 0, while $\alpha_{\mathtt{f}}(d,k)$ is the unique maximiser for $d > d_k$.

Additionally, we need the following elementary fact from linear algebra.

Fact 11 ([7, Lemma 2.5])**.** *Let A, B, C be matrices of sizes $M \times N$, $M' \times N$ and $M' \times N'$, respectively. Moreover, let $I \subseteq [N]$ be the set of non-zero columns of B and obtain B_0*

from B by replacing for every $i \in I \cap \mathcal{F}(A)$ the i-th column of B by zero. Unless I is a proper relation of A we have

$$\operatorname{nul}\begin{pmatrix} A & 0 \\ B & C \end{pmatrix} - \operatorname{nul} A + \operatorname{rk}(B_0\ C) = N'.$$

Further, we make a note of the degree distribution of the Tanner graph $G(\boldsymbol{A}^\dagger)$. Because the rows are chosen independently, the degrees of the variable nodes are asymptotically Poisson. More precisely, routine arguments show that the following is true.

Fact 12. *W.h.p. we have* $\sum_{\ell \geqslant 0} \exp(\ell)\,|\mathbb{P}\,[\mathrm{Po}(d) = \ell] - \sum_{i=1}^{n} \mathbf{1}\{d_{\boldsymbol{A}^\dagger}(v_i) = \ell\}| = o(n)$.

For the entropy of a probability distribution p on a finite set Ω we use the symbol

$$H(p) = -\sum_{\omega \in \Omega} p(\omega) \log p(\omega),$$

with the convention that $0 \log 0 = 0$. Finally, for a vector $\xi \in \mathbb{F}_q^N$ we write $\|\xi\|_h$ for the ℓ^h-norm of ξ, with the convention that $\|\xi\|_0 = |\mathrm{supp}\xi| = |\{i \in [N] : \xi_i \neq 0\}|$.

3 Warning Propagation

In this section we prove Propositions 5 and 6. We begin with some ruminations on short linear relations.

3.1 Short linear relations

The following lemma shows that if a matrix A possesses few short proper relations, then the same is true of any matrix A' obtained from A by adding a single row. Moreover, A and A' have more or less the same frozen variables.

Lemma 13. *For any $\delta' > 0$, $\ell' \geqslant 2$ there exist $\delta > 0, \ell \geqslant 2, N_0 > 0$ such that for any $N > N_0, M > 0$, any $M \times N$-matrix A and any matrix A' obtained from A by adding a single row the following is true. If A is (δ, ℓ)-free, then*

(i) A' is (δ', ℓ')-free, and

(ii) $|\mathcal{F}(A') \setminus \mathcal{F}(A)| < \delta' N$.

Proof. Set $\ell = 2\ell'$ and $\delta = \delta'^2 2^{-\ell-16}$. Assume for contradiction that A is (δ, ℓ)-free but that A' fails to be (δ', ℓ')-free. Let $\mathcal{I}'$ be the set of all proper relations I' of A' of size $|I'| = \ell'$ that fail to be proper relations of A. Since $\mathcal{F}(A) \subseteq \mathcal{F}(A')$, for any $y \in \mathbb{F}_q^{M+1}$ with

$$\emptyset \neq \mathrm{supp}(y^\top A') \subseteq I' \setminus \mathcal{F}(A') \subseteq I' \setminus \mathcal{F}(A)$$

we have $y_{M+1} \neq 0$. Furthermore, for sufficiently large N_0 the set $\mathcal{I}' \times \mathcal{I}'$ contains at least $(\delta' \binom{N}{\ell'})^2/8$ pairs (I', I'') such that $I' \cap I'' = \emptyset$. Given such a pair (I', I'') let $y, z \in \mathbb{F}_q^{M+1}$

be such that $\emptyset \neq \mathrm{supp}(y^\top A') \subseteq I' \setminus \mathcal{F}(A')$ and $\emptyset \neq \mathrm{supp}(z^\top A') \subseteq I'' \setminus \mathcal{F}(A')$. Since $y_{M+1}, z_{M+1} \neq 0$, there exists $\zeta \in \mathbb{F}_q \setminus \{0\}$ such that $y_{M+1} + \zeta z_{M+1} = 0$. Hence,

$$\begin{aligned}&\emptyset \neq \mathrm{supp}(((y_1 \cdots y_M) + \zeta(z_1 \cdots z_M))^\top A) \subseteq (I' \cup I'') \setminus \mathcal{F}(A)\\ &\text{and} \quad ((y_1 \cdots y_M) + \zeta(z_1 \cdots z_M))^\top A \neq 0.\end{aligned}$$

Consequently, $I' \cup I''$ is a proper relation of A of size $2\ell'$. Thus, A possesses at least $(\delta' \binom{N}{\ell'})^2/8$ such proper relations. However, choosing N_0 large enough, we obtain $(\delta' \binom{N}{\ell'})^2/8 > \delta \binom{N}{\ell}$, in contradiction to the fact that A is (δ, ℓ)-free. Concerning (ii), let $j, j' \in \mathcal{F}(A') \setminus \mathcal{F}(A)$ be two distinct indices that are frozen in A' but not in A. Then there exist vectors $y, z \in \mathbb{F}_q^{M+1}$ such that $\mathrm{supp}(y^\top A') = \{j\}$ and $\mathrm{supp}(z^\top A') = \{j'\}$. Since $j, j' \notin \mathcal{F}(A)$ we have $y_{M+1} \neq 0 \neq z_{M+1}$. Hence, there exists $\zeta \in \mathbb{F}_q \setminus \{0\}$ such that $y_{M+1} + \zeta z_{M+1} = 0$. Moreover,

$$\mathrm{supp}(((y_1 \cdots y_M) + \zeta(z_1 \cdots z_M))^\top A) = \{j, j'\}.$$

Thus, $\{j, j'\}$ is a proper relation of A. We therefore conclude that A possesses at least $\binom{|\mathcal{F}(A') \setminus \mathcal{F}(A)|}{2}$ proper relations of size two. Consequently, $\binom{|\mathcal{F}(A') \setminus \mathcal{F}(A)|}{2} < \delta \binom{N}{2}$, whence the desired bound $|\mathcal{F}(A') \setminus \mathcal{F}(A)| < \delta' N$ follows. □

Repeated application of Lemma 13 shows the following.

Corollary 14. *There exists $1 \ll \omega' = \omega'_n \ll \omega = \omega_n$ such that the following is true. Suppose that A is $(\omega, 1/\omega)$-free and that A' is obtained from A by adding at most ω' rows. Then A' is $(\omega', 1/\omega')$-free and $|\mathcal{F}(A') \setminus \mathcal{F}(A)| \leqslant n/\omega'$.*

3.2 Proof of Propositions 5

Proposition 5 posits that the standard WP messages from (2.1) are an approximate fixed point of the update rule (2.2) and that the labels defined in (2.3)–(2.4) match their intended semantics. The starting point of the proof is that the distribution of the random matrix $\boldsymbol{A}^\dagger$ remains asymptotically invariant under the following resampling operation.

Fact 15. *Let $\boldsymbol{A}^+$ be the matrix obtained from $\boldsymbol{A}^\dagger$ via the following operation.*

> *Choose a variable node $\boldsymbol{v} \in \{v_1, \ldots, v_n\}$ randomly, then independently for all $a \in \partial_{\boldsymbol{A}^\dagger} \boldsymbol{v}$ resample the neighbours of a other than $\boldsymbol{v}$ uniformly without replacement from $\{v_1, \ldots, v_n\} \setminus \{\boldsymbol{v}\}$.*

Then $\boldsymbol{A}^\dagger$ and $\boldsymbol{A}^+$ are identically distributed.

To establish the fixed point property (2.6) we are going to show that

$$\mathfrak{m}_{\boldsymbol{v} \to a}(\boldsymbol{A}^+) = \hat{\mathfrak{m}}_{\boldsymbol{v} \to a}(\boldsymbol{A}^+) \qquad \text{for all } a \in \partial_{\boldsymbol{A}^+} \boldsymbol{v} \text{ w.h.p.;} \tag{3.1}$$

then Markov's inequality implies that $\sum_{j=1}^n \sum_{a \in \partial_{\boldsymbol{A}^\dagger} v_j} \mathbf{1}\{\mathfrak{m}_{v_j \to a}(\boldsymbol{A}^\dagger) \neq \hat{\mathfrak{m}}_{v_j \to a}(\boldsymbol{A}^\dagger)\} = o(n)$ w.h.p. More specifically, we are going to exhibit an event $\mathcal{E}$ with $\mathbb{P}[\mathcal{E}] \sim 1$ such that (3.1) holds on $\mathcal{E}$ deterministically.

To define the event $\mathcal{E}$ pick a sequence $\Delta = \Delta(n) \gg 1$ that diverges slowly enough as $n \to \infty$. Moreover, obtain $\boldsymbol{A}^-$ from $\boldsymbol{A}^+$ by deleting all checks $a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}$. Now, let $\mathcal{E}$ be the event that the three following conditions hold.

E1 The second neighbourhood $\partial^2_{\boldsymbol{A}^+}\boldsymbol{v} = \{v_j : \exists a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v} : v_j \in \partial_{\boldsymbol{A}^+}a\} \setminus \{\boldsymbol{v}\}$ has size precisely $(k-1)|\partial_{\boldsymbol{A}^+}\boldsymbol{v}| \leqslant \Delta$.

E2 $\partial^2_{\boldsymbol{A}^+}\boldsymbol{v}$ is not a proper relation of $\boldsymbol{A}^-$.

E3 For all $a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}$ we have $\mathcal{F}(\boldsymbol{A}^+ \setminus \{a\}) \cap \partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}\} = \mathcal{F}(\boldsymbol{A}^-) \cap \partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}\}$.

Claim 16. *We have* $\mathbb{P}[\mathcal{E}] = 1 - o(1)$.

Proof. Condition **E1** asks that $\boldsymbol{v}$ have degree at most $\Delta/(k-1)$ and that the subgraph of $G(\boldsymbol{A}^\dagger)$ induced by the vertices of distance at most two from $\boldsymbol{v}$ be acyclic. Fact 12, Fact 15 and the independence of the positions of the non-zero entries in the different rows of $\boldsymbol{A}^\dagger$ imply that this is indeed the case w.h.p. Moreover, **E1** and the construction of $\boldsymbol{A}^+$ ensure that $\partial^2_{\boldsymbol{A}^+}\boldsymbol{v}$ is nothing but a random set of variable nodes of $G(\boldsymbol{A}^-)$ of size at most Δ. Since $\boldsymbol{A}^+$ contains the same $\boldsymbol{t}$ rows with ones in random positions that we added to $\boldsymbol{A}^\dagger$ by way of the pinning operation, Lemma 3 shows that $\boldsymbol{A}^+$ is $(\omega, 1/\omega)$-free with probability $1 - o(1/\omega)$ for a certain $\omega \gg 1$. Consequently, $\partial^2_{\boldsymbol{A}^+}\boldsymbol{v}$ is not a proper relation of $\boldsymbol{A}^+$ w.h.p., provided that $1 \ll \Delta \ll \omega$ diverges sufficiently slowly. Hence, **E2** holds w.h.p. Finally, $\boldsymbol{A}^+ \setminus \{a\}$ is obtained from $\boldsymbol{A}^-$ by adding at most Δ rows. Therefore, **E2** and Corollary 14 imply that **E3** is satisfied w.h.p., once again providing that $\Delta \to \infty$ sufficiently slowly. □

The following two claims deliver (3.1).

Claim 17. *Assume that $\mathcal{E}$ occurs and let $a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}$. If there exists $b \in \partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}$ such that $\mathtt{m}_{b\to\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{f}$, then $\mathtt{m}_{\boldsymbol{v}\to a}(\boldsymbol{A}^+) = \mathtt{f}$. Moreover, if $\mathtt{m}_{\boldsymbol{v}}(\boldsymbol{A}^+) \neq \mathtt{u}$, then $\boldsymbol{v} \in \mathcal{F}(\boldsymbol{A}^+)$.*

Proof. Let $b \in \partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}$ be such that $\mathtt{m}_{b\to\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{f}$. Then **E3** guarantees that $y \in \mathcal{F}(\boldsymbol{A}^-)$ for all $y \in \partial_{\boldsymbol{A}^+}b \setminus \{\boldsymbol{v}\}$. Therefore, for all $\sigma \in \ker(\boldsymbol{A}^+ \setminus \{a\}) \subseteq \ker(\boldsymbol{A}^- \setminus \{a\})$ and all $y \in \partial_{\boldsymbol{A}^+}b \setminus \{\boldsymbol{v}\}$ we have $\sigma_y = 0$, and consequently $\sigma_{\boldsymbol{v}} = 0$. Hence, $\boldsymbol{v} \in \mathcal{F}(\boldsymbol{A}^+ \setminus \{a\})$, and thus $\mathtt{m}_{\boldsymbol{v}\to a}(\boldsymbol{A}^+) = \hat{\mathtt{m}}_{\boldsymbol{v}\to a}(\boldsymbol{A}^+) = \mathtt{f}$ by (2.1). A similar argument yields the second assertion. □

Claim 18. *Assume that $\mathcal{E}$ occurs and let $a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}$. If $\mathtt{m}_{b\to\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{u}$ for all $b \in \partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}$, then $\mathtt{m}_{\boldsymbol{v}\to a}(\boldsymbol{A}^+) = \mathtt{u}$. Moreover, if $\mathtt{m}_{\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{u}$, then $\boldsymbol{v} \notin \mathcal{F}(\boldsymbol{A}^+)$.*

Proof. With Π, Π' suitable permutation matrices (to reshuffle the rows and columns appropriately), B a matrix of size $|\partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}| \times (n-1)$ and C a matrix of size $|\partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}| \times 1$, we can write

$$\boldsymbol{A}^+ \setminus \{a\} = \Pi \cdot \begin{pmatrix} \boldsymbol{A}^- \setminus \{\boldsymbol{v}\} & 0 \\ B & C \end{pmatrix} \cdot \Pi'. \tag{3.2}$$

Here the submatrix $(B\ C)$ corresponds to the checks $b \in \partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}$, and the last column $\binom{0}{C}$ represents $\boldsymbol{v}$. Obtain B_0 from B by replacing the columns corresponding to variable nodes $v_i \neq \boldsymbol{v}$ with $i \in \mathcal{F}(\boldsymbol{A}^-)$ by all-zero columns.

Now assume that $\mathcal{E}$ occurs and that for every $b \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}\setminus\{a\}$ there exists $u \in \partial_{\boldsymbol{A}^+}b\setminus\{\boldsymbol{v}\}$ such that $\mathtt{m}_{u\to b}(\boldsymbol{A}^+) = \mathtt{u}$. In fact, let $U = \{u \in \partial^2_{\boldsymbol{A}^+}\boldsymbol{v} : \mathtt{m}_{u\to b}(\boldsymbol{A}^+) = \mathtt{u}\}$. Then $U \cap \mathcal{F}(\boldsymbol{A}^-) = \emptyset$, because $\mathcal{F}(\boldsymbol{A}^-) \subseteq \mathcal{F}(\boldsymbol{A}^+ \setminus \{b\})$ for every $b \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}$. Due to **E1** for every column representing a variable $u \in U$ the u-column of B_0 contains precisely one non-zero entry. Therefore, $\mathrm{rk}(B_0\ C) = |\partial_{\boldsymbol{A}^*}\boldsymbol{v}\setminus\{a\}|$, i.e., the matrix $(B_0\ C)$ has full row rank. Since **E2** ensures that $\partial^2_{\boldsymbol{A}^+}\boldsymbol{v}$ is not a proper relation of $\boldsymbol{A}^-$, Fact 11 shows that

$$\mathrm{nul}\begin{pmatrix}\boldsymbol{A}^- \setminus \{\boldsymbol{v}\} & 0\\ B & C\end{pmatrix} = \mathrm{nul}(\boldsymbol{A}^+ \setminus \{a,\boldsymbol{v}\}) - |\partial_{\boldsymbol{A}^*}\boldsymbol{v} \setminus \{a\}| - 1. \tag{3.3}$$

Similarly, we can compute the rank of the matrix obtained by adding one more row with a single 1-entry in the last column, thereby expressly pinning $\boldsymbol{v}$:

$$\mathrm{nul}\begin{pmatrix}\boldsymbol{A}^- \setminus \{\boldsymbol{v}\} & 0\\ B & C\\ 0 & 1\end{pmatrix} = \mathrm{nul}(\boldsymbol{A}^+ \setminus \{a,\boldsymbol{v}\}) - |\partial_{\boldsymbol{A}^*}\boldsymbol{v} \setminus \{a\}| - 2 < \mathrm{nul}\begin{pmatrix}\boldsymbol{A}^- \setminus \{\boldsymbol{v}\} & 0\\ B & C\end{pmatrix}. \tag{3.4}$$

Combining (3.3)–(3.4), we conclude that the last coordinate n that represents $\boldsymbol{v}$ is unfrozen in $\begin{pmatrix}\boldsymbol{A}^- \setminus \{\boldsymbol{v}\} & 0\\ B & C\end{pmatrix}$; for otherwise the nullities on the left and right of (3.4) would have been equal. Hence, (3.2) shows that $\boldsymbol{v}$ is unfrozen in $\boldsymbol{A}^+ \setminus \{a\}$. Thus, $\mathtt{m}_{\boldsymbol{v}\to a}(\boldsymbol{A}^+) = \mathtt{u}$ by (2.1). A similar argument yields the second assertion. □

We proceed to investigate the check-to-variable messages.

Claim 19. *Assume that $\mathcal{E}$ occurs and let $a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}$. If $\mathtt{m}_{w\to a}(\boldsymbol{A}^+) = \mathtt{f}$ for all $w \in \partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}$, then $\mathtt{m}_{a\to\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{f}$.*

Proof. If $\mathtt{m}_{w\to a}(\boldsymbol{A}^+) = \mathtt{f}$, then $w \in \mathcal{F}(\boldsymbol{A}^+ \setminus \{a\})$ by the definition (2.1) of the standard messages. Hence, **E3** guarantees that $w \in \mathcal{F}(\boldsymbol{A}^-)$ for all $w \in \partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}$. Further, since $\mathcal{F}(\boldsymbol{A}^-) \subseteq \mathcal{F}(\boldsymbol{A}^- \setminus (\partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}))$ we obtain from (2.1) that $\mathtt{m}_{a\to\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{f}$. □

Claim 20. *Assume that $\mathcal{E}$ occurs and let $a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}$. If there exists $w \in \partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}\}$ such that $\mathtt{m}_{w\to a}(\boldsymbol{A}^+) = \mathtt{u}$, then $\mathtt{m}_{a\to\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{u}$.*

Proof. Let $w \in \partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}\}$ be such that $\mathtt{m}_{w\to a}(\boldsymbol{A}^+) = \mathtt{u}$. Then the definition (2.1) of $\mathtt{m}_{w\to a}(\boldsymbol{A}^+)$ ensures that $w \notin \mathcal{F}(\boldsymbol{A}^+ \setminus \{a\})$. Since $\mathcal{F}(\boldsymbol{A}^-) \subseteq \mathcal{F}(\boldsymbol{A}^+ \setminus \{a\})$, we conclude that $w \notin \mathcal{F}(\boldsymbol{A}^-)$. Further, for suitable permutation matrices Π, Π' we obtain $D \in \mathbb{F}_q^{n-1}$ and $\chi \in \mathbb{F}_q \setminus \{0\}$ such that

$$\boldsymbol{A}^+ \setminus (\partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}) = \Pi \cdot \begin{pmatrix}\boldsymbol{A}^- & 0\\ D & \chi\end{pmatrix} \cdot \Pi'; \tag{3.5}$$

thus, the permutation matrices Π, Π' are chosen such that they swap the $\boldsymbol{v}$-column to the last column and the a-row to the last row. Hence, the last row (D, χ) represents a. Now obtain D_0 from D by replacing all entries corresponding to variable nodes from $\mathcal{F}(\boldsymbol{A}^+) \setminus \{\boldsymbol{v}\}$ by 0. Then due to **E2**, Fact 11 shows that

$$\operatorname{nul}\begin{pmatrix}\boldsymbol{A}^- & 0\\ D & \chi\end{pmatrix} = \operatorname{nul}(\boldsymbol{A}^-) \qquad \text{and} \qquad \operatorname{nul}\begin{pmatrix}\boldsymbol{A}^- & 0\\ D & \chi\\ 0 & 1\end{pmatrix} = \operatorname{nul}\begin{pmatrix}\boldsymbol{A}^- & 0\\ D & \chi\end{pmatrix} - 1.$$

Hence, as in the proof of Claim 18 we obtain $\boldsymbol{v} \notin \mathcal{F}(\boldsymbol{A}^+ \setminus (\partial_{\boldsymbol{A}^+}\boldsymbol{v} \setminus \{a\}))$. Thus, $\mathtt{m}_{a\to\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{u}$ by (2.1). □

Proof of Proposition 5. Claims 16–20 directly imply that

$$\sum_{i=1}^{m}\sum_{j=1}^{n} \mathbf{1}\left\{\mathtt{m}_{v_j\to a_i}(\boldsymbol{A}^\dagger) \neq \hat{\mathtt{m}}_{v_j\to a_i}(\boldsymbol{A}^\dagger)\right\} = o(n) \text{and}$$
$$\sum_{i=1}^{m}\sum_{j=1}^{n} \mathbf{1}\left\{\mathtt{m}_{a_i\to v_j}(\boldsymbol{A}^\dagger) \neq \hat{\mathtt{m}}_{a_i\to v_j}(\boldsymbol{A}^\dagger)\right\} = o(n),$$

whence we obtain (2.6). Similarly, (2.7) follows from Claims 16–18.

Finally, in light of (2.7), to prove (2.8) it suffices to consider variables v_j with $j \notin \mathcal{F}(\boldsymbol{A}^\dagger)$. Hence, let $j, j' \in [n] \setminus \mathcal{F}(\boldsymbol{A}^\dagger)$ be two distinct indices such that $\{i, j\}$ is not a proper relation of $\boldsymbol{A}^\dagger$; Corollary 4 shows that this last property is violated for at most $o(n^2)$ pairs j, j'. Then the projection $\sigma \in \ker \boldsymbol{A}^\dagger \mapsto (\sigma_j, \sigma_{j'}) \in \mathbb{F}_q^2$ is an epimorphism. Therefore, for any $s, t \in \mathbb{F}_q^2$ we have $\left|\{\sigma \in \ker \boldsymbol{A}^\dagger : \sigma_i = s,\, \sigma_j = t\}\right| = q^{-2}|\ker \boldsymbol{A}^\dagger|$. Consequently, if $\boldsymbol{\sigma}^\dagger \in \ker \boldsymbol{A}^\dagger$ is drawn randomly, then for $(1 - \boldsymbol{\alpha} + o(1))^2 n^2$ pairs $j, j' \notin \mathcal{F}(\boldsymbol{A}^\dagger)$ the random variables $\boldsymbol{\sigma}_j^\dagger, \boldsymbol{\sigma}_{j'}^\dagger$ are independent and uniformly distributed. Thus, Chebyshev's inequality shows that given $\mathcal{E}$ for all $s \in \mathbb{F}_q$ we have

$$|\{j \in [n] \setminus \mathcal{F}(\boldsymbol{A}^\dagger) : \boldsymbol{\sigma}_j^\dagger = s\}| = (1 - \boldsymbol{\alpha} + o(1))|)n/q \qquad \text{w.h.p.},$$

whence we obtain (2.8). □

3.3 Proof of Proposition 6

The proof employs arguments broadly similar to those from the proof of Proposition 5. The main difference is that we are going to consider a uniformly random pair $(\boldsymbol{v}, \boldsymbol{v}')$ of variable nodes, rather than a single variable node. We begin by estimating the sizes $|\Delta_{z,\ell}(\mathtt{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger)) \times \Delta_{z',\ell'}(\mathtt{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger))|$ for $z, z' \in \{\mathtt{u}, \mathtt{s}, \mathtt{f}\}$, $\ell \in \mathcal{D}(z)$ and $\ell' \in \mathcal{D}(z')$. Similarly as in Section 3.2 obtain $\boldsymbol{A}^-$ from $\boldsymbol{A}^\dagger$ by deleting all checks $a \in \partial_{\boldsymbol{A}^\dagger}\boldsymbol{v} \cup \partial_{\boldsymbol{A}^\dagger}\boldsymbol{v}'$.

Fact 21. *Let $\boldsymbol{A}^+$ be the matrix obtained from $\boldsymbol{A}^\dagger$ via the following operation.*

> *Independently for all $a \in \partial_{\boldsymbol{A}^\dagger}\boldsymbol{v} \cup \partial_{\boldsymbol{A}^\dagger}\boldsymbol{v}'$ resample the neighbours of a other than $\boldsymbol{v}, \boldsymbol{v}'$ uniformly without replacement from $\{v_1, \ldots, v_n\} \setminus \{\boldsymbol{v}\}$.*

Then $\boldsymbol{A}^\dagger$ and $\boldsymbol{A}^+$ have total variation distance $o(1)$.

Proof. Given that $\boldsymbol{v}, \boldsymbol{v}'$ have distance at least four in both $G(\boldsymbol{A}^\dagger)$ and $G(\boldsymbol{A}^+)$, the Tanner graphs of $\boldsymbol{A}^\dagger$, $\boldsymbol{A}^+$ are identically distributed. Moreover, the probability that $\boldsymbol{v}, \boldsymbol{v}'$ have distance less than four is bounded by $n^{-1+o(1)}$. □

The plan is to derive the following joint probability formula, and then follow up with Chebyshev's inequality.

Lemma 22. *W.h.p. we have*

$$\mathbb{P}\left[\boldsymbol{v} \in \Delta_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^+)), \boldsymbol{v}' \in \Delta_{z',\ell'}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^+)) \mid \boldsymbol{A}^\dagger\right] = \bar{\Delta}_{z,\ell}(\boldsymbol{\alpha})\bar{\Delta}_{z',\ell'}(\boldsymbol{\alpha}) + o(1).$$

Towards the proof of Lemma 22 let $\mathcal{E}'$ be the event that the following statements hold; let $\Delta \gg 1$ diverge sufficiently slowly.

E0′ we have $|\partial_{\boldsymbol{A}^+}\boldsymbol{v}| = \ell_{\mathtt{uu}} + \ell_{\mathtt{fu}} + \ell_{\mathtt{uf}} + \ell_{\mathtt{ff}}$ and $|\partial_{\boldsymbol{A}^+}\boldsymbol{v}'| = \ell'_{\mathtt{uu}} + \ell'_{\mathtt{fu}} + \ell'_{\mathtt{uf}} + \ell'_{\mathtt{ff}}$.

E1′ the second neighbourhoods $\partial^2_{\boldsymbol{A}^+}\boldsymbol{v}, \partial^2_{\boldsymbol{A}^+}\boldsymbol{v}'$ satisfy

$$\boldsymbol{v}, \boldsymbol{v}' \notin \partial^2_{\boldsymbol{A}^+}\boldsymbol{v} \cup \partial^2_{\boldsymbol{A}^+}\boldsymbol{v}', |\partial^2_{\boldsymbol{A}^+}\boldsymbol{v}| = (k-1)|\partial_{\boldsymbol{A}^+}\boldsymbol{v}| \leqslant \Delta, |\partial^2_{\boldsymbol{A}^+}\boldsymbol{v}'| = (k-1)|\partial_{\boldsymbol{A}^+}\boldsymbol{v}'| \leqslant \Delta.$$

E2′ we have $|\mathcal{F}(\boldsymbol{A}^\dagger) \setminus \mathcal{F}(\boldsymbol{A}^-)| = o(n)$ and $|\mathcal{F}(\boldsymbol{A}^+) \setminus \mathcal{F}(\boldsymbol{A}^-)| = o(n)$.

E3′ for all $a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v} \cup \partial_{\boldsymbol{A}^+}\boldsymbol{v}'$ we have $\mathcal{F}(\boldsymbol{A}^+ \setminus \{a\}) \cap \partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}, \boldsymbol{v}'\} = \mathcal{F}(\boldsymbol{A}^-) \cap \partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}, \boldsymbol{v}'\}$.

E4′ for all $v \in \{\boldsymbol{v}, \boldsymbol{v}'\}$ and $a \in \partial_{\boldsymbol{A}^+}v$ we have $\mathfrak{m}_{v\to a}(\boldsymbol{A}^+) = \hat{\mathfrak{m}}_{v\to a}(\boldsymbol{A}^+)$, $\mathfrak{m}_{a\to v}(\boldsymbol{A}^+) = \hat{\mathfrak{m}}_{a\to v}(\boldsymbol{A}^+)$.

Thus, **E0′** provides that the degrees of $\boldsymbol{v}, \boldsymbol{v}'$ match the sum of the entries of ℓ, ℓ'. Moreover, **E1′** ensures that $\boldsymbol{v}, \boldsymbol{v}'$ have distance at least four and that their second neighbourhoods are acyclic. Further, **E2′** provides that $\boldsymbol{A}^-, \boldsymbol{A}^\dagger, \boldsymbol{A}^+$ have about the same number of frozen variables. In particular, **E3′** demands that the frozen variables in the second neighbourhood of $\boldsymbol{v}, \boldsymbol{v}'$ coincide in $\boldsymbol{A}^+$ and $\boldsymbol{A}^-$. Finally, **E4** posits that the messages that touch $\boldsymbol{v}, \boldsymbol{v}'$ are invariant under the WP update (2.2).

Claim 23. *We have* $\mathbb{P}[\mathcal{E}' \mid \mathbf{E0'}] = 1 - o(1)$ *and*

$$\mathbb{P}[\mathbf{E0'}] = \mathbb{P}[\mathrm{Po}(d) = \ell_{\mathtt{uu}} + \ell_{\mathtt{fu}} + \ell_{\mathtt{uf}} + \ell_{\mathtt{ff}}]\,\mathbb{P}[\mathrm{Po}(d) = \ell'_{\mathtt{uu}} + \ell'_{\mathtt{fu}} + \ell'_{\mathtt{uf}} + \ell'_{\mathtt{ff}}] + o(1). \quad (3.6)$$

Proof. The estimate (3.6) is an immediate consequence of Fact 12. Regarding the probability of $\mathcal{E}'$ given **E0′**, the same arguments as in the proof of Claim 16 show that **E1′**–**E3′** follow from Fact 12, Corollary 4 and Corollary 14. Furthermore, **E4′** follows from Eq. (2.7) from Proposition 5 and Fact 21. □

Proof of Lemma 22. Let $\boldsymbol{X} = |\partial_{\boldsymbol{A}^+}\boldsymbol{v}|$, $\boldsymbol{X}' = |\partial_{\boldsymbol{A}^+}\boldsymbol{v}'|$ be the degrees of $\boldsymbol{v}, \boldsymbol{v}'$. Moreover, let

$$\boldsymbol{X}_{\mathtt{f}} = \sum_{a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}} \mathbf{1}\{\partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}\} \subseteq \mathcal{F}(\boldsymbol{A}^-)\}, \qquad \boldsymbol{X}_{\mathtt{u}} = \boldsymbol{X} - \boldsymbol{X}_{\mathtt{f}},$$
$$\boldsymbol{X}'_{\mathtt{f}} = \sum_{a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}'} \mathbf{1}\{\partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}'\} \subseteq \mathcal{F}(\boldsymbol{A}^-)\}, \qquad \boldsymbol{X}'_{\mathtt{u}} = \boldsymbol{X}' - \boldsymbol{X}'_{\mathtt{f}}.$$

Additionally, let $\mathcal{X} = \{\boldsymbol{X}_{\mathtt{f}} = \ell_{\mathtt{ff}} + \ell_{\mathtt{fu}}, \boldsymbol{X}_{\mathtt{u}} = \ell_{\mathtt{uf}} + \ell_{\mathtt{uu}}\}$ and $\mathcal{X}' = \{\boldsymbol{X}'_{\mathtt{f}} = \ell'_{\mathtt{ff}} + \ell'_{\mathtt{fu}}, \boldsymbol{X}'_{\mathtt{u}} = \ell'_{\mathtt{uf}} + \ell'_{\mathtt{uu}}\}$. We are going to argue that Proposition 5 and Claim 23 imply

$$\mathbb{P}\left[\boldsymbol{v} \in \Delta_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^+)), \boldsymbol{v}' \in \Delta_{z',\ell'}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^+)) \mid \mathcal{E}'\right] = \mathbb{P}\left[\mathcal{X} \cap \mathcal{X}' \mid \mathcal{E}'\right] + o(1). \tag{3.7}$$

Indeed, the WP fixed point property **E4′** ensures that the WP messages that $\boldsymbol{v}, \boldsymbol{v}'$ send out to their neighbouring check nodes are determined by the incoming messages. Furthermore, **E3′** provides that for every $a \in \partial_{\boldsymbol{A}^+}\boldsymbol{v}$ we have $\mathfrak{m}_{a\to\boldsymbol{v}}(\boldsymbol{A}^+) = \mathtt{f}$ iff $\partial_{\boldsymbol{A}^+}a \setminus \{\boldsymbol{v}\} \subseteq \mathcal{F}(\boldsymbol{A}^-)$, and similarly for $\boldsymbol{v}'$. Consequently, (2.2), (2.3) and (2.4) show that on $\mathcal{E}'$ the random variables $\boldsymbol{X}_{\mathtt{f}}, \boldsymbol{X}_{\mathtt{u}}, \boldsymbol{X}'_{\mathtt{f}}, \boldsymbol{X}'_{\mathtt{u}}$ capture the salient information supplied by the incoming messages $\mathfrak{m}_{\cdot\to\boldsymbol{v}}(\boldsymbol{A}^+), \mathfrak{m}_{\cdot\to\boldsymbol{v}'}(\boldsymbol{A}^+)$, whence we obtain (3.7).

Further, we claim that if $\boldsymbol{A}^\dagger$ satisfies **E0′**, then

$$\mathbb{P}\left[\mathcal{X} \cap \mathcal{X}' \mid \boldsymbol{A}^\dagger\right] = \boldsymbol{\alpha}^{(k-1)(\ell_{\mathtt{fu}}+\ell_{\mathtt{ff}}+\ell'_{\mathtt{fu}}+\ell'_{\mathtt{ff}})}(1-\boldsymbol{\alpha}^{k-1})^{\ell_{\mathtt{uu}}+\ell_{\mathtt{uf}}+\ell'_{\mathtt{uu}}+\ell'_{\mathtt{uf}}} + o(1); \tag{3.8}$$

for by construction the new second neighbours of $\boldsymbol{v}, \boldsymbol{v}'$ are chosen uniformly. Hence, due to **E2′** the probability that any specific second neighbour belongs to $\mathcal{F}(\boldsymbol{A}^-)$ equals $\boldsymbol{\alpha} + o(1)$, and due to **E1′** these events are asymptotically independent. Finally, we combine (3.6), (3.7) and (3.8) to complete the proof. □

In order to estimate the sizes of the sets $\Gamma_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\dagger))$, we let $\boldsymbol{a}, \boldsymbol{a}'$ be a random pair of distinct check nodes. Let $\boldsymbol{A}^\#$ be the matrix obtained from $\boldsymbol{A}^\dagger$ by resampling the neighbours of $\boldsymbol{a}, \boldsymbol{a}'$ independently. Then $\boldsymbol{A}^\#$ and $\boldsymbol{A}^\dagger$ are identically distributed. In analogy to Lemma 22, we prove the following.

Lemma 24. *Let $z, z' \in \{\mathtt{u}, \mathtt{f}, \mathtt{s}\}$ and let $\ell \in \mathcal{G}(z), \ell' \in \mathcal{G}(z')$. W.h.p. we have*

$$\mathbb{P}\left[\boldsymbol{a} \in \Gamma_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\#)), \boldsymbol{a}' \in \Gamma_{z',\ell'}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^\#)) \mid \boldsymbol{A}^\dagger\right] = \bar{\Gamma}_{z,\ell}(\boldsymbol{\alpha})\bar{\Gamma}_{z',\ell'}(\boldsymbol{\alpha}) + o(1).$$

Proof. Consider the following event $\mathcal{A}$:

A1 the neighbourhoods $\partial_{\boldsymbol{A}^\#}\boldsymbol{a}, \partial_{\boldsymbol{A}^\#}\boldsymbol{a}'$ are disjoint.

A2 we have $|\mathcal{F}(\boldsymbol{A}^\#) \setminus \mathcal{F}(\boldsymbol{A}^\dagger \setminus \{\boldsymbol{a}, \boldsymbol{a}'\})| = o(n)$ and $|\mathcal{F}(\boldsymbol{A}^\dagger) \setminus \mathcal{F}(\boldsymbol{A}^\dagger \setminus \{\boldsymbol{a}, \boldsymbol{a}'\})| = o(n)$.

A3 we have $\mathcal{F}(\boldsymbol{A}^\# \setminus \{\boldsymbol{a}\}) \cap \partial_{\boldsymbol{A}^\#}\boldsymbol{a} = \mathcal{F}(\boldsymbol{A}^\dagger \setminus \{\boldsymbol{a}, \boldsymbol{a}'\}) \cap \partial_{\boldsymbol{A}^\#}\boldsymbol{a}$ and $\mathcal{F}(\boldsymbol{A}^\# \setminus \{\boldsymbol{a}'\}) \cap \partial_{\boldsymbol{A}^\#}\boldsymbol{a}' = \mathcal{F}(\boldsymbol{A}^\dagger \setminus \{\boldsymbol{a}, \boldsymbol{a}'\}) \cap \partial_{\boldsymbol{A}^\#}\boldsymbol{a}'$.

A4 for all $v \in \partial_{\boldsymbol{A}^\#}\boldsymbol{a}$ we have $\mathfrak{m}_{\boldsymbol{a}\to v}(\boldsymbol{A}^\#) = \hat{\mathfrak{m}}_{\boldsymbol{a}\to v}(\boldsymbol{A}^\#)$ and for all $v \in \partial_{\boldsymbol{A}^\#}\boldsymbol{a}'$ we have $\mathfrak{m}_{\boldsymbol{a}'\to v}(\boldsymbol{A}^\#) = \hat{\mathfrak{m}}_{\boldsymbol{a}'\to v}(\boldsymbol{A}^\#)$.

Then Corollary 4, Proposition 5 and Corollary 14 show that

$$\mathbb{P}\left[\mathcal{A}\right] = 1 - o(1). \tag{3.9}$$

Further, let

$$\boldsymbol{Y}_{\mathtt{f}} = |\partial_{\boldsymbol{A}^\#}\boldsymbol{a} \cap \mathcal{F}(\boldsymbol{A}^\dagger \setminus \{\boldsymbol{a}, \boldsymbol{a}'\})|, \qquad \boldsymbol{Y}_{\mathtt{u}} = k - \boldsymbol{Y}_{\mathtt{f}},$$
$$\boldsymbol{Y}'_{\mathtt{f}} = |\partial_{\boldsymbol{A}^\#}\boldsymbol{a}' \cap \mathcal{F}(\boldsymbol{A}^\dagger \setminus \{\boldsymbol{a}, \boldsymbol{a}'\})|, \qquad \boldsymbol{Y}'_{\mathtt{u}} = k - \boldsymbol{Y}'_{\mathtt{f}}.$$

Also let $\mathcal{Y} = \{\boldsymbol{Y}_{\mathtt{f}} = \ell_{\mathtt{ff}} + \ell_{\mathtt{fu}}\}$ and $\mathcal{Y}' = \{\boldsymbol{Y}'_{\mathtt{f}} = \ell'_{\mathtt{ff}} + \ell'_{\mathtt{fu}}\}$. We claim that

$$\mathbb{P}\left[\boldsymbol{a} \in \Gamma_{z,\ell}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^{\#})), \boldsymbol{v}' \in \Gamma_{z',\ell'}(\mathfrak{m}_{\cdot\to\cdot}(\boldsymbol{A}^{\#})) \mid \mathcal{A}\right] = \mathbb{P}\left[\mathcal{Y} \cap \mathcal{Y}' \mid \mathcal{A}\right] + o(1); \tag{3.10}$$

for **A4** provides that the messages that $\boldsymbol{a}, \boldsymbol{a}'$ send out to their neighbours are determined by the incoming messages via (2.2). Moreover, **A3** ensures that for $v \in \partial_{\boldsymbol{A}^{\#}}\boldsymbol{a}$ we have $\mathfrak{m}_{v\to\boldsymbol{a}}(\boldsymbol{A}^{\#}) = \mathtt{f}$ iff $v \in \mathcal{F}(\boldsymbol{A}^{\dagger} \setminus \{\boldsymbol{a}, \boldsymbol{a}'\})$, and similarly for $v' \in \partial_{\boldsymbol{A}^{\#}}\boldsymbol{a}'$.

Finally, since $\boldsymbol{A}^{\#}$ is obtained by resampling the neighbourhoods of $\boldsymbol{a}, \boldsymbol{a}'$, **A1**–**A2** show that

$$\mathbb{P}\left[\mathcal{Y} \cap \mathcal{Y}' \mid \boldsymbol{A}^{\dagger}\right] = \boldsymbol{\alpha}^{\ell_{\mathtt{fu}}+\ell_{\mathtt{ff}}+\ell'_{\mathtt{fu}}+\ell'_{\mathtt{ff}}}(1-\boldsymbol{\alpha})^{\ell_{\mathtt{uu}}+\ell_{\mathtt{uf}}+\ell'_{\mathtt{uu}}+\ell'_{\mathtt{uf}}} + o(1). \tag{3.11}$$

Thus, the assertion follows from (3.9)–(3.11). □

Proof of Proposition 6. The proposition follows from Fact 12, Lemmas 22 and 24 and Chebyshev. □

4 Moment computations

In this section we prove Proposition 7 and Corollary 8 and complete the proof of Theorem 2. Our principal tool will be moment computations. In particular, we will compute the mean of the number $\boldsymbol{X}_{\alpha}$ of α-extensions for $\alpha \in [0, 1]$. Crucially, because the definitions (2.22)–(2.23) prescribe the correct 'quenched' statistics provided by (2.16)–(2.21) as well as an approximate version of the WP fixed point property (2.5), the ensuing calculations turn out to be tight as well as relatively elegant. This manifests itself in the fact that we ultimately recover the function $\Phi_{d,k}$ from (1.1).

4.1 Counting WP fixed points

We begin by calculating the expected number of α-WP fixed points, for which we resort to the pairing model of the random bipartite Tanner graph. To this end we condition on the σ-algebra $\mathfrak{D}$ generated by the degrees $d_{\boldsymbol{A}^{\dagger}}(v_j)$ of the variable nodes and by $\boldsymbol{t}$. Given $\mathfrak{D}$ let

$$\mathfrak{V} = \bigcup_{j=1}^{n}\{v_j\} \times [d_{\boldsymbol{A}^{\dagger}}(v_j)] \qquad \text{and} \qquad \mathfrak{F} = \{b_1 \cup \cdots \cup b_{\boldsymbol{t}}\} \cup \bigcup_{i=1}^{m}\{a_i\} \times [k]$$

be sets of variable and check clones; here $b_1, \ldots, b_{\boldsymbol{t}}$ represent the checks that the pinning operation from Section 2.1 induces. A *pairing* is a bijection $\pi : \mathfrak{V} \to \mathfrak{F}$. Let $\mathfrak{P}$ be the set of all pairings. As usual, we construct a Tanner graph $G(\boldsymbol{\pi})$ by drawing a $\boldsymbol{\pi} \in \mathfrak{P}$ uniformly at random and contracting the clones into single vertices. This graph may possess multi-edges, in contrast to the random graph $G(\boldsymbol{A}^{\dagger})$. However, it is well known that once we condition on the event $\mathfrak{S}$ that $G(\boldsymbol{\pi})$ is simple, the distribution of $G(\boldsymbol{\pi})$ coincides with that of $G(\boldsymbol{A}^{\dagger})$. Moreover, routine arguments along the lines of [19, Chapter 9] show the following.

Fact 25. *For any $d > 0, k \geqslant 3$ w.h.p. we have $\mathbb{P}[\mathfrak{S} \mid \mathfrak{D}] = \Omega(1)$.*

In order to calculate the expected number of α-WP fixed points of $G(\boldsymbol{\pi})$ we compute the total number of pairings $\pi \in \mathfrak{P}$ together with appropriate $\{\mathtt{u}, \mathtt{f}\}$-valued annotations of the clones. To be precise, an *α-cover* $(\pi, \mathfrak{p})$ consists of a pairing π and a map $\mathfrak{p} : \mathfrak{V} \cup \mathfrak{F} \to \{\mathtt{u}, \mathtt{f}\}^2, (x, h) \mapsto \mathfrak{p}(x, h) = (\mathfrak{p}_1(x, h), \mathfrak{p}_2(x, h))$ that satisfy the following conditions.

COV1 For all $(x, h) \in \mathfrak{V} \cup \mathfrak{F}$ we have $(\mathfrak{m}_1(\pi(x, h)), \mathfrak{m}_2(\pi(x, h))) = (\mathfrak{m}_2(x, h), \mathfrak{m}_1(x, h))$.

COV2 For all but $o(n)$ pairs (v_j, l) with $j \in [n]$ and $l \in [d_{\boldsymbol{A}^\dagger}(v_i)]$ we have

$$\mathfrak{p}_2(v_j, l) = \begin{cases} \mathtt{f} & \text{if } \mathfrak{p}_1(v_j, h) = \mathtt{f} \text{ for some } h \in [d_{\boldsymbol{A}^\dagger}(v_j)] \setminus \{l\}, \\ \mathtt{u} & \text{otherwise.} \end{cases}$$

COV3 For all but $o(n)$ pairs (a_i, l) with $i \in [m]$ and $l \in [d_{\boldsymbol{A}^\dagger}(a_i)]$ we have

$$\mathfrak{p}_2\left(a_i, l\right) = \begin{cases} \mathtt{f} & \text{if } \mathfrak{p}_1\left(a_i, h\right) = \mathtt{f} \text{ for all } h \in [k] \setminus \{l\}, \\ \mathtt{u} & \text{otherwise.} \end{cases}$$

COV4 For any $z \in \{\mathtt{f}, \mathtt{s}, \mathtt{u}\}$, $\ell = (\ell_{\mathtt{uu}}, \ell_{\mathtt{uf}}, \ell_{\mathtt{fu}}, \ell_{\mathtt{ff}}) \in \mathcal{L}$, $i \in [m]$ and $j \in [n]$ let

$$\mathfrak{p}(v_j) = \begin{cases} \mathtt{f} \text{ if } \mathfrak{p}_1(v_j, l) = \mathtt{f} \text{ for at least two } l \in [d_{\boldsymbol{A}^\dagger}(v_j)], \\ \mathtt{s} \text{ if } \mathfrak{p}_1(v_j, l) = \mathtt{f} \text{ for precisely one } l \in [d_{\boldsymbol{A}^\dagger}(v_j)], \\ \mathtt{u} \text{ otherwise,} \end{cases} \tag{4.1}$$

$$\mathfrak{p}(a_i) = \begin{cases} \mathtt{f} \text{ if } \mathfrak{p}_1(a_i, l) = \mathtt{f} \text{ for all } l \in [d_{\boldsymbol{A}^\dagger}(a_i)], \\ \mathtt{s} \text{ if } \mathfrak{p}_1(a_i, l) = \mathtt{f} \text{ for all but precisely one } l \in [d_{\boldsymbol{A}^\dagger}(a_i)], \\ \mathtt{u} \text{ otherwise,} \end{cases} \tag{4.2}$$

$$\boldsymbol{\Delta}(z, \ell) = \sum_{i=1}^{n} \mathbf{1}\left\{\mathfrak{p}\left(v_j\right) = z\right\} \prod_{x,y \in \{\mathtt{u},\mathtt{f}\}} \mathbf{1}\left\{\left|\{l \in [d_{\boldsymbol{A}^\dagger}(v_j)] : \mathfrak{p}_1(v_j, l) = x,\ \mathfrak{p}_2(v_j, l) = y\}\right| = \ell_{xy}\right\}, \tag{4.3}$$

$$\boldsymbol{\Gamma}(z, \ell) = \sum_{i=1}^{m} \mathbf{1}\left\{\mathfrak{p}(a_i) = z\right\} \prod_{x,y \in \{\mathtt{u},\mathtt{f}\}} \mathbf{1}\left\{\left|\{l \in [d_{\boldsymbol{A}^\dagger}(a_i)] : \mathfrak{p}_1(a_i, l) = x,\ \mathfrak{p}_2(a_i, l) = y\}\right| = \ell_{xy}\right\}. \tag{4.4}$$

Then

$$\boldsymbol{\Delta}(z, \ell) = n\bar{\Delta}_{z,\ell}(\alpha) + o(n), \qquad \boldsymbol{\Gamma}(z, \ell) = m\bar{\Gamma}_{z,\ell}(\alpha) + o(n). \tag{4.5}$$

Condition **COV1** provides consistency of the labels associated with the paired clones. Moreover, **COV2**–**COV3** impose the fixed point condition (2.5) on $(\pi, \mathfrak{p})$. Similarly, the labels (4.1)–(4.2) mimic the definitions (2.3)–(2.4). Finally, (4.3)–(4.5) ensure that the statistics of the labels/messages are in line with the correct 'quenched' values (2.16)–(2.21) (see Proposition 6). The following lemma determines the size of the set $\mathfrak{C}(\alpha)$ of all α-covers.

Lemma 26. *W.h.p. we have* $\mathfrak{C}(\alpha) = \exp(o(n))(km)!k!^m \prod_{i=1}^n d_{\boldsymbol{A}^\dagger}(v_i)!$.

To prove Lemma 26 we begin with the following straightforward counting formula.

Claim 27. *With y, y' ranging over $\{\mathtt{u}, \mathtt{f}\}$, z ranging over $\{\mathtt{u}, \mathtt{s}, \mathtt{f}\}$ and ℓ ranging over $\mathcal{L}$ we have w.h.p.*

$$\frac{|\mathfrak{C}(\alpha)|}{(km)!} = \exp(-nH(\mathrm{Po}(d)) + o(n))\binom{n}{n(\bar{\Delta}_{z,\ell}(\alpha))_{z,\ell}}\binom{m}{m(\bar{\Gamma}_{z,\ell}(\alpha))_{z,\ell}} \cdot \binom{km}{\left(n\sum_{z,\ell}\ell_{yy'}\bar{\Delta}_{z,\ell}(\alpha)\right)_{y,y'}}^{-1} \prod_{z,\ell}\binom{\ell_{\mathtt{uu}}+\ell_{\mathtt{uf}}+\ell_{\mathtt{fu}}+\ell_{\mathtt{ff}}}{\ell_{\mathtt{uu}},\ell_{\mathtt{uf}},\ell_{\mathtt{fu}},\ell_{\mathtt{ff}}}^{n\bar{\Delta}_{z,\ell}(\alpha)+m\bar{\Gamma}_{z,\ell}(\alpha)}. \tag{4.6}$$

Proof. The first two multinomial coefficients account for the number of ways of assigning labels with the frequencies prescribed by (4.5) to the variables/checks. However, the first multinomial coefficient implicitly counts the assignment of the variable node degrees, on which we condition; this is because $\ell_{\mathtt{uu}}+\ell_{\mathtt{fu}}+\ell_{\mathtt{uf}}+\ell_{\mathtt{ff}}$ equals the degree of the corresponding variable. To correct for this overcounting, we divide by the multinomial coefficient

$$\binom{n}{(|\{i \in [n] : d_{\boldsymbol{A}^\dagger}(v_i) = h\}|)_{h\geqslant 0}}. \tag{4.7}$$

But since the variable node degrees are asymptotically Poisson by Fact 12, (4.7) equals $\exp(nH(\mathrm{Po}(d)) + o(n))$ w.h.p. The first multinomial coefficient on the second line of (4.6) counts the number of possible matchings of the clones that respect **COV1**. The last factor accounts for the number of ways of assigning labels to the clones of the individual variable/check nodes. Finally, the $\exp(o(n))$ error term swallows the approximations in (4.3)–(4.4). □

Claim 28. *Letting*

$$\mathfrak{l}_1 = \mathbb{E}[\log(\mathrm{Po}(d)!)], \qquad \mathfrak{l}_2 = -\sum_{z,\ell}\bar{\Delta}_{z,\ell}(\alpha)\log(\ell_{\mathtt{uu}}!\ell_{\mathtt{uf}}!\ell_{\mathtt{fu}}!\ell_{\mathtt{ff}}!),$$

$$\mathfrak{l}_3 = -\frac{d}{k}\sum_{z,\ell}\bar{\Gamma}_{z,\ell}(\alpha)\log(\ell_{\mathtt{uu}}!\ell_{\mathtt{uf}}!\ell_{\mathtt{fu}}!\ell_{\mathtt{ff}}!),$$

$$\mathfrak{h}_1 = H(\bar{\delta}(\alpha,z))_z + H(\mathrm{Po}(d(1-\alpha^{k-1}))) + \bar{\delta}(\alpha,f)H(\mathrm{Po}_{\geqslant 2}(d\alpha^{k-1})),$$

$$\mathfrak{h}_2 = \frac{d}{k}\left[H(\bar{\gamma}(\alpha,z))_z + \bar{\gamma}(\alpha,\mathtt{u})H(\mathrm{Bin}_{\geqslant 2}(k,1-\alpha))\right], \quad \mathfrak{h}_3 = d\left[H(\mathrm{Be}(\alpha^{k-1})) - H(\mathrm{Be}(\alpha))\right],$$

$$\mathfrak{h}_4 = -H(\mathrm{Po}(d))$$

w.h.p. we have $\frac{1}{n}\log\frac{|\mathfrak{C}(\alpha)|}{(k!)^m(km)!\prod_{i=1}^n d_{\boldsymbol{A}^\dagger}(v_i)!} = \mathfrak{l}_1+\mathfrak{l}_2+\mathfrak{l}_3+\mathfrak{h}_1+\mathfrak{h}_2+\mathfrak{h}_3+\mathfrak{h}_4+o(1)$.

Proof. In combination with (2.14)–(2.21), Stirling's formula shows that

$$\begin{aligned}&\frac{1}{n}\log\binom{n}{n(\bar\Delta_{z,\ell}(\alpha))_{z,\ell}} =\\ &H(\bar\delta(\alpha,z))_z + H(\mathrm{Po}(d(1-\alpha^{k-1}))) + \bar\delta(\alpha,f)H(\mathrm{Po}_{\geqslant 2}(d\alpha^{k-1})) + o(1),\end{aligned} \tag{4.8}$$

$$\frac{1}{n}\log\binom{m}{m(\bar\Gamma_{z,\ell}(\alpha))_{z,\ell}} = \frac{d}{k}\left(H(\bar\gamma(\alpha,z))_z + \bar\gamma(\alpha,\mathtt{u})H(\mathrm{Bin}_{\geqslant 2}(k,1-\alpha))\right)+o(1). \tag{4.9}$$

Similarly,

$$\frac{1}{n}\log\left[\frac{1}{(km)!}\prod_{y,y'}\left(n\sum_{z,\ell}\ell_{y,y'}\bar\Delta_{z,\ell}(\alpha)\right)!\right] = -d\left[H(\mathrm{Be}(\alpha)) - H(\mathrm{Be}(\alpha^{k-1}))\right]+o(1). \tag{4.10}$$

Further,

$$\sum_{z,\ell}\bar\Delta_{z,\ell}(\alpha)\log\frac{(\ell_{\mathtt{uu}}+\ell_{\mathtt{uf}}+\ell_{\mathtt{fu}}+\ell_{\mathtt{ff}})!}{\ell_{\mathtt{uu}}!\ell_{\mathtt{uf}}!\ell_{\mathtt{fu}}!\ell_{\mathtt{ff}}!} = \mathfrak{l}_1+\mathfrak{l}_2+o(1). \tag{4.11}$$

Finally, since $\ell_{\mathtt{uu}}+\ell_{\mathtt{uf}}+\ell_{\mathtt{fu}}+\ell_{\mathtt{ff}} = k$ for all ℓ such that $\bar\Gamma_{z,\ell}(\alpha)>0$, we have

$$-\frac{1}{n}\log((k!)^m) + \frac{m}{n}\sum_{z,\ell}\bar\Gamma_{z,\ell}(\alpha)\log\frac{(\ell_{\mathtt{uu}}+\ell_{\mathtt{uf}}+\ell_{\mathtt{fu}}+\ell_{\mathtt{ff}})!}{\ell_{\mathtt{uu}}!\ell_{\mathtt{uf}}!\ell_{\mathtt{fu}}!\ell_{\mathtt{ff}}!} = \mathfrak{l}_3+o(1). \tag{4.12}$$

Combining (4.8)–(4.12) with Claim 27 completes the proof. □

Proof of Lemma 26. Let $\lambda = \alpha^{k-1}d$ and $\mu = d-\lambda$. Since by Fact 12 the empirical distribution of the degrees $(d_{\boldsymbol{A}^\dagger}(v_i))_{i\in[n]}$ is approximately $\mathrm{Po}(d)$ and in light of (2.14)–(2.21), w.h.p. we have

$$\mathfrak{l}_1 = \frac{1}{n}\sum_{i=1}^n\log(d_{\boldsymbol{A}^\dagger}(v_i)!)+o(1), \quad \mathfrak{l}_2 = -\mathbb{E}\left[\log\left(\mathrm{Po}(\mu)!\right)\right] - \bar\delta(\alpha,\mathtt{f})\mathbb{E}\left[\log(\mathrm{Po}_{\geqslant 2}(\lambda)!)\right]+o(1), \tag{4.13}$$

$$\mathfrak{l}_3 = \frac{d}{k}\bar\gamma(\alpha,\mathtt{s})\log(k) + \frac{d}{k}\bar\gamma(\alpha,\mathtt{u})\mathbb{E}\left[\log\binom{k}{\mathrm{Bin}_{\geqslant 2}(k,1-\alpha)}\right]+o(1). \tag{4.14}$$

Furthermore, trite rearrangements reveal that

$$\mathfrak{h}_1 = d(1-H(\mathrm{Be}(\alpha^{k-1}))-\log d) + \mathbb{E}\left[\log(\mathrm{Po}(\mu)!)\right] + \bar\delta(\alpha,\mathtt{f})\mathbb{E}\left[\log(\mathrm{Po}_{\geqslant 2}(\lambda)!)\right], \tag{4.15}$$

$$\mathfrak{h}_2 = dH(\mathrm{Be}(\alpha)) - \frac{d}{k}\bar\gamma(\alpha,\mathtt{s})\log(k) - \frac{d}{k}\bar\gamma(\alpha,\mathtt{u})\mathbb{E}\left[\log\binom{k}{\mathrm{Bin}_{\geqslant 2}(k,1-\alpha)}\right], \tag{4.16}$$

$$\mathfrak{h}_4 = -d(1-\log d) - \mathbb{E}\left[\log(\mathrm{Po}(d)!)\right]. \tag{4.17}$$

The assertion follows from (4.13)–(4.17) and Claim 28. □

4.2 Proof of Proposition 7

Lemma 26 estimates of the number of α-WP fixed points. In order to prove Proposition 7 we now need to count the number of 'balanced' assignments of values to the unfrozen variables of a WP fixed point such that all checks are satisfied. Thus, let $(\pi, \mathfrak{p})$ be an α-cover. Call $\sigma \in \mathbb{F}_q^n$ *compatible* with $(\pi, \mathfrak{p})$ if

$$\sigma_j = 0 \text{ for all } j \in [n] \text{ with } \mathfrak{p}(v_j) \neq \mathtt{u}, \text{ and} \tag{4.18}$$

$$\sum_{\ell \geqslant 0} \sum_{s \in \mathbb{F}_q \setminus \{0\}} (\ell+1) \left| \sum_{j=1}^{n} \mathbf{1}\{d_{\boldsymbol{A}^\dagger}(v_j) = \ell,\, \mathfrak{p}(v_j) = \mathtt{u}\} \left(\mathbf{1}\{\sigma_j = s\} - q^{-1}\right) \right| = o(n). \tag{4.19}$$

Thus, we ask that the values of the variables v_j with $\mathfrak{p}(v_j) = \mathtt{u}$ be about uniformly distributed on $\mathbb{F}_q$, even when broken down to individual variable degrees. Further, a pairing $\pi \in \mathfrak{P}$ induces a matrix $A(\pi)$ by letting

$$A_{ij}(\pi) = \mathfrak{A}_{ij} \cdot \mathbf{1}\{\exists l \in [d_{\boldsymbol{A}^\dagger}(v_j),\, h \in [k] : \pi(v_j, l) = (a_i, h)\} \qquad (i \in [m],\, j \in [n]).$$

Finally, we say that $\sigma \in \mathbb{F}_q^n$ *essentially satisfies* $(\pi, \mathfrak{p})$ if $\|A(\pi)\sigma\|_0 = o(n)$. Recall that $\boldsymbol{\pi} \in \mathfrak{P}$ denotes a random pairing.

Lemma 29. *Let $\mathfrak{p} : \mathfrak{V} \cup \mathfrak{F} \to \{\mathtt{u}, \mathtt{f}\}$ and let $\sigma \in \mathbb{F}_q^n$. Let $\mathfrak{C}$ be the event that $(\boldsymbol{\pi}, \mathfrak{p})$ is an α-cover that σ is compatible with, and let $\mathfrak{E}$ be the event that σ is essentially satisfying. Then $\mathbb{P}[\mathfrak{E} \mid \mathfrak{C}, \mathfrak{D}] \leqslant q^{-m\bar{\gamma}(\alpha,\mathtt{u})+o(n)}$ w.h.p.*

Proof. Given $\mathfrak{C}, \mathfrak{D}$ let $\mathfrak{I}$ be the set of all pairs $(i,h) \in [m] \times [k]$ such that $\boldsymbol{\pi}(i,h) \in \{v_j\} \times \mathbb{N}$ for some variable v_j with $\mathfrak{p}(v_j) = \mathtt{u}$. Thus, $\mathfrak{I}$ contains the check clones 'hit' by an unfrozen variable. Further, let $\mathcal{I}$ contain all $i \in [m]$ such that $\{i\} \times [k] \cap \mathfrak{I} \neq \emptyset$. What remains random given $\mathfrak{C}, \mathfrak{D}, \mathfrak{I}$ is *which* unfrozen variable clones are matched to $\mathfrak{I}$. Our goal is to estimate the probability that all checks a_i, $i \in \mathcal{I}$, end up satisfied under this random matching. Let $\boldsymbol{\xi} = (\boldsymbol{\xi}_{ih})_{(i,h)\in\mathfrak{I}}$ be the vector that comprises the values under σ of the variables that the clones in $\mathfrak{I}$ get matched to. In symbols, $\boldsymbol{\xi}_{ih} = \sum_{j\in[n]} \sigma_j \mathbf{1}\{\boldsymbol{\pi}(a_i, h) \in \{v_j\} \times \mathbb{N}\}$.

To investigate $\boldsymbol{\xi}$ we introduce an auxiliary random vector $\boldsymbol{\chi} = (\boldsymbol{\chi}_{ih})_{(i,h)\in\mathfrak{I}}$ with independent uniformly distributed entries $\boldsymbol{\chi}_{ih} \in \mathbb{F}_q$. Consider the events

$$\mathfrak{R} = \left\{ \forall s \in \mathbb{F}_q \setminus \{0\} : \sum_{(i,h)\in\mathfrak{I}} \mathbf{1}\{\boldsymbol{\chi}_{ih} = s\} = \sum_{j=1}^{n} \mathbf{1}\{\sigma_j = s\} d_{\boldsymbol{A}^\dagger}(v_j) \right\},$$

$$\mathfrak{X} = \left\{ \sum_{i\in\mathcal{I}} \mathbf{1}\left\{ \sum_{h:(i,h)\in\mathfrak{I}} \boldsymbol{\chi}_{ih} \neq 0 \right\} = o(n) \right\}.$$

Given the event $\mathfrak{R}$ the vectors $\boldsymbol{\xi}$ and $\boldsymbol{\chi}$ are identically distributed. Hence,

$$\mathbb{P}[\mathfrak{E} \mid \mathfrak{C}, \mathfrak{D}] = \mathbb{P}[\mathfrak{X} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}, \mathfrak{R}]. \tag{4.20}$$

The unconditional probabilities $\mathbb{P}[\mathfrak{X} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}]$ and $\mathbb{P}[\mathfrak{R} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}]$ are computed easily. Indeed, because the $\boldsymbol{\chi}_{ih}$ are uniform and independent, for any $i \in \mathcal{I}$ the event $\sum_{h:(i,h)\in\mathfrak{I}} \boldsymbol{\chi}_{ih} = 0$ occurs with probability $1/q$. Hence,

$$\mathbb{P}[\mathfrak{S} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}] = q^{-|\mathcal{I}|+o(n)}. \tag{4.21}$$

Furthermore, conditions **COV1**–**COV4** and the definitions (2.19)–(2.21) of the coefficients $\Gamma_{z,\ell}(\alpha)$ ensure that w.h.p. given $\mathfrak{C}, \mathfrak{D}$ we have $|\mathfrak{I}| = m(\bar{\gamma}(\alpha, \mathfrak{u}) + o(1))$. Thus, (4.21) becomes

$$\mathbb{P}[\mathfrak{S} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}] = q^{-m\bar{\gamma}(\alpha,\mathfrak{u})+o(n)}. \tag{4.22}$$

Moreover, (4.19) ensures that $\mathbb{P}[\mathfrak{R} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}] = \exp(o(n))$. Combining (4.20) and (4.22) with Bayes' rule, we obtain

$$\mathbb{P}[\mathfrak{E} \mid \mathfrak{C}, \mathfrak{D}] = \mathbb{E}[\mathbb{P}[\mathfrak{S} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}, \mathfrak{R}] \mid \mathfrak{C}, \mathfrak{D}] \leqslant \mathbb{E}\left[\frac{\mathbb{P}[\mathfrak{S} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}]}{\mathbb{P}[\mathfrak{R} \mid \mathfrak{C}, \mathfrak{D}, \mathfrak{I}]} \mid \mathfrak{C}, \mathfrak{D}\right] \leqslant q^{-m\bar{\gamma}(\alpha,\mathfrak{u})+o(n)},$$

as desired. □

Proof of Proposition 7. As a first step we relate the number of α-WP fixed points of $\boldsymbol{A}^\dagger$ to the number of α-covers. Given $\mathfrak{D}, \mathfrak{S}$ the random matrix $A(\boldsymbol{\pi})$ has the same distribution as $\boldsymbol{A}^\dagger$. Hence, suppose that $\mathfrak{m}$ is an α-WP fixed point of $A(\boldsymbol{\pi})$. Then $\mathfrak{m}$ induces a map $\mathfrak{p}_{\boldsymbol{\pi}} : \mathfrak{V} \cup \mathfrak{F} \to \{\mathfrak{f}, \mathfrak{u}\}^2$ by letting $\mathfrak{p}_{\boldsymbol{\pi}}(a_i, h) = (\mathfrak{m}_{v_j \to a_i}, \mathfrak{m}_{a_i \to v_j})$, where $j \in [n]$ is the unique index such that $\boldsymbol{\pi}(a_i, h) \in \{v_j\} \times \mathbb{N}$. Similarly, $\mathfrak{p}_{\boldsymbol{\pi}}(v_j, h) = (\mathfrak{m}_{a_i \to v_j}, \mathfrak{m}_{v_j \to a_i})$ if $\boldsymbol{\pi}(v_j, h) \in \{a_i\} \times [k]$. The definitions (2.9)–(2.10) and (2.22)–(2.23) ensure that $(\boldsymbol{\pi}, \mathfrak{p}_{\boldsymbol{\pi}})$ satisfies **COV1**–**COV4**. Thus, $(\boldsymbol{\pi}, \mathfrak{p}_{\boldsymbol{\pi}})$ is an α-cover.

Before we proceed we need to deal with an overcounting issue. Specifically, given $\mathfrak{D}$ for any matrix $\boldsymbol{A}^\dagger$ there are $\Xi = (k!)^m \prod_{j=1}^n d_{\boldsymbol{A}^\dagger}(v_j)!$ pairings π that render $\boldsymbol{A}^\dagger$, i.e., that satisfy $A(\pi) = \boldsymbol{A}^\dagger$. At the same time, there are a total of $(km)!$ pairings π, and $\boldsymbol{A}^\dagger$ and $A(\boldsymbol{\pi})$ are identically distributed given $\mathfrak{S}$. In effect, Lemma 26, which counts the total number of α-covers, implies that the number W_α of α-WP fixed points of $\boldsymbol{A}^\dagger$ satisfies

$$\mathbb{E}[W_\alpha \mid \mathfrak{D}] = \exp(o(n)) \qquad \text{w.h.p.} \tag{4.23}$$

Now consider an extension σ of $\mathfrak{m}$. Then σ is *nearly* compatible with $(\boldsymbol{\pi}, \mathfrak{p}_{\boldsymbol{\pi}})$, except that (4.18) may be violated for $o(n)$ indices $j \in [n]$. To remedy this set $\tau_j = \mathbf{1}\{\mathfrak{p}(v_j) = \mathfrak{u}\}\sigma_j$. Then τ is compatible with $(\boldsymbol{\pi}, \mathfrak{p})$ and (2.24) implies

$$\sum_{j=1}^n \mathbf{1}\{\sigma_j \neq \tau_j\} = o(n). \tag{4.24}$$

Further, because $\sigma \in \ker \boldsymbol{A}^\dagger$, Fact 12 and (4.24) yield $\|A(\boldsymbol{\pi})\tau\|_0 = o(n)$. Hence, τ essentially satisfies $(\boldsymbol{\pi}, \mathfrak{p}_{\boldsymbol{\pi}})$.

Since (4.24) shows that the number of inverse images $(\mathfrak{m}, \sigma)$ that can give rise to a specific pair $(\mathfrak{p}_{\boldsymbol{\pi}}, \tau)$ is bounded by $\exp(o(n))$, in order to bound $\boldsymbol{X}_\alpha$ it suffices to bound the

expected number of pairs $(\mathfrak{p}_{\boldsymbol{\pi}}, \tau)$ given $\mathfrak{D}$. The estimate (4.23) shows that the expected number of α-covers $\mathfrak{p}_{\boldsymbol{\pi}}$ induced by α-WP fixed points is bounded by $\exp(o(n))$. Furthermore, given $\mathfrak{p}_{\boldsymbol{\pi}}$ the number of assignments τ that satisfy the condition (4.19) is bounded by $q^{\bar{\delta}(\alpha,\mathtt{u})n+o(n)}$. Moreover, Lemma 29 shows that such a τ is essentially satisfying with probability $q^{\bar{\gamma}(\alpha,\mathtt{u})m+o(n)}$. Combining these estimates and recalling the definitions (1.1), (2.14) and (2.15) of Φ, $\bar{\delta}(\alpha,\mathtt{u})$ and $\bar{\gamma}(\alpha,\mathtt{u})$, we obtain

$$\mathbb{E}\left[\boldsymbol{X}_\alpha \mid \mathfrak{D}\right] \leqslant q^{\bar{\delta}(\alpha,\mathtt{u})n+\bar{\gamma}(\alpha,\mathtt{u})m+o(n)} = q^{\Phi_{d,k}(\alpha)n+o(n)} \qquad \text{w.h.p.},$$

thereby completing the proof. □

4.3 Proof of Corollary 8

Fact 10 shows that for $d < d_k$ the function $\Phi_{d,k}(\alpha)$ attains its unique global maximum at $\alpha = 0$. Moreover, a glimpse at (1.1) reveals that $\Phi_{d,k}(0) = 1 - d/k$. Hence, for any $d < d_k$ there exists $\zeta > 0$ such that for any fixed $\xi > 0$ we have $n \max_{\alpha\in[\xi,1]} \Phi(\alpha) < n - m - 3\zeta n$. Hence, Propositions 5 and 7 show together with Markov's inequality that

$$\mathbb{P}\left[\operatorname{nul} \boldsymbol{A}^\dagger \geqslant n - m - 2\zeta n \mid \boldsymbol{\alpha} \in [\xi, 1]\right] = \mathbb{P}\left[\max_{\alpha\in[\xi,1]} \boldsymbol{X}_\alpha \geqslant q^{n-m-\zeta n}\right] + o(1) = o(1). \quad (4.25)$$

But since $\boldsymbol{A}^\dagger$ has $m + o(n)$ rows, we have $\operatorname{nul} \boldsymbol{A}^\dagger \geqslant n - m + o(n)$. Therefore, (4.25) shows that $\boldsymbol{\alpha} < \xi$ w.h.p. Letting $\xi \to 0$ sufficiently slowly as $n \to \infty$, we thus conclude that $\boldsymbol{\alpha} = o(1)$ w.h.p. Therefore, the assertion follows from Proposition 5.

4.4 Proof of Lemma 9

Recall that for $\sigma \in \mathbb{F}_q^n$ we let $\rho(\sigma) = (\rho_s(\sigma))_{s\in\mathbb{F}_q}$ with $\rho_s(\sigma) = \frac{1}{n}\sum_{j=1}^n \mathbf{1}\{\sigma_j = s\}$. Let $\mathcal{R} = \{\rho(\sigma) : \sigma \in \mathbb{F}_q^n\}$ be the set of all conceivable $\rho(\sigma)$-vectors. Further, for $\chi = (\chi_1, \ldots, \chi_n) \in \mathbb{F}_q^n$ let $\chi^\perp = \{\sigma \in \mathbb{F}_q^n : \sum_{j=1}^n \sigma_j\chi_j = 0\}$. The following claim yields the approximate probability that a random vector whose entries are drawn independently from a distribution $r \in \mathcal{R}$ close to the uniform distribution $q^{-1}\mathbf{1}$ belongs to $\chi^\perp$.

Claim 30. *Let $\chi \in \mathbb{F}_q^n$ be a vector with $|\mathrm{supp}\chi| = k \geqslant 3$. Then uniformly for $r \in \mathcal{R}$ with $\|r - q^{-1}\mathbf{1}\| < \varepsilon$ we have*

$$\varphi_\chi(r) = \sum_{\sigma\in\chi^\perp} \prod_{s\in\mathbb{F}_q} r_s^{n\rho_s(\sigma)} = \frac{1}{q} + O(\varepsilon^3) \qquad \text{as } \varepsilon \to 0.$$

Proof. Let $\mathcal{X}(\chi) = \{\sigma \in \mathbb{F}_q^{\mathrm{supp}\chi} : \sum_{j\in\mathrm{supp}\chi} \sigma_j\chi_j = 0\}$ and for $\sigma \in \mathcal{X}(\chi)$ and $s \in \mathbb{F}_q$ let $R_s(\sigma) = |\{j \in \mathrm{supp}\chi : \sigma_j = s\}|$. Then $\varphi_\chi(r) = f_\chi(r)$, where

$$f_\chi(r) = \sum_{\sigma\in\mathcal{X}(\chi)} \prod_{s\in\mathbb{F}_q} r_s^{R_s(\sigma)}.$$

We are going to expand $f_\chi(r)$ to the second order. Clearly, $f_\chi(q^{-1}\mathbf{1}) = q^{-1}$, because $\mathcal{X}(\chi) \subseteq \mathbb{F}_q^{\mathrm{supp}\chi}$ is a linear subspace of codimension one and thus $|\mathcal{X}(\chi)| = \mathbb{F}_q^{k-1}$. Further, the partial derivatives of $f_\chi(r)$ come out as

$$\frac{\partial f_\chi}{\partial r_t} = \sum_{\sigma\in\mathcal{X}(\chi)} R_t(\sigma) r_t^{R_t(\sigma)-1} \prod_{s\in\mathbb{F}_q\setminus\{t\}} r_s^{R_s(\sigma)} \qquad (t\in\mathbb{F}_q), \tag{4.26}$$

$$\frac{\partial^2 f_\chi}{\partial r_t \partial r_u} = \sum_{\sigma\in\mathcal{X}(\chi)} R_t(\sigma)R_u(\sigma) r_t^{R_t(\sigma)-1} r_u^{R_u(\sigma)-1} \prod_{s\in\mathbb{F}_q\setminus\{t,u\}} r_s^{R_s(\sigma)} \quad (t,u\in\mathbb{F}_q,\, t\neq u), \tag{4.27}$$

$$\frac{\partial^2 f_\chi}{\partial r_t^2} = \sum_{\sigma\in\mathcal{X}(\chi)} R_t(\sigma)(R_t(\sigma)-1) r_t^{R_t(\sigma)-2} \prod_{s\in\mathbb{F}_q\setminus\{t\}} r_s^{R_s(\sigma)} \qquad (t\in\mathbb{F}_q). \tag{4.28}$$

To evaluate (4.26) at $r = q^{-1}\mathbf{1}$, we observe that the affine subspace $\{\sigma\in\mathcal{X}(\chi) : \sigma_j = t\}$ has dimension $k-2$ for every $t\in\mathbb{F}_q$ and $j\in\mathrm{supp}\chi$, because $k = |\mathrm{supp}\chi| \geqslant 3$. Hence,

$$\left.\frac{\partial f_\chi}{\partial r_t}\right|_{r=q^{-1}\mathbf{1}} = q^{1-k} \sum_{j\in\mathrm{supp}\chi} \sum_{\sigma\in\mathcal{X}(\chi)} \mathbf{1}\{\sigma_j = t\} = \frac{k}{q}. \tag{4.29}$$

Similarly, since the affine subspaces $\{\sigma\in\mathcal{X}(\chi) : \sigma_j = t, \sigma_{j'} = u\}$ for $t,u\in\mathbb{F}_q$ and $j,j'\in\mathrm{supp}\chi$, $j\neq j'$, have dimension $k-3$, (4.27)–(4.28) evaluated at $r = q^{-1}\mathbf{1}$ boil down to

$$\left.\frac{\partial^2 f_\chi}{\partial r_t \partial r_u}\right|_{r=q^{-1}\mathbf{1}} = \left.\frac{\partial^2 f_\chi}{\partial^2 r_t}\right|_{r=q^{-1}\mathbf{1}} = \frac{k(k-1)}{q}. \tag{4.30}$$

Further, all third partial derivatives remain bounded, i.e.,

$$\frac{\partial^3 f_\chi}{\partial r_s \partial r_t \partial r_u} = O(1) \qquad \text{for all } s,t,u\in\mathbb{F}_q. \tag{4.31}$$

Finally, since for every $r\in\mathcal{R}$ we have $\sum_{s\in\mathbb{F}_q} r_s = 1$ and the only eigenspaces with non-zero eigenvalues of the Jacobi matrix $Df_\chi(q^{-1}\mathbf{1})$ and of the Hessian $D^2_\chi(q^{-1}\mathbf{1})$ are spanned by $\mathbf{1}$, the assertion follows from (4.29)–(4.31) and Taylor's formula. □

Proof of Lemma 9. Given the value of $\boldsymbol{t}$ the random matrix $\boldsymbol{A}^\dagger$ consists of m rows of support size k and $\boldsymbol{t}$ unary rows. These rows are stochastically independent. Therefore, Claim 30 shows that for any $r\in\mathcal{R}$ and any $\sigma\in\mathbb{F}_q^n$ with $\rho(\sigma) = r$ we have

$$\mathbb{P}\left[\sigma\in\ker\boldsymbol{A}^\dagger \mid \boldsymbol{t}\right] = q^{-m-\boldsymbol{t}}\exp(O(n\|r-q^{-1}\mathbf{1}\|_1^3)). \tag{4.32}$$

Further, we recall that the entropy function $H(r)$ has the expansion

$$H(r) = \log q - \frac{q}{2}\sum_{s\in\mathbb{F}_q}(r_s - q^{-1})^2 + O(\|r-q^{-1}\mathbf{1}\|_1^3). \tag{4.33}$$

Combining (4.32)–(4.33) and applying the Laplace method, we thus obtain for small enough $\varepsilon > 0$,

$$\begin{aligned}
&\mathbb{E}\left|\ker \boldsymbol{A}^\dagger \cap \{\sigma \in \mathbb{F}_q^n : \|\rho(\sigma) - q^{-1}\mathbf{1}\|_1 < \varepsilon\} \mid \boldsymbol{t}\right| \\
&= (1+o(1))q^{n-m-\boldsymbol{t}} \sum_{r \in \mathcal{R}: \|r - q^{-1}\mathbf{1}\|_1 < \varepsilon} \frac{\exp(-q\|r-q^{-1}\mathbf{1}\|_2^2/2 + O(\|r-q^{-1}\mathbf{1}\|_1^3))}{\sqrt{(2\pi n)^{q-1}\prod_{s\in\mathbb{F}_q} r_s}} \\
&\sim q^{n-m-\boldsymbol{t}},
\end{aligned}$$

as claimed. □

5 Proof of Theorem 2 (ii)

The proof of the second part of Theorem 2 is based on the interpolation method from mathematical physics [25]. The interpolation method has been applied previously in order to estimate the rank of random matrices from a more general model [7], and in fact the upper bound on the rank obtained in [7] implies Theorem 2 (ii). Nonetheless, for the sake of completeness here we present a simplified version of the interpolation argument tailored to the specific random matrix model $\boldsymbol{A}^\dagger$.

The basic idea is to construct a family $\boldsymbol{A}^\dagger(\theta)$ of matrices parametrised by $\theta \in [0,1]$. The first matrix $\boldsymbol{A}^\dagger(0)$ (essentially) coincides with the random matrix $\boldsymbol{A}^\dagger$, while at the other end $\boldsymbol{A}^\dagger(1)$ we have a matrix whose nullity is easy to compute explicitly. We will then differentiate $\mathbb{E}[\operatorname{nul} \boldsymbol{A}^\dagger(\theta)]$ to compare $\mathbb{E}[\operatorname{nul} \boldsymbol{A}^\dagger(0)]$ and $\mathbb{E}[\operatorname{nul} \boldsymbol{A}^\dagger(1)]$. Thus, we obtain a lower bound on the nullity of $\boldsymbol{A}^\dagger(0)$, and hence of $\boldsymbol{A}^\dagger$. Since $\operatorname{nul}(\boldsymbol{A}^\dagger) + \operatorname{rk}(\boldsymbol{A}^\dagger) = n$, this lower bound on the nullity translates into the desired upper bound on the rank of $\boldsymbol{A}^\dagger$.

The interpolating family $\boldsymbol{A}^\dagger(\theta)$ is constructed as follows. Let $\boldsymbol{m}_\theta, \boldsymbol{m}'_\theta$ be two independent Poisson variables with means $(1-\theta)dn/k$ and $d\theta\alpha_{\mathtt{f}}^{k-1}n$, respectively; here $\alpha_{\mathtt{f}} = \alpha_{\mathtt{f}}(d,k) > 0$ is the maximum fixed point of $\phi_{d,k}$ (see Fact 10). Both $\boldsymbol{m}_\theta, \boldsymbol{m}'_\theta$ are also independent of the uniform random variable $\boldsymbol{t} \in [T]$. The random matrix $\boldsymbol{A}^\dagger(\theta)$ has size $(\boldsymbol{m}_\theta + \boldsymbol{m}'_\theta + \boldsymbol{t}) \times n$. As in the definition (1.2) of $\boldsymbol{A}$, the first $\boldsymbol{m}_\theta$ rows of $\boldsymbol{A}^\dagger(\theta)$ have entries

$$\boldsymbol{A}^\dagger_{ij}(\theta) = \mathfrak{A}_{ij}\mathbf{1}\{j \in \boldsymbol{e}_i\} \qquad (i \in [\boldsymbol{m}_\theta], j \in [n]),$$

where $(\boldsymbol{e}_i)_{i\geqslant 1}$ is a family of uniformly random subsets of $[n]$ of size k; these sets are mutually independent as well as independent of $\boldsymbol{m}_t, \boldsymbol{m}'_t$ and $\boldsymbol{t}$. Further, for $\boldsymbol{m}_t < i \leqslant 1 + \boldsymbol{m}'_t + \boldsymbol{t}$ the i-th row of $\boldsymbol{A}^\dagger$ contains a single one in a uniformly random column $j \in [n]$, while all other entries are zero. The positions of these 1-entries are drawn independently of each other and of everything else.

Lemma 31. *We have* $\mathbb{E}[\operatorname{nul} \boldsymbol{A}^\dagger(0)] = \mathbb{E}[\operatorname{nul} \boldsymbol{A}] + o(n)$ *and* $\mathbb{E}[\operatorname{nul} \boldsymbol{A}^\dagger(1)] = n\exp(-d\alpha_{\mathtt{f}}^{k-1}) + o(n)$.

Proof. By construction the first $\boldsymbol{m}_0 \wedge m$ rows of $\boldsymbol{A}^\dagger(0)$ and $\boldsymbol{A}$ are identically distributed. Moreover, w.h.p. we have $\boldsymbol{m}_0 = m + o(n)$. Since adding or removing a single row can alter the nullity by at most one, the first assertion follows.

Regarding the second assertion, observe that the rows of $\boldsymbol{A}^\dagger(1)$ are all-zero, except for a single one entry that sits in an independent and uniformly random position. Hence, the nullity of $\boldsymbol{A}^\dagger(1)$ is simply the number of all-zero columns. Further, since $\mathbb{E}[\boldsymbol{m}'_1] = d\alpha_{\mathtt{f}}^{k-1}n$, the expected number of non-zero entries per column equals $d\alpha_{\mathtt{f}}^{k-1} + o(1)$. Since the $\boldsymbol{m}'_\theta$ is a Poisson variable, we expect $n\exp(-d\alpha_{\mathtt{f}}^{k-1} + o(1))$ all-zero columns. □

The main step of the interpolation method is to compute the derivative $\frac{\partial}{\partial\theta}\mathbb{E}[\operatorname{nul}\boldsymbol{A}(\theta)]$.

Lemma 32. *We have* $\frac{1}{n}\frac{\partial}{\partial\theta}\mathbb{E}[\operatorname{nul}\boldsymbol{A}(\theta)] \leqslant -d\alpha_{\mathtt{f}}^{k-1} + \frac{d}{k}(k-1)\alpha_{\mathtt{f}}^{k} + \frac{d}{k} + o(1)$.

Proof. Since $\boldsymbol{m}_\theta, \boldsymbol{m}'_\theta$ are Poisson variables, we calculate

$$\begin{aligned}\frac{1}{n}\frac{\partial}{\partial\theta}\mathbb{P}\left[\boldsymbol{m}_\theta = m\right] &= \frac{d}{k}\left[\mathbb{P}\left[\boldsymbol{m}_\theta = m\right] - \mathbb{P}\left[\boldsymbol{m}_\theta = m-1\right]\right],\\ \frac{1}{n}\frac{\partial}{\partial\theta}\mathbb{P}\left[\boldsymbol{m}'_\theta = m\right] &= d\alpha_{\mathtt{f}}^{k-1}\left[\mathbb{P}\left[\boldsymbol{m}'_\theta = m-1\right] - \mathbb{P}\left[\boldsymbol{m}'_\theta = m\right]\right].\end{aligned}$$

Therefore,

$$\begin{aligned}&\frac{1}{n}\frac{\partial}{\partial\theta}\mathbb{E}[\operatorname{nul}\boldsymbol{A}(\theta)] =\\ &\frac{1}{n}\sum_{m,m'\geqslant 0}\mathbb{E}\left[\operatorname{nul}\boldsymbol{A}^\dagger(\theta)\mid \boldsymbol{m}_\theta = m, \boldsymbol{m}'_\theta = m'\right]\frac{\partial}{\partial\theta}\mathbb{P}\left[\boldsymbol{m}_\theta = m\right]\mathbb{P}\left[\boldsymbol{m}'_\theta = m'\right]\\ &= d\alpha_{\mathtt{f}}^{k-1}\sum_{m'\geqslant 0}\left[\mathbb{E}\left[\operatorname{nul}\boldsymbol{A}^\dagger(\theta)\mid \boldsymbol{m}'_\theta = m'+1\right] - \mathbb{E}\left[\operatorname{nul}\boldsymbol{A}^\dagger(\theta)\mid \boldsymbol{m}'_\theta = m'\right]\right]\mathbb{P}\left[\boldsymbol{m}'_\theta = m'\right]\\ &\quad - \frac{d}{k}\sum_{m\geqslant 0}\left[\mathbb{E}\left[\operatorname{nul}\boldsymbol{A}^\dagger(\theta)\mid \boldsymbol{m}_\theta = m+1\right] - \mathbb{E}\left[\operatorname{nul}\boldsymbol{A}^\dagger(\theta)\mid \boldsymbol{m}_\theta = m\right]\right]\mathbb{P}\left[\boldsymbol{m}_\theta = m\right]. \quad (5.1)\end{aligned}$$

Hence, obtain $\boldsymbol{A}^\dagger_+(\theta)$ from $\boldsymbol{A}^\dagger(\theta)$ by adding one more row with precisely one non-zero entry in a uniformly random position, chosen independently of everything else. Let $\boldsymbol{a}^+$ signify this new row. Similarly, obtain $\boldsymbol{A}^\dagger_-(\theta)$ from $\boldsymbol{A}^\dagger(\theta)$ by adding the row $\boldsymbol{a}^-$ with entries

$$\boldsymbol{a}^-_j = \mathfrak{A}_{\boldsymbol{m}_\theta+1\, j}\mathbf{1}\{j \in \boldsymbol{e}_{m_\theta+1}\}.$$

Then (5.1) shows that

$$\begin{aligned}&\frac{1}{dkn}\frac{\partial}{\partial\theta}\mathbb{E}[\operatorname{nul}\boldsymbol{A}(\theta)] = \quad (5.2)\\ &k\alpha_{\mathtt{f}}^{k-1}\mathbb{E}\left[\operatorname{nul}(\boldsymbol{A}^\dagger_+(\theta)) - \operatorname{nul}(\boldsymbol{A}^\dagger(\theta))\right] - \mathbb{E}\left[\operatorname{nul}(\boldsymbol{A}^\dagger_-(\theta)) - \operatorname{nul}(\boldsymbol{A}^\dagger(\theta))\right].\end{aligned}$$

Let $\boldsymbol{\alpha}_\theta = |\mathcal{F}(\boldsymbol{A}^\dagger(\theta))|/n$. We claim that

$$\mathbb{E}\left[\operatorname{nul}(\boldsymbol{A}^\dagger_+(\theta)) - \operatorname{nul}(\boldsymbol{A}^\dagger(\theta))\right] = -\mathbb{E}\left[1 - \boldsymbol{\alpha}_\theta\right]. \quad (5.3)$$

Indeed, let $\boldsymbol{j}^+ \in [n]$ be the position of the non-zero entry of $\boldsymbol{a}^+$. Then adding $\boldsymbol{a}^+$ to $\boldsymbol{A}^\dagger(\theta)$ decreases the nullity iff $j^+ \notin \mathcal{F}(\boldsymbol{A}^\dagger(\theta))$. Since $\boldsymbol{j}^+$ is uniformly random and independent of $\boldsymbol{A}^\dagger(\theta)$, we obtain (5.3).

Further, we claim

$$\mathbb{E}\left[\operatorname{nul}(\boldsymbol{A}^\dagger_-(\theta)) - \operatorname{nul}(\boldsymbol{A}^\dagger(\theta))\right] = -\mathbb{E}\left[1 - \boldsymbol{\alpha}_\theta^k\right] + o(1). \tag{5.4}$$

To see this, let $\mathcal{E}_\theta$ be the event that $\boldsymbol{A}^\dagger(\theta)$ is $(o(1), k)$-free. Since the construction of $\boldsymbol{A}^\dagger(\theta)$ incorporates $\boldsymbol{t}$ random unary equations as in the pinning lemma (Lemma 3), we have $\mathbb{P}[\mathcal{E}_\theta] = 1 - o(1)$. Furthermore, since $\boldsymbol{a}^-$ is independent of $\boldsymbol{A}^\dagger(\theta)$, the probability that the positions $1 \leqslant \boldsymbol{j}_1^- < \cdots < \boldsymbol{j}_k^- \leqslant n$ of the non-zero entries of $\boldsymbol{a}^-$ form a proper relation of $\boldsymbol{A}^\dagger(\theta)$ is $o(1)$ on the event $\mathcal{E}_\theta$. Hence, assume that $\boldsymbol{j}_1^-, \dots, \boldsymbol{j}_k^-$ do not form a proper relation. Then the nullity drops upon addition of row $\boldsymbol{a}^-$ unless $\boldsymbol{j}_1^-, \dots, \boldsymbol{j}_k^- \in \mathcal{F}(\boldsymbol{A}^\dagger(\theta))$. Since $\mathbb{P}\left[\boldsymbol{j}_1^-, \dots, \boldsymbol{j}_k^- \in \mathcal{F}(\boldsymbol{A}^\dagger(\theta)) \mid \boldsymbol{A}^\dagger(\theta)\right] = \boldsymbol{\alpha}_\theta^k + o(1)$, we obtain (5.4).

Combining (5.2)–(5.4), we find

$$\frac{1}{dkn}\frac{\partial}{\partial\theta}\mathbb{E}[\operatorname{nul}\boldsymbol{A}(\theta)] = \mathbb{E}\left[1 - \boldsymbol{\alpha}_\theta^k - k\alpha_{\mathtt{f}}^{k-1}(1 - \boldsymbol{\alpha}_\theta)\right] + o(1). \tag{5.5}$$

To complete the proof, we notice that

$$1 - \boldsymbol{\alpha}_\theta^k - k\alpha_{\mathtt{f}}^{k-1}(1 - \boldsymbol{\alpha}_\theta) + \left(k\alpha_{\mathtt{f}}^{k-1} - (k-1)\alpha_{\mathtt{f}}^k - 1\right) = -\boldsymbol{\alpha}_\theta^k + k\boldsymbol{\alpha}_\theta\alpha_{\mathtt{f}}^{k-1} - (k-1)\alpha_{\mathtt{f}}^k \leqslant 0, \tag{5.6}$$

because $X^k - kXY^{k-1} + (k-1)Y^k \geqslant 0$ for all $X, Y \in [0,1]$ and all $k \geqslant 2$. The assertion follows from (5.5) and (5.6). □

Proof of Theorem 2 (ii). Suppose that $d > d_k$. Integrating on $\theta \in [0,1]$, we learn from Fact 10 and Lemma 31 that

$$\frac{1}{n}\mathbb{E}[\operatorname{nul}\boldsymbol{A}^\dagger] \geqslant \Phi_{d,k}(\alpha_{\mathtt{f}}) + o(1) > 1 - d/k. \tag{5.7}$$

Furthermore, Azuma–Hoeffding shows that $\operatorname{nul}\boldsymbol{A}^\dagger$ is tightly concentrated, because adding or removing a single row alters the nullity by at most one. Thus, since $\boldsymbol{A}^\dagger$ is obtained from $\boldsymbol{A}$ via the addition of $o(n)$ rows, we conclude that $n^{-1}\operatorname{nul}\boldsymbol{A} \geqslant \Phi_{d,k}(\alpha_{\mathtt{f}}) + o(1)$ w.h.p. Therefore, (5.7) shows that $\operatorname{rk}\boldsymbol{A} < m - \Omega(n)$ w.h.p. □

Acknowledgements

Amin Coja-Oghlan is supported by DFG CO 646/3 and DFG CO 646/5. Mihyun Kang is supported by a Friedrich Wilhelm Bessel research award of the Alexander von Humboldt Foundation (AUT 1204138 BES). This work is supported in part by the Austrian Science Fund (FWF) [10.55776/I6502]. For the purpose of open access, the second author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. An extended abstract version of this work appeared in Proc. EUROCOMB 2023.

References

[1] D. Achlioptas, M. Molloy: The solution space geometry of random linear equations. Random Structures and Algorithms **46** (2015) 197–231.

[2] D. Achlioptas, A. Naor, Y. Peres: Rigorous location of phase transitions in hard optimization problems. Nature **435** (2005) 759–764.

[3] M. Aizenman, R. Sims, S. Starr: An extended variational principle for the SK spin-glass model. Phys. Rev. B **68** (2003) 214403.

[4] P. Ayre, A. Coja-Oghlan, P. Gao, N. Müller: The satisfiability threshold for random linear equations. Combinatorica **40** (2020) 179–235.

[5] A. Coja-Oghlan, O. Cooley, M. Kang, J. Lee, J. Ravelomanana: The sparse parity matrix. Proc. 33rd SODA (2022) 822–833.

[6] A. Coja-Oghlan, P. Gao, M. Hahn-Klimroth, J. Lee, N. Müller, M. Rolvien: The full rank condition for sparse random matrices. Proc. 27th RANDOM (2023) #54.

[7] A. Coja-Oghlan, A. Ergür, P. Gao, S. Hetterich, M. Rolvien: The rank of sparse random matrices. Random Structures and Algorithms **62** (2023) 68–130.

[8] O. Cooley, J. Lee, J. Ravelomanana: Warning Propagation: stability and subcriticality. arXiv:2111.15577 (2021).

[9] C. Cooper: The cores of random hypergraphs with a given degree sequence. Random Structures and Algorithms **25** (2004) 353–375.

[10] C. Cooper, A. Frieze, W. Pegden: On the rank of a random binary matrix. Electron. J. Comb. **26** (2019) P4.12.

[11] N. Creignou, H. Daude, O. Dubois: Approximating the satisfiability threshold for random k-XOR-formulas. arXiv:cs/0106001 (2001).

[12] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, M. Rink: Tight thresholds for cuckoo hashing via XORSAT. Proc. 37th ICALP (2010) 213–225.

[13] J. Ding, A. Sly, N. Sun: Proof of the satisfiability conjecture for large k. Annals of Mathematics **196** (2022) 1–388.

[14] O. Dubois, J. Mandler: The 3-XORSAT threshold. Proc. 43rd FOCS (2002) 769–778.

[15] D. Fernholz, V. Ramachandran: Cores and connectivity in sparse random graphs. UTCS Technical Report TR04-13 (2004).

[16] A. Goerdt, L. Falke: Satisfiability thresholds beyond k-XORSAT. Proc. 7th International Computer Science Symposium in Russia (2012) 148–159.

[17] M. Ibrahimi, Y. Kanoria, M. Kraning, A. Montanari: The set of solutions of random XORSAT formulae. Annals of Applied Probability **25** (2015) 2743–2808.

[18] S. Janson, M. Luczak: A simple solution to the k-core problem. Random Structures and Algorithms **30** (2007) 50–62.

[19] S. Janson, T. Luczak, A. Rucinski: Random graphs. Wiley (2000).

[20] J.H. Kim: Poisson cloning model for random graphs. Proceedings of the International Congress of Mathematicians (2006) 873–897.

[21] M. Mézard, A. Montanari: Information, physics and computation. Oxford University Press (2009).

[22] M. Mézard, F. Ricci-Tersenghi, R. Zecchina: Two solutions to diluted p-spin models and XORSAT problems. Journal of Statistical Physics **111** (2003) 505–533.

[23] M. Molloy: Cores in random hypergraphs and Boolean formulas. Random Structures and Algorithms **27** (2005) 124–135.

[24] A. Montanari: Estimating random variables from random sparse observations. European Transactions on Telecommunications **19**(4) (2008) 385–403.

[25] D. Panchenko, M. Talagrand: Bounds for diluted mean-fields spin glass models. Probab. Theory Relat. Fields **130** (2004) 319–336.

[26] B. Pittel, G. Sorkin: The satisfiability threshold for k-XORSAT. Combinatorics, Probability and Computing **25** (2016) 236–268.

[27] B. Pittel, J. Spencer, N. Wormald: Sudden emergence of a giant k-core in a random graph. Journal of Combinatorial Theory, Series B **67** (1996) 111–151.

[28] P. Raghavendra, N. Tan: Approximating CSPs with global cardinality constraints using SDP hierarchies. Proc. 23rd SODA (2012) 373–387.

[29] O. Riordan: The k-core and branching processes. Combinatorics, Probability and Computing **17** (2008) 111–136.