The satisfiability threshold of random linear equations over finite commutative rings

by

Theodore Morrison

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2026

© Theodore Morrison 2026

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis is based on a paper co-authored with Jane Gao which has been submitted for publication [30]. The introductory sections up to Section 1.1 have largely been re-written for this thesis. Sections 1.2, 1.3, and 1.4 appear as in [30] with some additions. Chapters 4 and 5 appear verbatim in [30] except for minor additions and corrections. Chapters 6 and 7 have been reorganized to include the proofs of Lemmas 6.1.3 and 6.1.4 in Section 7.2, but otherwise appear as in [30].

Abstract

In this thesis we study systems of random linear equations as a random constraint satisfaction problem (CSP). Determining the satisfiability threshold is one of the most fundamental problems in the study of any random CSP: how many constraints can a typical CSP have before it can no longer be satisfied? This question has been answered in the setting of linear equations over a finite field [7]. We investigate natural generalizations of [7] in the setting of finite commutative rings.

Under very mild assumptions, the satisfiability threshold of a random CSP is either linear or sublinear in the number of variables. We provide a complete characterization of the linear and sublinear regimes for our model of random linear equations over a ring. In the sublinear regime, we show that the threshold is coarse and determine its exact value.

In the linear regime, we determine the exact threshold under the assumption that the underlying ring is a principal ideal ring. We show that the threshold over principal ideal rings is sharp, and coincides exactly with the satisfiability threshold over fields. This extends a striking phenomenon in [7], where it is observed that the satisfiability threshold over finite fields is the same for every choice of field.

We show by example that this phenomenon does not extend to non-principal ideal rings. We construct a non-principal ideal ring R, and determine the threshold for a family of distributions of linear equations over R. Not only does the satisfiability threshold differ from linear equations over a field, we find that the satisfiability threshold depends on parameters of the distribution in ways that are not observed for fields and principal ideal rings.

The final section of this thesis is devoted to contiguity results among models of random bipartite graphs. In particular, we show that the models of [7] and [14] are contiguous for some choice of parameters, and we correct a mistake made in [13]. These results are used in earlier sections of the thesis, and may be of independent interest.

Acknowledgements

I would like to thank my supervisor Jane Gao for the patience, generosity, and advice that she has given me during my time in Waterloo. This thesis would not have been possible without her guidance.

The final draft of thesis has benefited from the feedback of my thesis readers Sophie Spirkl and Ren Yeats. I thank them for their time, and for many helpful suggestions regarding clarity and exposition.

I would also like to thank the people that make department of Combinatorics and Optimization an exceptional place to be graduate student. I thank the department staff for making my life easier, the instructors for sharing their diverse expertise, and my classmates for partaking in stimulating discussions.

Finally, I would like to thank my friends in Waterloo and my family back in BC.

Table of Contents

Author's Declaration ii

Statement of Contributions iii

Abstract iv

Acknowledgements v

List of Figures viii

List of Symbols ix

1 Introduction 1

1.1 Motivation and related work . . . 2

1.2 Definitions and preliminaries . . . 4

1.2.1 Factor graphs . . . 6

1.3 Contributions . . . 7

1.3.1 Principal ideal rings . . . 7

1.3.2 Non-principal ideal rings . . . 9

1.4 Open problems . . . 13

2 Random Models and Contiguity 15

3 Overview of the Proofs 18

3.1 Theorems 1.3.2, 1.3.5, 1.3.7, and 1.3.6 . . . 18

3.2 Theorem 1.3.3 . . . 19

3.3 Theorem 1.3.9 . . . 20

3.4 Dependency diagrams . . . 20

4 Sublinear SAT Thresholds **23**

4.1 Proof of Theorem 1.3.5 23

4.2 Proofs of Theorems 1.3.2, 1.3.6, and 1.3.7 27

4.3 Proofs of Lemmas 4.2.2, 4.2.6, 4.2.5, and 4.2.4 33

5 Linear SAT threshold for principal rings: proof of Theorem 1.3.3 **41**

6 Linear SAT threshold for nonprincipal rings: proof of Theorem 1.3.9 **44**

6.1 Satisfiability 45

6.2 Unsatisfiability 47

7 Proofs of Theorems 2.0.3 and 2.0.4 and Lemmas 6.1.3 and 6.1.4 **50**

7.1 Proof of Lemma 7.0.1 51

7.2 Proof of Lemma 7.0.2 55

7.3 Factor graphs with a random number of constraint vertices 60

7.4 Proof of Lemmas 6.1.3 and 6.1.4 63

References **70**

List of Figures

1.1 $d_{3,\rho}$ as a function of $\rho \in [0,1]$ 13

3.1 Dependency diagram for Theorems 1.3.2, 1.3.5, 1.3.6, and 1.3.7 21
3.2 Dependency diagram for Theorem 1.3.3 21
3.3 Dependency diagram for Theorem 1.3.9 22

4.1 $T_0 = T_0^+$ 28
4.2 T_1^+ 28
4.3 T_1 28
4.4 A smallest obstruction for (1.1) in Example 4.2.7 33

List of Symbols

$\mathrm{Unif}(S)$	The uniform distribution on a set S
$\mathrm{Po}(\lambda)$	The Poisson distribution with mean λ
$\mathrm{Bin}(\ell, p)$	The binomial distribution with ℓ trials and probability of success p
$\mathbf{1}_{\{A\}}$	The indicator function of an event A
$\hat{P}$	A distribution on R^k
P	A distribution on $R \setminus \{0\}$
Q	A distribution on R
$\mathbf{B}$	The random matrix defined by (1.1)
C_m	A set of m constraint vertices
V_n	A set of n variable vertices
$F(B)$	The factor graph of a matrix $B \in R^{C_m \times V_n}$
$F(B, b)$	The weighted factor graph of a system of equations $Bx = b$ for $B \in R^{C_m \times V_n}$ and $b \in R^{C_m}$
$E(F)$	The set of edges of a factor graph F
$C(F)$	The set of constraint vertices of a factor graph F
$V(F)$	The set of edges variable vertices of a factor graph F
$e(F)$	The number of edges of a factor graph F
$c(F)$	The number of constraint vertices of a factor graph F
$v(F)$	The number of variable vertices of a factor graph F
$e_F(a, v)$	The number of edges between vertices a and v in a factor (multi)graph F
$\deg_F(u)$	The degree of a vertex u in a factor graph F
$N(u)$	The set of neighbours of a vertex u in a factor graph F
$E(F)$	The set of edges of a factor graph F
$(B)_{\text{2-core}}$	The 2-core of a random matrix B, defined in Sections 1.2
d_k	The constant defined by (1.2)
$d_{k,\rho}$	The constant defined by (1.6)
$\{I_\ell(P)\}_{\ell \in \mathbb{N}}$	The sequence of ideals defined in Definition 1.3.1
$\{I_\ell^+(P)\}_{\ell \in \mathbb{N}}$	The sequence of ideals defined in Definition 4.2.3
$\ell(P, Q)$	The integer defined in Definition 1.3.1

$\ell^+(P,Q)$	The integer defined in Definition 4.2.3
$\alpha_{P,Q}$	The exponent of the coarse satisfiability threshold in Theorems 1.3.2 and 1.3.5
$\mathcal{T}(P,Q)$	The set of unsatisfiable weighted acyclic factor graphs defined in Section 1.3.2
T_ℓ^+, T_ℓ	Connected acyclic factor graphs defined in Definition 4.2
$G_{n,\boldsymbol{m}',\boldsymbol{k}}$ $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ $G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$	Models of simle random factor graphs defined in Chapter 2
$M_{n,\boldsymbol{m}',\boldsymbol{k}}$ $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$	Models of random factor graphs defined in Chapter 2 that allow multiple edges
$\mathcal{D}_n$	The event $\sum_{i=1}^n \boldsymbol{d}'_{v_i} = \sum_{i=1}^{\boldsymbol{m}'} \boldsymbol{k}'_{a_i}$ defined in (2.1)
$\mathbf{B}_{XY}$	A random matrix in $\mathbb{F}_q^{m\times 2n}$ defined in Chapter 6
$\mathbf{C}$, $\mathbf{C}'$, $\mathbf{C}''$	Coupled random matrices defined in Section 6.1
$\mathcal{S}_n$	The set of simple factor graphs on n variable vertices
$\triangleright$	The contiguity relation between random graph models defined in Definition 2.0.2
$\lhd\rhd$	The mutual contiguity relation between random graph models defined in Definition 2.0.2

Chapter 1

Introduction

Constraint satisfaction problems (CSPs) provide a unified framework for studying a wide range of problems in computer science, combinatorics, and statistical physics. A general CSP instance consists of a set of n variables and a set of m constraints; the variables are assigned values from a fixed set, and the goal is to find an assignment of the variables that satisfies every constraint. The study of CSPs often focuses on one fundamental question: how many constraints, in terms of the number of variables, can a CSP have before it becomes unsatisfiable? In deterministic settings, answers to this question come in the form of extremal results. For example, r-colouring a graph G can be viewed as a CSP, where each vertex of G is a variable, and colourings of the vertices (ie. assignments of the variables) are subject to the constraint that no edge of G is monochromatic. The maximum number of constraints in a satisfiable instance of this CSP is the maximum number of edges in an r-colourable graph on n vertices, which is given by the Turán graph. In this thesis we follow a large body of work that attempts to answer a similar question in the setting of random CSPs: given n, what is the maximum number of constraints m such that a randomly generated CSP with m constraints and n variables is likely to be satisfiable?

Answers to this question come in the form of a *satisfiability threshold function*, which marks a phase transition between asymptotically satisfiable CSPs and asymptoticaly unsatisfiable CSPs. More precisely, a satisfiability threshold function of a random CSP is a function $t(n)$ such that[1]

$$\lim_{n\to\infty} \mathbb{P}[\text{A random } n\text{-variable, } m\text{-constraint CSP is satisfiable}] = \begin{cases} 1 & m = o(t(n)) \\ 0 & m = \omega(t(n)). \end{cases}$$

Satisfiability thresholds can be further distinguished between sharp and coarse thresholds.

[1] In this thesis, the Landau notation $o(\cdot), \omega(\cdot), O(\cdot), \Omega(\cdot)$ is always taken with respect to the number of variables n. We will sometimes use the notation $f(n) \gg g(n)$ and $f(n) \ll g(n)$ instead of $f(n) = \omega(g(n))$ and $f(n) = o(g(n))$.

A threshold $t(n)$ is sharp if, for all fixed $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}[\text{A random } n\text{-variable, } m\text{-constraint CSP is satisfiable}] = \begin{cases} 1 & m < (1-\varepsilon)t(n) \\ 0 & m > (1+\varepsilon)t(n). \end{cases}$$

Any threshold that is not sharp is called a coarse threshold. Determining whether the satisfiability threshold is sharp or coarse is typically the first step in studying the satisfiability threshold of a random CSP. The exact asymptotics of $t(n)$ can be determined only if the threshold is sharp.

The particular problem we consider is that of solving systems of linear equations over finite commutative rings. We primarily study systems of equations as a CSP, and attempt to determine the satisfiability threshold for a naturally defined random model. Our setting is also relevant to the vast literature on random matrices. The singularity of random matrices is particularly relevant, as well as questions about rank and nullity [14, 48, 20, 21]. In the particular setting of random matrices over a ring, questions about the distribution of the kernel and co-kernel may also be of interest [44, 36]. While we do not attempt to answer questions about random matrices beyond determining the satisfiability threshold, we pose some open problems in this area in Section 1.4.

1.1 Motivation and related work

As noted above, random CSPs are used to study problems in a wide range of fields. Statistical physicists have been interested in random CSPs since at least the 1990s, when it was observed that random CSPs are closely related to the spin glass models used by physicists to study disordered particle systems [41]. An extensive sequence of work by statistical physicists has shown how heuristic arguments such as the cavity method can be adapted to give precise predictions about the behavior of random CSPs [38, 41, 43, 46]. In particular, [43] predicted the existence multiple structural transitions in the solution space of random CSPs, which are marked by thresholds similar to the satisfiability threshold. The smallest of these is the "clustering" threshold. When the number of constraints m is below the clustering threshold, the entire space of solutions can be traversed by iteratively changing only a small number of variables at a time. At the clustering threshold, the solution space breaks into exponentially many "clusters" which are separated from each other in terms of their Hamming distance. A third "condensation" threshold lies between the clustering threshold and the satisfiability threshold, and marks the condensation of solutions into a few large clusters. The heuristic predictions of physicists have been steadily vindicated for many CSP models, including the existence and location of clustering [37, 2, 42], condensation [15, 8], and satisfiability [26, 25, 1] thresholds.

The k-SAT problem, where each constraint is a disjunctive boolean clause containing exactly k literals, is the most classical problem in the CSP literature. The origins of k-SAT

date back to the beginning of modern complexity theory, where Cook and Levin independently proved that the k-SAT problem is NP-complete [19, 34]. Rigorous results on the random k-SAT satisfiability threshold began to appear in the 1990s, motivated by a large body of experimental work [12, 47, 22, 39] and later by statistical physics predictions. The first major breakthrough came when Friedgut and Bourgain independently proved the existence of a sharp (but nonuniform) satisfiability threshold [28]. Around the same time, an asymptotically correct upper bound on the threshold for large k was given using a first moment argument [32]. The asymptotic lower bound proved much more difficult, and required a sequence of refinements [4, 6, 16] culminating in [17] with a lower bound that agrees with the previously established upper bound and statistical physics predictions. The exact k-SAT threshold for fixed (but sufficiently large) k was finally determined in [26].

Beyond k-SAT, satisfiability thresholds have been studied for a wide range of other CSPs. These include boolean satisfiability problems, such as 1-in-k-SAT, k-NAE-SAT, and k-XORSAT, and combinatorial problems such as graph colouring and regular subgraphs [1, 25, 5, 40, 11]. The most relevant of these is the k-XORSAT problem. The satsifaibility threshold was first established for 3-XORSAT [27], and later extended to k-XORSAT [24, 45]. Since the k-XORSAT problem is equivalent to linear equations over $\mathbb{F}_2$, it is natural to ask how the satisfiability threshold of linear equations behaves over other algebraic structures. The first result in this direction showed that the satisfiability threshold for linear equations in k variables over $\mathbb{F}_3$ and $\mathbb{F}_4$ coincides exactly with the k-XORSAT threshold for each k [31]. This was later extended to show that the satisfiability threshold of linear equations over any finite field in coincides with the k-XORSAT threshold [7].

Connamacher and Molloy conjectured that this remarkable phenomenon is due to the "unique extendability" property that all linear equations over a field share: for any assignment of $k-1$ variables in an equation over a field, there exists a unique assignment of the remaining variable that extends that assignment to a solution. More specifically, they conjectured that the satisfiability threshold of a separate family of uniquely extendable CSPs also coincides with the k-XORSAT threshold, and they confirmed this conjecture in the case where $k = 3$ and where variables are assigned values in a set of size 4 [18]. Recently, Gao and Morrison proved that a similar conjecture holds for a class of "reducible" uniquely extendable CSPs [29].

The setting of linear equations over a ring is motivated as a natural generalization of linear equations over a field, and as a setting in which to study the relationship between uniquely extendability and the satisfiability threshold. In general, we say that a constraint is *extendable* if there exists at least one way of extending any assignment of a subset of its variables to a solution. Linear equations over a ring provide examples of uniquely extendable constraints, constraints that are extendable in multiple ways, and constraints that are not extendable. The broader setting of rings therefore provides a rich source of examples that interpolate between many classes of CSPs, and provides a setting to ask questions about the relationship between extendability and satisfiability. For instance, over what rings other than fields does the satisfiability threshold of linear equations coincide with the k-XORSAT threshold? What algebraic properties of rings govern the location of the satisfiability thresh-

old? Are these properties related to extendability, or is more structure of the ring needed to determine the threshold? Our contributions, which are explained in Section 1.3, provide partial answers to each of these questions.

1.2 Definitions and preliminaries

We adopt a model of random linear equations similar to [7]. Let $k \geq 3$ be an integer, and let $\hat{P}$ be a distribution over $(R \setminus \{0\})^k$, where $(R, +, \cdot)$ is a finite commutative ring. For each $m, n \in \mathbb{N}$, let $C_m = \{a_1, \dots, a_m\}$ and $V_n = \{v_1, \dots, v_n\}$ be disjoint index sets, and define a random $m \times n$ matrix $\mathbf{B} \in R^{C_m \times V_n}$ over R as follows: independently for each $a \in C_m$, choose a k-tuple $(u_{a,1}, \dots, u_{a,k}) \in V_n^k$ of distinct elements of V_n uniformly at random, and independently sample $(\alpha_{a,1} \dots, \alpha_{a,k})$ according to $\hat{P}$. Define $\mathbf{B}$ by setting $\mathbf{B}_{a,u_{a,i}} = \alpha_{a,i}$ for all $a \in C_m,\ i \in [k]$, and setting $\mathbf{B}_{a,v} = 0$ for all $a \in C_m,\ v \in V_n \setminus \{u_{a,1}, \dots, u_{a,k}\}$. Note that the definition of $\mathbf{B}$ is not symmetric in n and m, as every row of $\boldsymbol{B}$ contains exactly k non-zero entries, and the number of non-zero entries in each column is randomly distributed.

Let Q be a distribution over R and choose a random $\mathbf{b} \in R^{C_m}$ such that $\mathbf{b}$ is independent of $\mathbf{B}$ and the entries of $\mathbf{b}$ are chosen independently from Q. Consider the system of random linear equations over R given by

$$\mathbf{B}x = \mathbf{b}. \tag{1.1}$$

We say that the system (1.1) is satisfiable (SAT) if there exists $x \in R^n$ satisfying all the equations in (1.1). Otherwise, we say that (1.1) is unsatisfiable (UNSAT). We say that a sequence of events $\mathcal{E}_n$ occurs asymptotically almost surely (a.a.s.) if $\lim_{n\to\infty} \mathbb{P}[\mathcal{E}_n] = 1$.

When Q is uniform over R (and thus $\mathbf{b}$ is uniform in R^m), the system (1.1) is unsatisfiable a.a.s. when $m - n = \omega(1)$, as the column space of $\mathbf{B}$ is a subspace of R^m of size at most $|R|^n$. This gives an immediate upper bound on the satisfiability threshold of (1.1). In general, this upper bound is not tight, as the kernel of $\mathbf{B}$ may have an exponential size when $m \sim n$, and thus (1.1) may be unsatisfiable a.a.s. when $m \sim n$.

In the case where R is a field, we can get around this issue by considering a full rank subsystem of (1.1) called the 2-core of $\mathbf{B}x = \mathbf{b}$. For any matrix $B \in R^{C_m \times V_n}$ with k non-zero entries in each row, we define the 2-core of B by iteratively removing some of its rows and columns. In each iteration, we remove a column if that column contains less than 2 non-zero entries. If a column that we remove contains a non-zero entry, we also remove the row that contains that non-zero entry. The process terminates when we obtain a matrix where every column has at least 2 non-zero entries, and we define the 2-core of B to be this matrix. It is not hard to show that the 2-core of B defined by this process is the unique largest submatrix of B that has k non-zero entries in each row and at least 2 non-zero entries in each column. Thus, the 2-core does not depend on the order in which columns are removed. We denote the 2-core of B by $(B)_{\text{2-core}}$. Given a system of equations $Bx = b$, we define the

2-core subsystem $(B)_{2\text{-core}}x = (b)_{2\text{-core}}$ by letting $(b)_{2\text{-core}}$ be the restriction of b to the rows contained in $(B)_{2\text{-core}}$.

When $Bx = b$ is a system of equations over a field, solutions to $(B)_{2\text{-core}}x = (b)_{2\text{-core}}$ can be extended to solutions to $Bx = b$ by iteratively working backwards through the 2-core process. More precisely, if x is an assessment of variables that satisfies every row remaining after i iterations of the 2-core process, and v is the column removed at iteration i, then there exists an extension of x to the column v that satisfies every row remaining after $i-1$ iterations of the 2-core process. Such an extension exists by the unique extendability property of linear equations over a field discussed in Section 1.1. By repeating this process of extensions, we get a solution to the original system $Bx = b$.

Since any solution to the subsystem $(B)_{2\text{-core}}x = (b)_{2\text{-core}}$ can be extended to a solution to $Bx = b$, the system (1.1) is satisfiable if and only if $(\mathbf{B})_{2\text{-core}}x = (\mathbf{b})_{2\text{-core}}$ is satisfiable. Therefore the threshold for when $(\mathbf{B})_{2\text{-core}}$ has more rows than columns is an upper bound on the satisfiability threshold of (1.1). The dimensions of $(\mathbf{B})_{2\text{-core}}$ are well-studied in the literature [3, 38], and the threshold for when the number of rows exceeds the number of columns is given by (1.2) below. The corresponding lower bound was established in [7], where it was shown that $(\mathbf{B})_{2\text{-core}}$ has full rank a.a.s. The exact satisfiability threshold for linear equations over a field is therefore given as follows.

Theorem 1.2.1 ([7, Theorem 1.1])**.** *Let $k \geq 3$ be a fixed integer, and let*

$$\begin{aligned} \rho_{k,d} &= \sup\left\{x \in [0,1] : x = 1 - \exp(-dx^{k-1})\right\} \\ d_k &= \inf\left\{d > 0 : \rho_{d,k} - d\rho_{d,k}^{k-1} + (1-1/k)d\rho_{d,k}^{k} < 0\right\}. \end{aligned} \tag{1.2}$$

Suppose that R is a finite field and that Q is the uniform distribution on R. Then, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}[\textit{System (1.1) is satisfiable}] = \begin{cases} 1 & m < (d_k - \varepsilon)n/k \\ 0 & m > (d_k + \varepsilon)n/k. \end{cases}$$

Note that the entries of $\mathbf{B}$ are not independent under the assumptions of Theorem 1.2.1, as $\hat{P}$ is not assumed to by a product measure. However, the choice $\hat{P} = P^{\otimes k}$ for some distribution P over $R \setminus \{0\}$ is among the most natural and important cases, and will be our primary focus.

The second assumption of Theorem 1.2.1, that Q is the uniform distribution on R, is overly restrictive in the setting where R is a ring. The following example illustrates this.

Example 1.2.2. *Suppose $\hat{P} = P^{\otimes k}$ for some distribution P on $R \setminus \{0\}$ such that* $\mathrm{supp}(P)$ *contains a nonunit $r \in R$. Suppose Q is uniform on R. Then, for each $a \in C_m$, the event $\{\alpha_{a,i} = r,\ i \in [k]\} \cap \{\mathbf{b}_a = 1\}$ occurs with positive probability. Therefore each row of system (1.1) is unsatisfiable with positive probability. Since every row is chosen independently, this implies that (1.1) is unsatisfiable a.a.s. whenever the number of rows $m = m(n)$ is unbounded.*

Since we would like to study linear (or at least $\omega(1)$) satisfiability thresholds for more general distributions P, we therefore do not assume that Q is uniform on R.

Even in the case where Q is not uniform over R, the system (1.1) is still trivially satisfiable and trivially unsatisfiable for some choices of P and Q. To avoid cases were (1.1) is trivially unsatisfiable, we assume that $(\hat{P}, Q)$ is not inconsistent in the following sense.

Definition 1.2.3. *Let $\hat{P}$ be a distribution on R^k and let Q be a distribution on R. We say that $(\hat{P}, Q)$ is inconsistent if, for all $(a_1, \ldots, a_k) \in \mathrm{supp}(\hat{P})$ and all $b \in \mathrm{supp}(Q)$, b is not in the ideal generated by $\{a_1, \ldots, a_k\}$.*

If $(\hat{P}, Q)$ are inconsistent, then every equation $\sum_{i=1}^{k} a_i x_i = b$ with $(a_1, \ldots, a_k) \in \mathrm{supp}(\hat{P})$, $b \in \mathrm{supp}(Q)$ is unsatisfiable, and so (1.1) is trivially unsatisfiable for any value of $m > 0$. We assume the following throughout this thesis.

(A1) $(\hat{P}, Q)$ is not inconsistent.

Next, we consider an example where system (1.1) is trivially satisfiable.

Definition 1.2.4. *Let $\hat{P}$ be a distribution on R^k and let Q be a distribution on R. We say that $(\hat{P}, Q)$ admits a k-ary constant solution $r \in R$ if, for all $(a_1, \ldots, a_k) \in \mathrm{supp}(\hat{P})$ and $b \in \mathrm{supp}(Q)$,*

$$\sum_{i=1}^{k} a_i r = b.$$

Clearly, if $(\hat{P}, Q)$ admits a k-ary constant solution r, then the variable assignment $x_v = r$ for all $v \in V_n$ satisfies (1.1). This is similar to a homogeneous system with $b = 0$, but there are examples of (P, Q) with $\mathrm{supp}(Q) \neq \{0\}$ that admit a k-ary constant solution. For example, $R = \mathbb{Z}_4$ with $\mathrm{supp}(P) = \{1, 3\}$, $\mathrm{supp}(Q) = \{2\}$, and $k = 3$ admits a constant solution 2. To avoid this second trivial case, we make another universal assumption about $(\hat{P}, Q)$.

(A2) $(\hat{P}, Q)$ does not admit any k-ary constant solutions.

1.2.1 Factor graphs

Factor graphs have been used throughout the CSP literature as a graphical representation of the variables that each constraint takes as its arguments [33]. In this thesis, we use factor graphs to study the satisfiability of system (1.1), and we prove contiguity results among more general models of random factor graphs considered in the CSP literature. In the more general setting, we define a factor graph to be any subgraph of the complete bipartite graph on vertices $C_m \cup V_n$. Given a factor graph F, let $E(F)$ denote the edge set of F, let $C(F)$

denote the set of vertices of F contained in C_m, and let $V(F)$ denote the set of vertices of F contained in V_n. We call $C(F)$ the set of constraint vertices of F, and we call $V(F)$ the set of variable vertices of F. For $u \in C_m \cup V_n$, the degree of u is denoted by $\deg(u)$, and the set of neighbours of u is denoted by $N(u)$.

Given a system of equations $Bx = b$ for some $B \in R^{C_m \times V_n}$ and $b \in R^{C_m}$, we define an associated factor graph $F(B)$ and weighted factor graph $F(B, b) = (F(B), w)$ as follows: let $F(B)$ be the factor graph on vertices $C_m \cup V_n$ with edge set $\{(a, v) \in C_m \times V_n : B_{a,v} \neq 0\}$; let w be the weight function on edges and constraint vertices of $F(B)$ given by

$$w(a) = b_a, \quad a \in C_m, \qquad\qquad w(av) = B_{a,v}, \quad av \in E(F(B)).$$

Clearly, this gives a bijection between the set of matrices in $R^{C_m \times V_n}$ and the set of weighted factor graphs with vertex set $C_m \cup V_n$. Given a weighted factor graph (F, w) we say that (F, w) is satisfiable (resp. unsatisfiable) if the unique system of equations $Bx = b$ with $F(B, b) = (F, w)$ is satisfiable (resp. unsatisfiable).

We say that a factor graph F is k-uniform if all of its constraint vertices have degree k. In most of this thesis, we focus on k-uniform factor graphs, as the constraint vertices of $F(\mathbf{B})$ for the model $\mathbf{B}$ from (1.1) all have degree k. In Chapter 2, we consider other models of random factor graphs that are not k-uniform, and give contiguity results for these models. These contiguity results are used Chapter 6, and may be of independent interest.

1.3 Contributions

As mentioned in Section 1.1, the SAT thresholds of many classically studied CSPs are linear in the number of variables. These include linear equations over a field, as Theorem 1.2.1 shows.

For our model of linear equations over rings, we find that the satisfiability threshold can be either linear or sublinear in the number of variables, depending on the support of the distributions P and Q. In both the sublinear and linear regimes, the satisfiability threshold is easier to determine under the assumption that R is a principal ideal ring (PIR). We discuss this special case before moving to the more general setting.

1.3.1 Principal ideal rings

Our first result provides a complete characterization of the sublinear SAT threshold, determines the associated coarse threshold, and establishes that the phase transition is not sharp in the sublinear regime. To state the result we need a technical definition. We use the notation (r) and (S) to denote the ideals generated by $r \in R$ and $S \subseteq R$ respectively. We use $(I : J)$ to denote the ideal quotient $(I : J) = \{t \in R : tJ \subseteq I\}$ for any two ideals $I, J \subseteq R$.

Definition 1.3.1. *Let P and Q be distributions over R. Define a sequence of ideals $I_0(P), I_1(P), \ldots$ inductively by setting $I_0(P) = \bigcap_{r \in \mathrm{supp}(P)}(r)$, and setting*

$$I_\ell(P) = \bigcap_{r \in \mathrm{supp}(P)} r\left(\bigcap_{s \in \mathrm{supp}(P)} (I_{\ell-1}(P) : (s)) \right), \quad \ell \geq 1.$$

Define

$$\ell(P, Q) = \inf\{\ell \geq 0 : I_\ell(P) \not\supseteq \mathrm{supp}(Q)\}$$

with the convention that $\inf(\emptyset) = \infty$.

Roughly speaking, the integer $\ell(P, Q)$ measures the height of the smallest unsatisfiable acyclic factor graph with edge weights in $\mathrm{supp}(P)$ and constraint vertex weights in $\mathrm{supp}(Q)$. Section 4.2, particularly Lemmas 4.2.2 and 4.2.6, gives a more precise statement of the relation between $\ell(P, Q)$ and the satisfiability of acyclic factor graphs. By calculating the threshold at which the smallest unsatisfiable acyclic factor graph appears, we get the following theorem.

Theorem 1.3.2. *Let P and Q be distributions on a principal ring R and let $\hat{P} = P^{\otimes k}$. Suppose $(\hat{P}, Q)$ satisfies* **(A1)** *and* **(A2)**. *Set*

$$\alpha_{P,Q} = \frac{\sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}}{1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}}, \tag{1.3}$$

with $\alpha_{P,Q} = 1$ if $\ell(P, Q) = \infty$. Then,

$$\lim_{n\to\infty} \mathbb{P}[\textit{System (1.1) is satisfiable}] = \begin{cases} 1 & m \ll n^{\alpha_{P,Q}}, \\ 0 & m \gg n^{\alpha_{P,Q}}. \end{cases}$$

Moreover, if $\alpha_{P,Q} < 1$ and $m = Cn^{\alpha_{P,Q}}$, then for all fixed $C > 0$,

$$0 < \liminf_{n\to\infty} P[\textit{System (1.1) is satisfiable}] \leq \limsup_{n\to\infty} P[\textit{System (1.1) is satisfiable}] < 1. \tag{1.4}$$

Remark. If (P, Q) is such that $(\hat{P}, Q)$ is inconsistent, then $\alpha_{P,Q} = 0$, and (1.1) is trivially unsatisfiable for every $m \geq 1$. Thus, the inequality (1.4) does not hold. Hence it is necessary to assume **(A1)** in this theorem.

This theorem shows that the SAT phase transition is not sharp when $\alpha_{P,Q} < 1$, in contrast to Theorem 1.2.1 in the field case. What remains to be understood is the case where $\alpha_{P,Q} = 1$. Is there a sharp phase transition for SAT in this regime? We confirm it in the case that R is principal. Moreover, we show that the SAT threshold coincides exactly with the threshold d_k in Theorem 1.2.1. As discussed in Section 1.1, d_k has been shown to be the satisfiability threshold for many classes of uniquely extendable CSPs [18, 7]. Our result extends this observation to linear equations over a PIR.

Theorem 1.3.3. *Suppose that R is principal, that P, Q are distributions over R such that $\alpha_{P,Q} = 1$. Suppose $\hat{P} = P^{\otimes k}$ and $(\hat{P}, Q)$ satisfies* **(A2)**. *Let d_k be as defined in Theorem 1.2.1. Then, for every fixed $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}[\textit{System (1.1) is satisfiable}] = \begin{cases} 1 & \text{if } m < (d_k - \varepsilon)n/k, \\ 0 & \text{if } m > (d_k + \varepsilon)n/k. \end{cases}$$

Remark. Note that **(A1)** is satisfied by the assumption that $\alpha_{P,Q} = 1$.

The following example shows that there are choices of R, P and Q that satisfy the assumptions of Theorem 1.3.3, such that non-uniquely extendable equations appear in the resulting system (1.1).

Example 1.3.4. *Let $R = \mathbb{Z}_6$, and let P, Q be such that* $\operatorname{supp}(P) = \{1, 3, 5\}$ *and* $\operatorname{supp}(Q) = \{3, 0\}$. *Suppose* $\gcd(k-1, 6) = 1$. *Then $\alpha_{P,Q} = 1$, and there exist $r_1, \ldots, r_k \in \operatorname{supp}(P)$, $b \in \operatorname{supp}(Q)$, and $x_1, \ldots x_{k-1} \in R$ such that $r_1x_1 + \ldots + r_{k-1}x_{k-1} + r_kx_k = b$ is not satisfiable for any $x_k \in R$.*

Proof. For the second assertion, choose $r_1, \ldots, r_{k-1} = 1$, $r_k = 3$, $b = 0$, and $x_1, \ldots, x_{k-1} = 1$. Then $r_1x_1 + \ldots + r_{k-1}x_{k-1} = k - 1 \pmod 6$ is invertible in $\mathbb{Z}_6$ and $r_k = 3$ is not invertible in $\mathbb{Z}_6$. Therefore there is no value of $x_k \in \mathbb{Z}_6$ such that $r_1x_1 + \ldots + r_{k-1}x_{k-1} + r_kx_k = 0$.

To show that $\alpha_{P,Q} = 1$, by (1.3) it suffices to show that $I_\ell(P) \supseteq (3)$ for all $\ell \geq 0$, as we then have $\ell(P, Q) = \infty$ by Definition 1.3.1. We prove by induction on $\ell \geq 0$. For $\ell = 0$, we have $I_0(P) = (1) \cap (3) \cap (5) = (3)$. For $\ell \geq 1$, we have

$$((3) : (1)) = (3), \quad ((3) : (3)) = \mathbb{Z}_6, \quad ((3) : (5)) = (3).$$

Thus, assuming $I_{\ell-1}(P) \supseteq (3)$, we get

$$I_\ell(P) \supseteq \bigcap_{r \in \operatorname{supp}(P)} r \cdot (3) = (3) \cap (9) \cap (15) = (3).$$

Therefore the claim follows by induction. □

1.3.2 Non-principal ideal rings

In this section we drop the assumption that R is a principal ideal ring. In this setting, we are still able to characterize the sublinear SAT threshold and establish that the phase transition is not sharp in the sublinear regime. However, determining the location of the sublinear threshold requires solving an optimization problem over a set of acyclic factor graphs, which we define below. The advantage of the principal ideal case is that we do not need to solve this optimization problem to find the threshold.

Recall from Section 1.2.1 that every pair $B \in R^{C_m \times V_n}$, $b \in R^{C_m}$ is associated with a weighted factor graph $F(B,b)$, and $F(B,b)$ is said to be unsatisfiable if $Bx = b$ is unsatisfiable. For any distributions P, Q on R, let $\mathcal{T}(P,Q)$ be the set of unsatisfiable connected acyclic k-uniform weighted factor graphs whose edge weights are in $\operatorname{supp}(P)$ and whose constraint vertex weights are in $\operatorname{supp}(Q)$. Recall that $C(T)$ is the set of constraint vertices of a factor graph T.

Theorem 1.3.5. *Let P and Q be distributions on a ring R, and let $\hat{P} = P^{\otimes k}$. Suppose $\hat{P}$ and Q satisfy* **(A1)** *and* **(A2)**. *Define*

$$\alpha_{P,Q} = 1 - \frac{1}{\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\}}.$$

with $\alpha_{P,Q} = 1$ if $\mathcal{T}(P,Q) = \emptyset$. Then

$$\lim_{n\to\infty} \mathbb{P}[\textit{System (1.1) is satisfiable}] = \begin{cases} 1 & m \ll n^{\alpha_{P,Q}}, \\ 0 & m \gg n^{\alpha_{P,Q}}. \end{cases}$$

Moreover, if $\alpha_{P,Q} < 1$ and $m = Cn^{\alpha_{P,Q}}$, then for all fixed $C > 0$,

$$0 < \liminf_{n\to\infty} P[\textit{System (1.1) is satisfiable}] \le \limsup_{n\to\infty} P[\textit{System (1.1) is satisfiable}] < 1. \quad (1.5)$$

As in Theorem 1.3.2, the inequality (1.5) shows that the satisfiability threshold is coarse in the sublinear regime. While $\ell(P,Q)$ does not determine the value of $\alpha_{P,Q}$ as it does in Theorem 1.3.3, the proof of Theorem 1.3.3 still yields the following bounds on $\alpha_{P,Q}$.

Theorem 1.3.6. *Let P and Q be distributions on a ring R, and let $p = |\operatorname{supp}(P)|$. If $\ell(P,Q) < \infty$, then*

$$\frac{\sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}}{1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}} \le \alpha_{P,Q} \le \frac{\sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}p^{2i}}{1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}p^{2i}}.$$

The bounds in Theorem 1.3.6 are not tight (see Example 4.2.7), but they do imply the following characterization of the linear and sublinear regimes over arbitrary finite commutative rings.

Theorem 1.3.7. *Let P and Q be distributions on a ring R. Then $\alpha_{P,Q} = 1$ if and only if $\ell(P,Q) = \infty$.*

We conjecture that there is a sharp SAT threshold at $\beta n/k$ for some $\beta > 0$ when $\alpha_{P,Q} = 1$, although the value of β may not coincide with d_k.

Conjecture 1.3.8. *Suppose that P, Q are distributions over R such that $\alpha_{P,Q} = 1$ and $\hat{P} = P^{\otimes k}$ satisfies* **(A2)**. *There exists a constant $\beta = \beta(P,Q,R,k) > 0$ such that for all fixed $\varepsilon > 0$,*

$$\lim_{n\to\infty} P[\textit{System (1.1) is satisfiable}] = \begin{cases} 1 & m < (\beta - \varepsilon)n/k, \\ 0 & m > (\beta + \varepsilon)n/k. \end{cases}$$

The value of $\beta(P, Q, R, k)$ is given by Theorem 1.3.3 in the case where R is a principal ideal ring, but we do not attempt to determine the value of $\beta(P, Q, R, k)$ for general rings in this thesis. In the theorem below, we determine $\beta(P, Q, R, k)$ for some examples of (P, Q, R, k) where R is non-principal ideal ring, which establishes that $\beta(P, Q, R, k)$ can be different from d_k in Theorem 1.3.3. Moreover, the value of $\beta(P, Q, R, k)$ can be different for distributions P that have the same support. This is in contrast to the case that R is a finite field or a finite commutative principal ideal ring.

Given a prime power q, let $\mathbb{F}_q$ denote the finite field of order q. Let $\mathbb{F}_q[X, Y]$ denote the polynomial ring in variables X and Y over $\mathbb{F}_q$. For the theorem below, we consider $R = \mathbb{F}_q[X, Y]/(X^2, Y^2)$, i.e. the quotient ring of $\mathbb{F}_q[X, Y]$ by the ideal generated by $\{X^2, Y^2\}$. For any $p(X, Y) \in \mathbb{F}_q[X, Y]$, let $\overline{p(X, Y)}$ denote its projection in R.

Theorem 1.3.9. *Let $R = \mathbb{F}_q[X, Y]/(X^2, Y^2)$ and let $\rho \in [0, 1]$. Suppose that P is supported on $\{\bar{X}, \bar{Y}\}$, and that P takes value $\bar{X}$ with probability ρ and value $\bar{Y}$ with probability $1 - \rho$. Suppose that $\mathrm{supp}(Q) = \{\bar{X}\bar{Y}, 0\}$. Define $d_{k,\rho}$ by*

$$d_{k,\rho} = \sup\{d > 0 : \Phi_{d,k,\rho}(z) < \Phi_{d,k,\rho}(0) \text{ for all } z \in (0, 1]\}, \tag{1.6}$$

where

$$\Phi_{d,k,\rho}(z) = e^{-\rho d z^{k-1}} + e^{-(1-\rho) d z^{k-1}} - d(1 - 1/k)z^k + dz^{k-1} - d/k.$$

Then, for any fixed $\varepsilon > 0$,

$$\lim_{n\to\infty} P[\text{System (1.1) is satisfiable}] = \begin{cases} 1 & \text{if } m < (d_{k,\rho} - \varepsilon)n/k, \\ 0 & \text{if } m > (d_{k,\rho} + \varepsilon)n/k. \end{cases}$$

We conclude this section by comparing $d_{k,\rho}$ to the field threshold d_k for some values of $\rho \in [0, 1]$. This illustrates how the satisfiability threshold in Theorem 1.3.9 depends on ρ, in contrast to linear equations over fields and principal ideal rings. We first show analytically that $d_{k,\rho}$ depends on the value of ρ.

Claim 1.3.10. *For any $k \geq 3$, $d_{k,0} = d_{k,1} = d_k$ and $d_{k,1/2} = 2d_k$.*

Proof. Taking a derivative with respect to z shows that

$$\Phi'_{d,k,\rho}(z) = dz^{k-2}(k - 1)\phi_{d,k,\rho}(z),$$

where

$$\phi_{d,k,\rho}(z) = 1 - z - \rho e^{-\rho d z^{k-1}} - (1 - \rho)e^{-(1-\rho) d z^{k-1}}.$$

Let $\eta_{d,k,\rho} = \sup\{z \in [0, 1] : \phi_{d,k,\rho}(z) = 0\}$, and let $d^* = \inf\{d > 0 : \eta_{d,k,\rho} > 0\}$. For $d \leq d^*$, we have $\phi_{d,k,\rho}(z) \leq 0$ for all $z \in [0, 1]$, and so $\Phi_{d,k,\rho}(z) < \Phi_{d,k,\rho}(0)$ for all $z \in (0, 1]$. For $d > d^*$, the function $\phi_{d,k,\rho}$ has two zeros on $(0, 1)$. The largest zero is $\eta_{d,k,\rho}$, which is a local minimum of $\Phi'_{d,k,\rho}$. The smaller zero is a local maximum of $\Phi_{d,k,\rho}$. Therefore, using the fact

that $\Phi_{d,k,\rho}$ is continuously differentiable, we have $\Phi_{d,k,\rho}(z) \geq \min\{\Phi_{d,k,\rho}(0), \Phi_{d,k,\rho}(\eta_{d,k,\rho})\}$ for $z \in [0,1]$, and hence

$$d_{k,\rho} = \sup\{d > 0 : \Phi_{d,k,\rho}(0) - \Phi_{d,k,\rho}(\eta_{d,k,\rho}) > 0\}.$$

For $\rho = 0$, we have $\eta_{d,k,\rho} = \rho_{d,k}$, where $\rho_{d,k}$ is defined in (1.2). Since $\rho_{d,k}$ satisfies $\rho_{d,k} = 1 - e^{d\rho_{d,k}^{k-1}}$, we get

$$\begin{aligned}\Phi_{d,k,0}(0) - \Phi_{d,k,\rho}(\eta_{d,k,0}) &= 1 - e^{-d\rho_{d,k}^{k-1}} + d(1-1/k)\rho_{d,k}^k - d\rho_{d,k}^{k-1} \\ &= \rho_{d,k} + d(1-1/k)\rho_{d,k}^k - d\rho_{d,k}^{k-1}.\end{aligned}$$

The implicit function theorem shows that $\Phi_{d,k,0}(0) - \Phi_{d,k,\rho}(\eta_{d,k,0})$ is differentiable and decreasing as a function of d for $d > d^*$, and so

$$d_{k,0} = \sup\{d > 0 : \Phi_{d,k,\rho}(0) - \Phi_{d,k,\rho}(\eta_{d,k,\rho}) > 0\} = d_k,$$

A symmetric argument shows that $d_{k,1} = d_k$. For $\rho = 1/2$, we have

$$\eta_{d,k,a/2} = \sup\{z \in [0,1] : z = 1 - e^{-dz^{k-1}/2}\} = \rho_{d/2,k}.$$

and hence

$$\begin{aligned}\Phi_{d,k,1/2}(0) - \Phi_{d,k,1/2}(\eta_{d,k,1/2}) &= 2 - 2e^{-d\rho_{d/2,k}^{k-1}/2} + d(1-1/k)\rho_{d/2,k}^k + d\rho_{d/2,k}^{k-1} \\ &= 2\rho_{d/2,k} + d(1-1/k)\rho_{d/2,k}^k + d\rho_{d/2,k}^{k-1},\end{aligned}$$

Again using the implicit function theorem, we get

$$d_{k,1/2} = \sup\{d > 0 : \Phi_{d,k,1/2}(0) - \Phi_{d,k,1/2}(\eta_{d,k,1/2}) > 0\} = 2d_k.$$

□

Using the definition of $d_{k,\rho}$ in Theorem 1.3.9, we can approximate the value of $d_{k,\rho}$ numerically. Figure 1.1 below plots $d_{3,\rho}$ as a function of ρ on the interval $[0,1]$. The graph appears to be concave, with a unique maximum at $\rho = 1/2$. From numerical calculations, we also see that $d_{3,0} \approx 2.754$ and $d_{3,1/2} \approx 5.508$. Figure 1.1 was generated by numerically approximating $d_{3,\rho}$ at 100 values of $\rho \in [0,1]$ to an accuracy of 0.001.

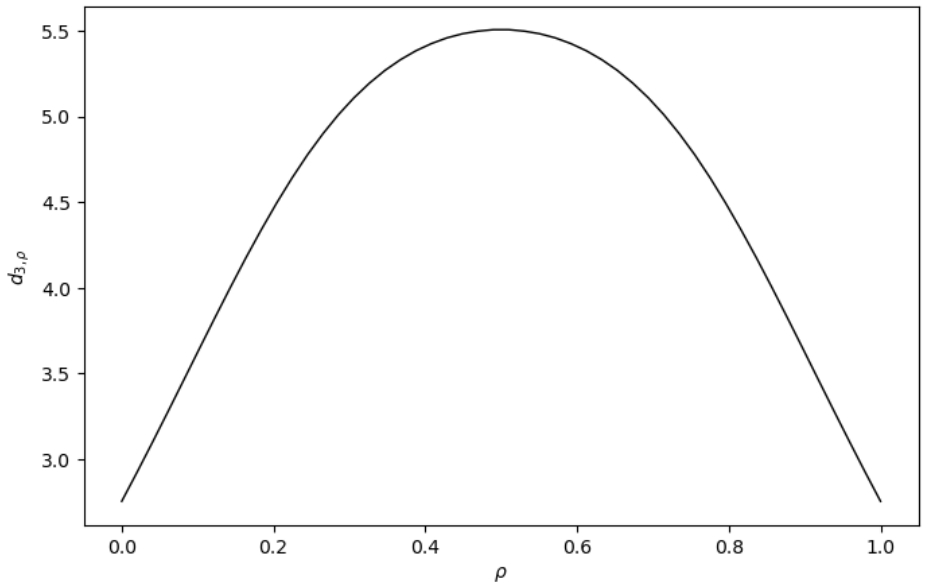


Figure 1.1: $d_{3,\rho}$ as a function of $\rho \in [0,1]$

1.4 Open problems

In Theorem 1.3.2 we determined the coarse satisfiability threshold of the random system (1.1) to be $n^{\alpha_{P,Q}}$ where $\alpha_{P,Q}$ is given explicitly as an easily computable function of P, Q and R under the assumption that R is a finite commutative principal ideal ring. This explicit characterization does not extend to the case of non-principal ideal rings; see Theorem 1.3.5. Although $\alpha_{P,Q}$ is well defined for non-principal ideal rings, its definition involves the solution of an optimization problem over the space $\mathcal{T}(P,Q)$. Evaluating this optimization problem is challenging, largely because the members of $\mathcal{T}(P,Q)$ are not well understood. Our first open problem is to find an explicit expression for $\alpha_{P,Q}$ for non-principal ideal rings.

Problem 1.4.1. *Find an explicit expression for $\alpha_{P,Q}$ as an easily computable function of P, Q and R when R is not a principal ideal ring.*

Conjecture 1.3.8 states that a sharp satisfiability threshold exists for distributions (P,Q) over any finite commutative ring R if $\alpha_{P,Q} = 1$. Theorem 1.3.3 verifies this conjecture and determines the location of the threshold in the case where R is a principal ideal ring. Our next open problem is to determine the value of this threshold for general rings, assuming the validity of Conjecture 1.3.8.

Problem 1.4.2. *Determine $\beta(P,Q,R,k)$ in Conjecture 1.3.8 for non-principal ideal rings.*

Other open problems arise from studying $\mathbf{B}$ as a random matrix, as opposed to studying the system $\mathbf{B}x = \mathbf{b}$ as a CSP. Determining the rank of the random matrix $\mathbf{B}$ is one possible

direction that has received considerable interest in the field setting [13, 14]. However, the rank of a matrix is not as easy to define over an arbitrary ring as it is over a field. Some possibilities include the "determinant rank," which is the size of the largest square submatrix with an invertible determinant, and the "inner rank," which is the smallest value r such that an $m \times n$ matrix can be factored into a product of $m \times r$ and $r \times n$ matrices. Full row rank does not correspond directly to satisfiability of the system $\mathbf{B}x = \mathbf{b}$ under either of these definitions, and thus we do not study the rank of $\mathbf{B}$ in this thesis, but the problem of determining the rank of sparse random matrices over a ring may be of independent interest.

While the rank of matrix is difficult to define, the kernel and co-kernel of a matrix are well-defined over any ring, and can serve as an analog for rank and nullity. The co-kernels of random matrices have recently been studied for integral matrices and matrices over arbitrary commutative rings [44, 35], although much of this work has been for denser random matrix models than the one we consider here. Determining the asymptotic size of the kernel would be a first step in this direction.

Problem 1.4.3. *Let $\mathbf{B}$ be the random matrix defined as above (1.1) and suppose that $m \sim dn$ where $d > 0$ is fixed. Determine the asymptotic values of*

$$\limsup_{n\to\infty} n^{-1} \log |ker(\mathbf{B})| \text{ and } \liminf_{n\to\infty} n^{-1} \log |ker(\mathbf{B})|.$$

When R is a field, the above two limits are equal, and their precise value as a function of d was determined in [13].

Our final open problems concern the solution space geometry of $\mathbf{B}x = \mathbf{b}$. As discussed in Section 1.1, many CSPs are observed to undergo a clustering phase transition at a threshold value slightly below the satisfiability threshold. To define solution clusters more precisely, we say that two solutions x', x'' to the system $\mathbf{B}x = \mathbf{b}$ are reachable from each other if there exists a sequence of solutions $x_1, \dots, x_\ell$ such that $x_1 = x'$, $x_\ell = x''$, and such that the Hamming distance between x_i and x_{i+1} is at most $\log(n) \log\log(n)$ for $i \in [\ell - 1]$. Reachability clearly defines an equivalence relation on the solution space of $\mathbf{B}x = \mathbf{b}$, and the solution clusters are the equivalence classes of the connectedness relation. [7, Theorem 1.4] shows that linear equations over a field undergo a clustering phase transition at which the solution space slits into exponentially many clusters. The following conjecture asserts that linear equations over a finite commutative ring have a similar sharp clustering threshold in the linear regime.

Conjecture 1.4.4. *Suppose P and Q are distributions on R such that $\alpha_{P,Q} = 1$. Then there is a constant $\gamma = \gamma(P, Q, R, k) > 0$ such that the following hold for all fixed $\varepsilon > 0$:*

1. *If $m < (\gamma - \varepsilon)n$, then a.a.s. every solution to (1.1) is in the same cluster.*
2. *If $m > (\gamma + \varepsilon)n$, then a.a.s. the solution space of (1.1) is partitioned into exponentially many clusters, and the minimum Hamming distance between any two solutions in distinct clusters is $\Omega(n)$.*

Problem 1.4.5. *Determine $\gamma(P, Q, R, k)$ in Conjecture 1.4.4 for general rings R.*

Chapter 2

Random Models and Contiguity

In this section we describe general models of random factor graphs used in the literature, and present contiguity results among these models. Contiguity results for factor graphs are useful in the setting of random linear equations because, as discussed in Chapter 1, a system of equations can be specified by choosing a factor graph and a weight function. The results presented here will be used in Chapter 6, and may be of independent interest for translating high probability properties from one model to another.

Coja-Oghlan, Gao, Hahn-Klimroth, Lee, Müller, and Rolvien in [14] determined the satisfiablity random linear equations over finite fields in great generality. The factor graph for their model of random linear equations is described as follows. Let $\boldsymbol{d}, \boldsymbol{k} \geq 0$ be integer valued random variables such that $\mathbb{E}[\boldsymbol{k}^{2+\varepsilon}], \mathbb{E}[\boldsymbol{d}^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$. Set $d = \mathbb{E}[\boldsymbol{d}]$ and $k = \mathbb{E}[\boldsymbol{k}]$. Let $\boldsymbol{m}', \{\boldsymbol{k}'_{a_i}\}_{i\in\mathbb{N}}, \{\boldsymbol{d}'_{v_i}\}_{i\in\mathbb{N}}$ be mutually independent random variables, where $\boldsymbol{m}'$ is a Poisson random variable with mean dn/k, $\{\boldsymbol{k}'_a\}_{a\in C_m}$ are i.i.d. copies of $\boldsymbol{k}$, and $\{\boldsymbol{d}'_v\}_{v\in V_n}$ are i.i.d. copies of $\boldsymbol{d}$. Let $\mathcal{D}_n$ be the event

$$\sum_{i=1}^{n} \boldsymbol{d}'_{v_i} = \sum_{i=1}^{\boldsymbol{m}'} \boldsymbol{k}'_{a_i}. \tag{2.1}$$

Choose $\boldsymbol{m}, \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}$ from the conditional distribution of $\boldsymbol{m}', \{\boldsymbol{k}'_a\}_{a\in C_m]}\{\boldsymbol{d}'_v\}_{v\in V_n}$ given $\mathcal{D}$. Let $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ be the random factor graph choosen from the uniform distribution on simple bipartite graphs on $C_{\boldsymbol{m}} \cup V_n$ with degree sequence $\{\boldsymbol{k}_a\}_{a\in C_{\boldsymbol{m}}}$, $\{\boldsymbol{d}_v\}_{v\in V_n}$. This is precisely the random graph model used in [14].

Many parts of our results rely on reductions to a system $Bx = b$ of linear equations over a finite field, where $F(B)$ is distributed approximately as $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ by choosing appropriate distributions for $\boldsymbol{d}$ and $\boldsymbol{k}$. However, there are several drawbacks of the model $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ that make it hard to apply results in [14]. One is the i.i.d. condition on $\{\boldsymbol{d}'_v\}_{v\in V_n}$. For instance, by taking $\mathbf{B}$ with distribution in (1.1), the distribution of variable vertex degrees $\{\deg_{F(\mathbf{B})}(v)\}_{v\in V_n}$ of $F(\mathbf{B})$ is approximately independent Poisson but not exactly. The second drawback is the conditioning on event $\mathcal{D}_n$, which is an event with vanishing probability.

Next we define a sequence of models which we show to be contiguous with the random graph $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$. Let $\{\boldsymbol{k}_a\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{d}_v\}_{v\in V_n}$ and $\{\boldsymbol{k}'_a\}_{a\in C_{\boldsymbol{m}'}}, \{\boldsymbol{d}'_v\}_{v\in V_n}$ be given as above.

- $G_{n,\boldsymbol{m}',\boldsymbol{k}}$: Independently for each $a \in C_{\boldsymbol{m}'}$, choose a uniform random subset of V_n of size $\boldsymbol{k}'_a$ and add and edge from a to each variable in this set.
- $M_{n,\boldsymbol{m}',\boldsymbol{k}}$: Independently for each $a \in C_{\boldsymbol{m}'}$, generate $(u_1, \dots, u_{\boldsymbol{k}'_a},)$ where each u_i is chosen uniformly and independently from V_n. Then, add the $\boldsymbol{k}'_a$ edges $\{au_i : i \in [\boldsymbol{k}'_a]\}$. Note that the resulting random graph $M_{n,\boldsymbol{m}',\boldsymbol{k}}$ can be a multigraph since the elements in $\{u_1, \dots, u_{\boldsymbol{k}'_a}\}$ are not necessarily distinct for every $a \in C_{\boldsymbol{m}'}$.
- $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$: Choose a random bipartite multigraph with degree sequence $\{\boldsymbol{k}_a\}_{a\in C_{\boldsymbol{m}}}$, $\{\boldsymbol{d}_v\}_{v\in V_n}$ generated by the pairing model, described as follows. Let Γ be a random uniformly chosen perfect matching of the complete bipartite graph with vertex classes $\bigcup_{a\in C_{\boldsymbol{m}}}\{a\} \times [\boldsymbol{k}_a]$ and $\bigcup_{v\in V_n}\{v\} \times [\boldsymbol{d}_v]$. For each $a \in C_{\boldsymbol{m}}, v \in V_n$, add an edge between a and v in $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ for each edge between the vertex sets $\{a\} \times [\boldsymbol{k}_a]$ and $\{v\} \times [\boldsymbol{d}_v]$ in Γ.
- $G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$: Choose from the conditional distribution of $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ given that $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ has no parallel edges.

We also derive contiguity results for these models conditioned on having a deterministic number of constraint vertices $m = m(n)$. For our arguments to work, we need the probability of having exactly m constraint vertices to be at least $\Omega(1/\sqrt{n})$ for each of the models. This is ensured by the following assumptions on n and m.

1. $|m - dn/k| = O(\sqrt{n})$,
2. If $\mathrm{Var}(\boldsymbol{k}) = 0$, then $\gcd(\boldsymbol{d}) \mid km$, where $\gcd(\boldsymbol{d})$ is the gcd of the support of $\boldsymbol{d}$.
3. If $\mathrm{Var}(\boldsymbol{d}) = 0$, then $\gcd(\boldsymbol{k}) \mid dn$, where $\gcd(\boldsymbol{k})$ is the gcd of the support of $\boldsymbol{k}$.
4. If $\mathrm{Var}(\boldsymbol{k}) = \mathrm{Var}(\boldsymbol{d}) = 0$, then $dn = km$.

Define random graphs $G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$, $G_{n,m,\boldsymbol{k}}$, $M_{n,m,\boldsymbol{k}}$, $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$, and $G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ by conditioning the respective models above on the event that they have exactly m constraint vertices.

Remark 2.0.1.

(a) Note that, for $\mathbf{B}$ distributed as in (1.1), $F(\mathbf{B})$ has the same distribution as $G_{n,m,\boldsymbol{k}}$ where $\boldsymbol{k} = k$.

(b) [13] assumed that $G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ and $G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ have the same distribution, which is untrue. We fix this error here by showing that these two models are asymptotically equivalent (see Theorem 2.0.4 below).

Our main contiguity results are stated in the theorems below, using the following standard definition of contiguity for random graphs.

Definition 2.0.2. *Let $\{N_n\}_{n\in\mathbb{N}}$ and $\{N'_n\}_{n\in\mathbb{N}}$ be sequences of random graphs. We say that $\{N_n\}_{n\in\mathbb{N}}$ is contiguous with respect to $\{N'_n\}_{n\in\mathbb{N}}$, and write $\{N_n\}_{n\in\mathbb{N}} \rhd \{N'_n\}_{n\in\mathbb{N}}$, if $\lim_{n\to\infty}\mathbb{P}[N_n \in \mathcal{E}_n] = 0$ implies $\lim_{n\to\infty}\mathbb{P}[N'_n \in \mathcal{E}_n] = 0$ for all sequences of sets of graphs $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$. If $\{N_n\}_{n\in\mathbb{N}} \rhd \{N'_n\}_{n\in\mathbb{N}}$ and $\{N'_n\}_{n\in\mathbb{N}} \rhd \{N_n\}_{n\in\mathbb{N}}$, we say that $\{N_n\}_{n\in\mathbb{N}}, \{N'_n\}_{n\in\mathbb{N}}$ are mutually contiguous, and write $\{N_n\}_{n\in\mathbb{N}} \lhd\rhd \{N'_n\}_{n\in\mathbb{N}}$.*

Theorem 2.0.3. *Suppose $\boldsymbol{d}$ is a Poisson random variable. Then $G_{n,m,\boldsymbol{k}} \lhd\rhd G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \lhd\rhd G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$, and $G_{n,\boldsymbol{m},\boldsymbol{k}} \lhd\rhd G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \lhd\rhd G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$.*

Theorem 2.0.4. *Let $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ be a sequence of sets of factor graphs. Then*

$$\lim_{n\to\infty}\left|\mathbb{P}[G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] - \mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n]\right| = 0,$$
$$\lim_{n\to\infty}\left|\mathbb{P}[G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] - \mathbb{P}[G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n]\right| = 0.$$

The proofs of Theorems 2.0.3 and 2.0.4 are deferred to Chapter 7. The results for models $G_{n,m,\boldsymbol{k}}$, $G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$, $G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ with deterministic numbers of constraint vertices are proven in Sections 7.1 and 7.2. The results for models $G_{n,\boldsymbol{m},\boldsymbol{k}}$, $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$, $G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ with random numbers of constraint vertices are proven in Section 7.3.

Chapter 3

Overview of the Proofs

3.1 Theorems 1.3.2, 1.3.5, 1.3.7, and 1.3.6

To prove Theorem 1.3.5, we use the first and the second moment methods to determine that the threshold at which a given tree T whose constraint vertices all have degree k appears in $F(\mathbf{B})$ is $\Theta(n^{1-1/c(T)})$, where $c(T)$ is the number of constraint vertices in T. Consequently, if $m \ll n^{\alpha_{P,Q}}$ then a.a.s. there will be no $T \in \mathcal{T}(P,Q)$ occuring in $F(\mathbf{B})$ and thus every component of $F(\mathbf{B})$ is a tree that is satisfiable for any assignment of entries of $(\mathbf{B}, b)$, and so the system (1.1) is satisfiable. On the other hand, if $m \gg n^{\alpha_{P,Q}}$ then there will be $\omega(1)$ disjoint copies of $T \in \mathcal{T}(P,Q)$ occurring in $F(\mathbf{B})$. Since the nonzero entries of $\mathbf{B}$ are assigned independently of $F(\mathbf{B})$, independently for each occurrence of $T \in \mathcal{T}(P,Q)$ in $F(\mathbf{B})$, there is a positive probability away from zero such that (T, w) is unsatisfiable, where (T, w) is the weighted factor graph of $\mathbf{B}x = b$ induced by the set of vertices in T. This implies that a.a.s. the system (1.1) is unsatisfiable.

By Theorem 1.3.5, it suffices to estimate $\inf\{c(T) : T \in \mathcal{T}(P,Q)\}$ to prove Theorem 1.3.6. We inductively define a pair of rooted trees $T_{\ell(P,Q)}$, $T^+_{\ell(P,Q)}$, both of height $2\ell(P,Q)+1$, such that any $T \in \mathcal{T}(P,Q)$ contains $T^+_{\ell(P,Q)}$ (Lemma 4.2.6), and $(T_{\ell(P,Q)}, w)$ is unsatisfiable for some weight function w (Lemma 4.2.2). Consequently, $c(T^+_{\ell(P,Q)}) \leq \inf\{c(T) : T \in \mathcal{T}(P,Q)\} \leq c(T_{\ell(P,Q)})$, which yields Theorem 1.3.6. The appropriate weight function w on $T_{\ell(P,Q)}$ is derived from Q and the sequence of ideals $I_0(P), I_1(P), \ldots, I_{\ell(P,Q)}(P)$.

Unfortunately, $T_{\ell(P,Q)}$ and $T^+_{\ell(P,Q)}$ are not the same size, and therefore do not give tight bounds on $\inf\{c(T) : T \in \mathcal{T}(P,Q)\}$ and $\alpha_{P,Q}$. However, for Theorem 1.3.7, it suffices to show that $\mathcal{T}(P,Q) \neq \emptyset$ if and only if $\ell(P,Q) = \infty$, which immediately follows from Theorem 1.3.6.

Finally, we prove Theorem 1.3.2 by finding the smallest obstruction in $\mathcal{T}(P,Q)$, which gives the SAT threshold by Theorem 1.3.5. We begin by defining a second inductive sequence of ideals $I^+_0(P), I^+_1(P), \ldots$ similar to $I_0(P), I_1(P), \ldots$ (Definition 4.2.3), which agrees with $I_0(P), I_1(P), \ldots$ when R is a principal ideal ring (Lemma 4.2.4). Letting $\ell^+(P,Q) = \inf\{\ell \geq 0 : I_\ell(P) \not\subseteq \operatorname{supp}(Q)\}$, we use the sequence $I_0(P), I_\ell(P) \ldots$ to derive a weight function w

on $T^+_{\ell^+(P,Q)}$ such that $(T^+_{\ell^+(P,Q)}, w)$ is unsatisfiable (Lemma 4.2.5). Since $\ell^+(P,Q) = \ell(P,Q)$ when R is principal, this shows that the bound on $\alpha_{P,Q}$ given by Lemma 4.2.6 is tight, and proves Theorem 1.3.2.

3.2 Theorem 1.3.3

We first prove Theorem 1.3.3 in a special case where R is a local ring, which is stated as Theorem 3.2.2 below. We then then give an overview of how to reduce Theorem 1.3.3 to this special case.

Definition 3.2.1. *A commutative ring R is local if it has a unique maximal ideal.*

Theorem 3.2.2. *Suppose P, Q are distributions on a local ring R, and let $\hat{P} = P^{\otimes k}$. Suppose $(\hat{P}, Q)$ satisfies* **(A2)** *and P is supported on the units of R. Then, for fixed $\varepsilon > 0$, we have*

$$\lim_{n\to\infty} P[\textit{System (1.1) is satisfiable}] = \begin{cases} 1 & m < (d_k - \varepsilon)n/k \\ 0 & m > (d_k + \varepsilon)n/k. \end{cases}$$

Proof. Let M be the unique maximal ideal of R. Let $\bar{\mathbf{B}}$ and $\bar{\mathbf{b}}$ be the matrix and vector obtained from $\mathbf{B}$ and $\mathbf{b}$ by taking entry-wise projections on to the residue filed R/M. By the assumption that P is supported on the units of R, every nonzero element of $\mathbf{B}$ is mapped to a nonzero field element in R/M by the projection. Let C and V denote the set of rows and columns of $\mathbf{B}$. Given $C' \subseteq C$ and $V' \subseteq V$, let $\mathbf{B}[C', V']$ denote the submatrix of $\mathbf{B}$ induced by rows in C' and coloumns in V'.

In the subcritical case $m < (d_k - \varepsilon)n/k$, by following the identical to the argument in [7], we can show that there exists a subset $V' \subseteq V$ with $|V'| = |C|$ such that $\bar{\mathbf{B}}[C, V']$ is invertible over R/M. Hence, $\det(\bar{\mathbf{B}}[C, V']) \neq 0$ in R/M, which implies that $\det(\mathbf{B}[C, V'])$ is a unit in R since R is local. Therefore, $\mathbf{B}[C, V']$ is invertible over R, and so $\mathbf{B}x = \mathbf{b}$ is a.a.s. satisfiable.

The proof for the supercritical case $m > (d_k + \varepsilon)n/k$ follows closely that of [7, Theorem 1.5], and so we only briefly outline the main ideas here. Start with a system $\mathbf{B}'x = \mathbf{b}'$ of $m' = (d_k - \varepsilon_n)n/k$ equations for some $\varepsilon_n \to 0$ sufficiently slowly, with rows in $\mathbf{B}'$ and entries in $\mathbf{b}'$ drawn from the same distribution as in $\mathbf{B}x = \mathbf{b}$. Then, consider the set of solutions Σ of the subsystem $\mathbf{B}'_{\text{2-core}}x = \mathbf{b}'_{\text{2-core}}$, where $\mathbf{B}'_{\text{2-core}}$ is the maximum submatrix of $\mathbf{B}'$ where every column contains at least 2 non-zero entries and every row contains k non-zero entries. We can show that $|\Sigma| = |R|^{O(\varepsilon_n n)}$ by the same argument as Claim 2.2 in [29], using the fact that any invertible submatrix of $\bar{\mathbf{B}}'$ (over R/M) is also an invertible submatrix of $\mathbf{B}'$ (over R). Then, we add $m - m' > \varepsilon n/k$ equations to $\mathbf{B}'x = \mathbf{b}'$ to get a system with the same distribution as $\mathbf{B}x = \mathbf{b}$. Let R' be the subset of equations that were added to $\mathbf{B}'$ which have non-zero entries only in the columns that are contained in the submatrix $\mathbf{B}'_{\text{2-core}}$. As argued in [29], we can show that a.a.s. $|R'| = \Omega(n)$. Using **(A2)** and the same argument as Claim 2.3 of [29], there is a $\delta > 0$ such that, given an $x \in \Sigma$, the probability that x satisfies each additional equation

in R' is at most $1-\delta$. Then we can show that the expected number of solutions in Σ that satisfy all the additional $m-m'$ equations is bounded by $q^{O(\varepsilon_n n)}(1-\delta)^{\Omega(\varepsilon n/k)} = \exp(-\Omega(n))$. This immediately implies that a.a.s. $\mathbf{B}x = \mathbf{b}$ is unsatisfiable. □

In the first step in proving both Theorem 1.3.3, we note that any finite ring R can be written as a product of local rings (Lemma 4.3.1), and therefore it suffices to prove Theorem 1.3.3 in the case where R is a local ring with a unique maximal ideal M. We show that $\mathrm{supp}(P)$ and $\mathrm{supp}(Q)$ have a very simple structure when $\alpha_{P,Q} = 1$ (Lemma 5.0.1). In particular, either every $r \in \mathrm{supp}(P)$ generates the same ideal, or $\mathrm{supp}(Q) = \{0\}$. Satisfiability is trivial if $\mathrm{supp}(Q) = \{0\}$. In the case where every $r \in \mathrm{supp}(P)$ generates the same ideal I, the quotient I/IM is a one dimensional vector space over the residue field R/M, with $r \notin IM$ for all $r \in \mathrm{supp}(P)$. This allows us to apply Theorem 3.2.2 to show that satisfiability threshold is $m \sim d_k n/k$.

3.3 Theorem 1.3.9

We show that system (1.1) is satisfiable if and only if a related system $\mathbf{B}_{XY}\, x = \mathbf{b}_{XY}$ is satisfiable, where $\mathbf{B}_{XY}$ is an $m \times 2n$ $\{0,1\}$-matrix over $\mathbb{F}_q$ whose distribution is contiguous to the one given as follows. Independently for each row a, let k_a^X be a random variable drawn from distribution $\mathrm{Bin}(k,\rho)$ and let $k_a^Y = k - k_a^X$. Then, uniformly choose k_a^X entries from the first n columns and k_a^Y entries from the last n columns. These k chosen entries form the set of nonzero entries of row a.

Let $d_1, \ldots, d_{2n}$ denote the number of nonzero entries in the columns of this random matrix. The distribution of $(d_1, \ldots, d_{2n})$ is approximately that of i.i.d. copies of the random variable $\boldsymbol{d}$ where with probability $1/2$ it is a Poisson variable with mean $\rho km/n$ and with probability $1/2$ it is a Poisson variable with mean $(1-\rho)km/n$. The threshold $d_{k,\rho}$ in this theorem corresponds exactly to the threshold of satisfiability of $\mathbf{B}x = \mathbf{b}$ if $\mathbf{B}$ was drawn from the model in [14] described in Section 2. That is, $(d_1, \ldots, d_{2n})$ is drawn from i.i.d. $\boldsymbol{d}$, and $\boldsymbol{m}$ is a Poisson variable with mean m, and $\boldsymbol{k} = k$.

However, the distribution of our random matrix is not exactly the model above; in particular, $d_1, \ldots, d_{2n}$ are not independent. The main challenge in proving the satisfiablity statement of Theorem 1.3.9 is to come up with a coupling scheme that allows us to compare our random linear equations with the model above with i.i.d. d_is and conclude that if the latter is satisfiable then our system is also satisfiable. The detailed description of the coupling scheme as well as the full proof of Theorem 1.3.9 is given in Section 6.

3.4 Dependency diagrams

In each of the following dependency diagrams, an arrow from statement A to statement B indicates that A is used in the proof of B.

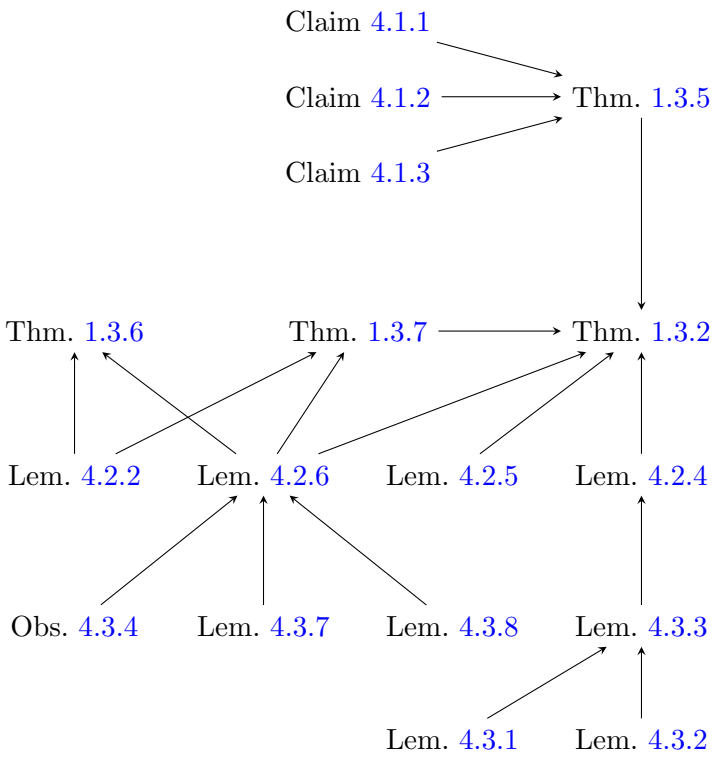


Figure 3.1: Dependency diagram for Theorems 1.3.2, 1.3.5, 1.3.6, and 1.3.7

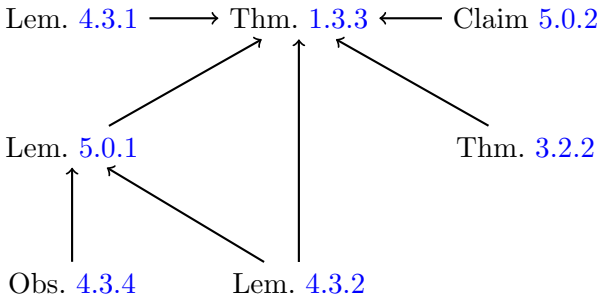


Figure 3.2: Dependency diagram for Theorem 1.3.3

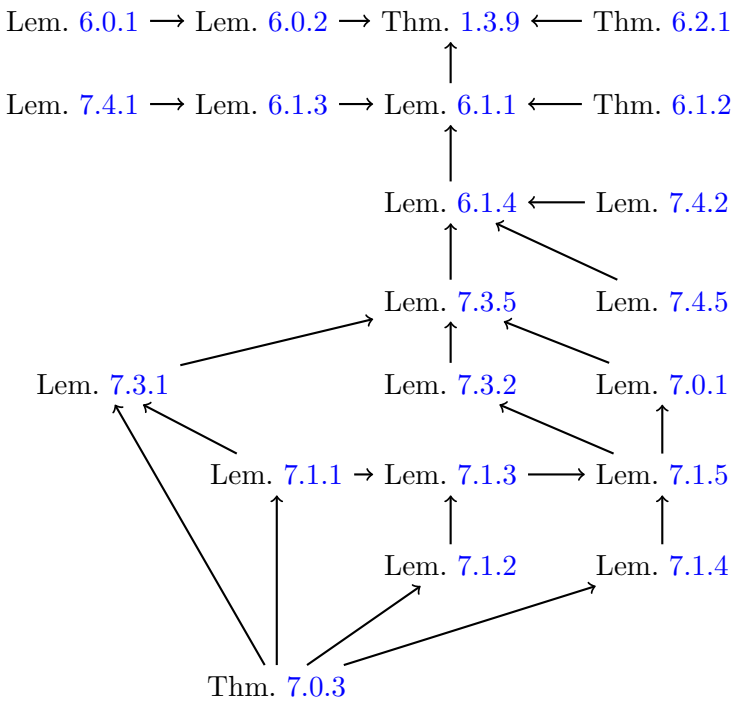


Figure 3.3: Dependency diagram for Theorem 1.3.9

Chapter 4

Sublinear SAT Thresholds

4.1 Proof of Theorem 1.3.5

Let R be a general finite commutative ring, let P and Q be distributions on R, and suppose $\hat{P} = P^{\otimes k}$, Q satisfy **(A1)** and **(A2)**. Recall from Section 1.3.2 that $\mathcal{T}(P, Q)$ is the set of unsatisfiable connected acyclic k-uniform weighted factor graphs whose edge weights are in $\text{supp}(P)$ and whose constraint vertex weights are in $\text{supp}(Q)$. Theorem 1.3.5 shows that the coarse satisfiability threshold for the system $\mathbf{B}x = \mathbf{b}$ is determined by the size of the smallest weighted graph in $\mathcal{T}(P, Q)$. In the following sections, we will refer to elements of $\mathcal{T}(P, Q)$ as *obstructions* for the system (1.1).

If the weighted factor graph $F(\mathbf{B}, \mathbf{b})$ contains an obstruction, then clearly $\mathbf{B}x = \mathbf{b}$ is unsatisfiable. Therefore, for each $T \in \mathcal{T}(P, Q)$, the satisfiability threshold is bounded above by the threshold for the existence of T as a subgraph of $F(\mathbf{B}, \mathbf{b})$. We begin by determining the existence threshold for any fixed unweighted tree T. The proof follows a standard first and second moment argument used in many other settings (see, for example, [10, Chapter 4]). For a factor graph F, let $c(F) = |C(F)|$, $v(F) = |V(F)|$, $e(F) = |E(F)|$ and let $a(F)$ be the number of automorphisms of F. Note that any acyclic k-uniform factor graph T satisfies $v(T) \geq (k-1)c(T) + 1$, and equality holds in the case where T is connected.

Claim 4.1.1. *Suppose T is a connected acyclic k-uniform factor graph, and let X_T be the number of isomorphic copies of T in $F(\boldsymbol{B})$. If $m \ll n^{1-1/c(T)}$, then $\mathbb{P}[X_T = 0] = 1 - o(1)$. If $m = \Theta(n^{1-1/c(T)})$, then $\mathbb{E}[X_T] = O(1)$ and $\mathbb{P}[X_T > 0] = \Omega(1)$. If $m \gg n^{1-1/c(T)}$, then $\mathbb{P}[X_T = \omega(1)] = 1 - o(1)$.*

Proof. Since T is k-uniform, connected, and acyclic, $v(T) = (k-1)c(T)+1$. Hence

$$\begin{aligned}\mathbb{E}[X_T] &= \frac{m!n!}{(m-c(T))!(n-v(T))!a(T)}\binom{n}{k}^{-c(T)} \\ &= \Theta(m^{c(T)}n^{v(T)-kc(T)}) \\ &= \Theta(m^{c(T)}n^{-c(T)+1}),\end{aligned} \tag{4.1}$$

where $\frac{m!n!}{(m-c(T))!(n-v(T))!a(T)}$ counts the number of ways that T can appear in the complete bipartite graph on $C_m \cup V_n$, and $\binom{n}{k}^{-c(T)}$ is the probability that each of these subgraphs appear in $F(\mathbf{B})$. Therefore $\mathbb{P}[X_T = 0] = 1 - o(1)$ if $m \ll n^{1-1/c(T)}$.

To prove the remaining statements of the claim, we calculate the second moment of X_T. Suppose $m = \Omega(n^{1-1/c(T)})$. Let $\mathcal{T}$ be the set of isomorphic copies of T in the complete bipartite graph on $C_m \cup V_n$, and let $\mathcal{H}$ be the set of k-uniform subgraphs of T. We can write the second moment of $\mathbb{E}[X_T]$ as

$$\begin{aligned}\mathbb{E}[X_T^2] &= \sum_{T_1,T_2\in\mathcal{T}} \mathbb{P}[T_1 \subseteq F(\mathbf{B}), T_2 \subseteq F(\mathbf{B})] \\ &= \sum_{H\in\mathcal{H}} \sum_{\substack{T_1,T_2\in\mathcal{T} \\ T_1\cap T_2\cong H}} \mathbb{P}[T_1 \subseteq F(\boldsymbol{B}), T_2 \subseteq F(\boldsymbol{B})].\end{aligned}$$

For $H \in \mathcal{H}$ with $c(H) = 0$, we have

$$\sum_{\substack{T_1,T_2\in\mathcal{T} \\ T_1\cap T_2\cong H}} \mathbb{P}[T_1 \subseteq F(\boldsymbol{B}), T_2 \subseteq F(\boldsymbol{B})] \le \sum_{T_1,T_2\in\mathcal{T}} \mathbb{P}[T_1 \subseteq F(\boldsymbol{B})]\mathbb{P}[T_2 \subseteq F(\boldsymbol{B})] = (\mathbb{E}[X_T])^2. \tag{4.2}$$

and for $H = T$,

$$\sum_{\substack{T_1,T_2\in\mathcal{T} \\ T_1\cap T_2\cong T}} \mathbb{P}[T_1 \subseteq F(\boldsymbol{B}), T_2 \subseteq F(\boldsymbol{B})] = \mathbb{E}[X_T]. \tag{4.3}$$

For any $H \in \mathcal{H}$ with $c(T) > c(H) > 0$,

$$\begin{aligned}\sum_{\substack{T_1,T_2\in\mathcal{T} \\ T_1\cap T_2\cong H}} \mathbb{P}[T_1 \subseteq F(\boldsymbol{B}), T_2 \subseteq F(\boldsymbol{B})] &= O(m^{2c(T)-c(H)}n^{2v(T)-v(H)})\binom{n}{k}^{-2c(T)+c(H)} \\ &= O(m^{2c(T)-c(H)}n^{2v(T)-v(H)-2kc(T)+kc(H)}).\end{aligned}$$

Since T is acyclic, H satisfies $v(H) \ge (k-1)c(H)+1$. With (4.1), this implies

$$\begin{aligned}O(m^{2c(T)-c(H)}n^{2v(T)-v(H)-2kc(T)+kc(H)}) &\le O(m^{2c(T)-c(H)}n^{-2c(T)+c(H)+1}) \\ &= (\mathbb{E}[X_T])^2 O(m^{-c(H)}n^{c(H)-1}),\end{aligned} \tag{4.4}$$

and since $m = \Omega(n^{1-1/c(T)})$, we have $m^{-c(H)}n^{c(H)-1} = O(n^{c(H)/c(T)-1}) = o(1)$. Combining (4.2), (4.3), and (4.4) gives

$$\mathbb{E}[X_T^2] \leq (\mathbb{E}[X_T])^2 + \mathbb{E}[X_T] + o(1)(\mathbb{E}[X_T])^2$$

and so, for $C > 0$, the Paley-Zygmund inequality shows that

$$\mathbb{P}[X_T \geq C\mathbb{E}[X_T]] \geq (1-C)^2\frac{(\mathbb{E}[X_T])^2}{\mathbb{E}[X_T^2]} = (1-C)^2\frac{(\mathbb{E}[X_T])^2}{(\mathbb{E}[X_T])^2 + \mathbb{E}[X_T]} - o(1).$$

For $m = \Theta(n^{1-1/c(T)})$, we have $\mathbb{E}[X] = \Theta(1)$ by (4.1), and hence $\mathbb{P}[X_T > 0] = \Omega(1)$. For $m \gg n^{1-1/c(T)}$, we have $\mathbb{E}[X_T] = \omega(1)$ by (4.1), and taking $C = 1/\sqrt{\mathbb{E}[X_T]}$ shows that $\mathbb{P}[X_T \geq \sqrt{\mathbb{E}[X_T]}] = 1 - o(1)$. □

The following two claims are where we use assumptions **(A1)** and **(A2)**. Claim 4.1.2 shows that the satisfiability threshold is not superlinear. Claim 4.1.3 is necessary in the proof of Theorem 1.3.5 to show that the satisfiability threshold is coarse in the sublinear regime.

Claim 4.1.2. *Suppose P and Q are distributions such that $(P^{\otimes k}, Q)$ satisfies* **(A2)**. *If $m \gg n$ then*

$$\lim_{n\to\infty} \mathbb{P}[\textit{System (1.1) is satisfiable}] = 0$$

Proof. Let

$$\gamma = \min_{r\in R} \mathbb{P}[\sum_{v\in V_n} \mathbf{B}_{a,v} r \neq \mathbf{b}_a].$$

Since the rows of $\boldsymbol{B}x = \boldsymbol{b}$ are i.i.d., γ does not depend on the row $a \in C_m$. Since $(P^{\otimes k}, Q)$ satisfy **(A2)** and therefore do not admit a constant solution, $\gamma > 0$. Now, given $x \in R^{V_n}$, let $r_x \in R$ be such that $|\{v \in V_n : x_v = r_x\}| \geq n/|R|$. Then x does not satisfy $(\mathbf{B}x)_a = b_a$ if $x_{u_{a,i}} = r_x$ for all $i \in [k]$, where $(u_{a,1}, \ldots, u_{a,k})$ are the positions of the non-zero entries of $\mathbf{B}$ in row a, as defined in (1.1). Since $\{u_{a,1}, \ldots, u_{a,k}\}$ is a uniformly chosen random subset of V_n of size k, we get

$$\mathbb{P}[(\mathbf{B}x)_a \neq b_a] \geq \gamma\mathbb{P}[\{u_{a,1}, \ldots, u_{a,k}\} \subseteq \{v \in V_n : x_v = r_x\}] \geq (1 - o(1))\gamma|R|^{-k}$$

Since the rows of $(\boldsymbol{B}, \boldsymbol{b})$ are independent, the expected number of solutions to $\boldsymbol{B}x = \boldsymbol{b}$ is therefore at most $(1-\gamma|R|^{-k}+o(1))^m|R|^n$. If $m \gg n$, we get $(1-\gamma|R|^{-k}+o(1))^m|R|^n = o(1)$, and therefore (1.1) is unsatisfiable a.a.s. □

Claim 4.1.3. *For any distributions P and Q that satisfy* **(A1)** *and any k-uniform acyclic factor graph F, there is a weight function w with edge weights in* $\mathrm{supp}(P)$ *and constraint-vertex weights in* $\mathrm{supp}(Q)$ *such that (F, w) is satisfiable.*

Proof. Suppose $(P^{\otimes k}, Q)$ satisfy **(A1)**, and let $r_1, \dots, r_k \in \mathrm{supp}(P)$ and $b \in \mathrm{supp}(Q)$ be such that b is in the generated ideal $(r_1, \dots, r_k)$. We prove the claim for connected F by induction on $|C(F)|$. The claim for disconnected F follows by taking components.

For $|C(F)| = 1$, assign weight b to the one constraint vertex in $C(F)$, and assign weights $r_1, \dots, r_k$ to its k incident edges. Since $b \in (r_1, \dots, r_k)$, (F, w) is satisfiable.

Now suppose $|C(F)| > 0$. Since F is a tree, there is an $a \in C(F)$ such that $N_F(a) = \{v_1, \dots, v_k\}$, where $\deg_F(v_1) \geq 2$ and $\deg_F(v_i) = 1$ for $i = 2, \dots, k$. Let F' be the factor graph obtained by removing a and $v_2, \dots v_k$ from F. By the induction hypothesis, there is weighting w' of F' such that (F', w) is satisfiable. Let $a' \in C(F')$ be a constraint vertex adjacent to v_1 in F', and let $v_2', \dots, v_k' \in V(F')$ be such that $N_{F'}(a') = \{v_1, v_2', \dots, v_k'\}$. Extend w' to a weighting w of F by setting $w(a) = w(a')$ and $w(av_1) = w(a'v_1)$ and $w(av_i) = w(av_i')$ for $i = 2, \dots, k$. Then any satisfying assignment x of (F', w) can be extended to a satisfying assignment of (F, w) by setting $x_{v_i} = x_{v_i'}$ for $i = 2, \dots, k$. Therefore (F, w) is satisfiable. □

Proof of Theorem 1.3.5. First consider the case where $\alpha_{P,Q} = 1$. If $m \ll n^{\alpha_{P,Q}}$, then the first moment method shows that $F(\mathbf{B})$ is acyclic. Since $\mathcal{T}(P, Q) = \emptyset$, system (1.1) is therefore satisfiable a.a.s. If $m \gg n$, then Claim 4.1.2 shows that (1.1) is unsatisfiable a.a.s.

Now suppose that $\alpha_{P,Q} < 1$, and let

$$c^* = \inf\{|C(T)| : T \in \mathcal{T}(P, Q)\}, \quad \mathcal{T}^*(P, Q) = \{(T, w) \in \mathcal{T}(P, Q) : c(T) = c^*\}.$$

Suppose $m = O(n^{\alpha_{P,Q}})$. Then Claim 4.1.1 shows that $F(\mathbf{B})$ does not contain any tree with more than c^* constraint vertices a.a.s., and there is an $M > 0$ such that

$$\mathbb{P}[|\{T \subseteq F(\mathbf{B}) : c(T) = c^*\}| \leq M] = \Omega(1).$$

By Claim 4.1.3 each unweighted tree in $\mathcal{T}^*(P, Q)$ is satisfiable for some assignment of edge weights in $\mathrm{supp}(P)$ and constraint vertex weights in $\mathrm{supp}(Q)$. Therefore, conditioned on the event $\{|\{T \subseteq F(\mathbf{B}) : c(T) = c^*\}| \leq M\}$ the probability that all weighted trees in $\{(T, w) \subseteq F(\mathbf{B}, \mathbf{b}) : c(T) = c^*\}$ are satisfiable is $\Omega(1)$. Therefore (1.1) is satisfiable with probability $\Omega(1)$. Moreover, if $m \ll n^{\alpha_{P,Q}}$, then Claim 4.1.1 shows that $F(\mathbf{B})$ does not contain any trees with c^* constraint vertices a.a.s., and therefore (1.1) is satisfiable a.a.s.

Now suppose $m = \Omega(n^{\alpha_{P,Q}})$. Then Claim 4.1.1 shows that $F(\mathbf{B})$ contains a copy of T for some $(T, w) \in \mathcal{T}^*(P, Q)$ with probability $\Omega(1)$. Conditioned on $F(\mathbf{B})$, each copy of T in $F(\mathbf{B})$ is assigned weights w in $F(\mathbf{B}, \mathbf{b})$ with probability $\Omega(1)$. Therefore $F(\mathbf{B}, \mathbf{b})$ contains (T, w) with probability $\Omega(1)$, and so 1.1 is unsatisfiable with probability $\Omega(1)$. If $m \gg n$, then Claim 4.1.1 shows that the number of unweighted trees in $\mathcal{T}^*(P, Q)$ contained in $F(\boldsymbol{B})$ is unbounded a.a.s. Since the assignment of weights to each of these trees is independent, we get that (1.1) is unsatisfiable a.a.s. □

4.2 Proofs of Theorems 1.3.2, 1.3.6, and 1.3.7

By Theorem 1.3.5, it suffices to show that if R is principal and $\mathcal{T}(P,Q) \neq \emptyset$ then

$$\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\} = 1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}. \tag{4.5}$$

Suppose that $T_{\min}$ is a weighted tree factor graph in $\mathcal{T}(P,Q)$ which attains the infimum of the left hand side of (4.5). We seek tight upper and lower bounds for $|C(T_{\min})|$ by identifying tree factors that are members of $\mathcal{T}(P,Q)$. This motivates the sequences of unweighted tree factor graphs $\{T_i\}_{i\in\mathbb{N}}$ and $\{T_i^+\}_{i\in\mathbb{N}}$ defined as follows.

Definition 4.2.1. *We define two sequences of k-uniform factor graphs $T_0, T_1, \ldots$ and $T_0^+, T_1^+, \ldots$ inductively as follows. Let T_0 and T_0^+ be the graph with one constraint vertex and k adjacent variable vertices. Let $p = |\operatorname{supp} P|$. Let $L(T)$ denote the set of leaves of a tree T.*

For $\ell \geq 1$, construct T_ℓ from $T_{\ell-1}$ by adding p^2 constraint vertices $a_{v,1}, \ldots, a_{v,p^2}$ adjacent to each leaf $v \in L(T_{\ell-1})$, and adding $k-1$ new variable vertices $u_{a_{v,i},1}, \ldots, u_{a_{v,i},k-1}$ adjacent to each $a_{v,i}$ for $v \in L(T_{\ell-1}), i \in [p^2]$.

For $\ell \geq 1$, construct T_ℓ^+ from $T_{\ell-1}^+$ by adding a constraint vertex a_v adjacent to each leaf $v \in L(T_{\ell-1}^+)$, and adding $k-1$ new variable vertices $u_{a_v,1}, \ldots, u_{a_v,k-1}$ adjacent to a_v for each $v \in L(T_{\ell-1}^+)$.

As an illustration, $T_0 = T_0^+$ is given in Figure 4.1 below for the case $k = 3$ (ignoring the weight b on the constraint vertices, drawn as a square, and the weights r on the edges. The variable vertices are drawn as small disks). Figures 4.2 and 4.3 illustrate how to obtain T_1^+ and T_1 from $T_0^+ = T_0$ for the case $k = 3$ and $p = 2$ (again ignoring the weights).

We show in Lemma 4.2.2 that there is an assignment of weights that makes $T_{\ell(P,Q)}$ unsatisfiable. We briefly illustrate how to assign weights to $T_{\ell(P,Q)}$ in the cases $\ell(P,Q) = 0$ and $\ell(P,Q) = 1$.

Suppose that $\ell(P,Q) = 0$. Then $\operatorname{supp} Q \not\subseteq I_0(P)$, and there exists $(b,r) \in (\operatorname{supp} Q, \operatorname{supp} P)$ such that $b \notin (r)$. Assign weights to T_0 as demonstrated in Figure 4.1. Obviously, this weighted T_0 is unsatisfiable. Slightly less trivially, suppose $\ell(P,Q) = 1$. Then $\operatorname{supp}(Q) \subseteq I_0(P)$, but $\operatorname{supp}(Q) \not\subseteq I_1(P)$. Since $\operatorname{supp}(Q) \subseteq I_0(P)$, we have $b \in (r)$ for all $b \in \operatorname{supp}(Q)$ and $r \in \operatorname{supp}(P)$, and so T_0 is satisfiable for all choices of weights. However, since $\operatorname{supp}(Q) \not\subseteq I_1(P)$, there is a $b_1 \in \operatorname{supp}(Q)$ such that

$$b_1 \notin I_1(P) = \bigcap_{r\in\operatorname{supp}(P)} r \bigcap_{s\in\operatorname{supp}(P)} \{t \in R : st \in I_0(P)\}.$$

Therefore there is an $r_1 \in \operatorname{supp}(P)$ such that

$$b_1 \notin r_1 \bigcap_{s\in\operatorname{supp}(P)} \{t \in R : st \in I_0(P)\} = r_1 \bigcap_{s,r\in\operatorname{supp}(P)} \{t \in R : st \in (r)\}. \tag{4.6}$$

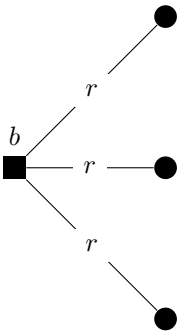


Figure 4.1: $T_0 = T_0^+$

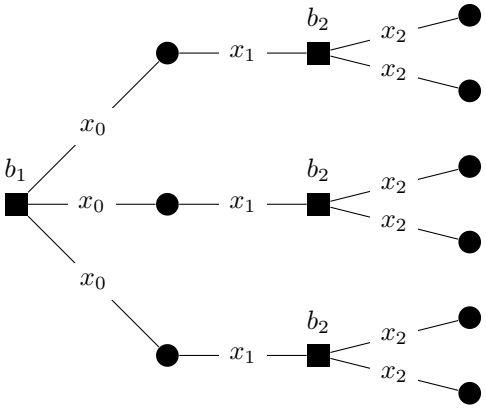


Figure 4.2: T_1^+

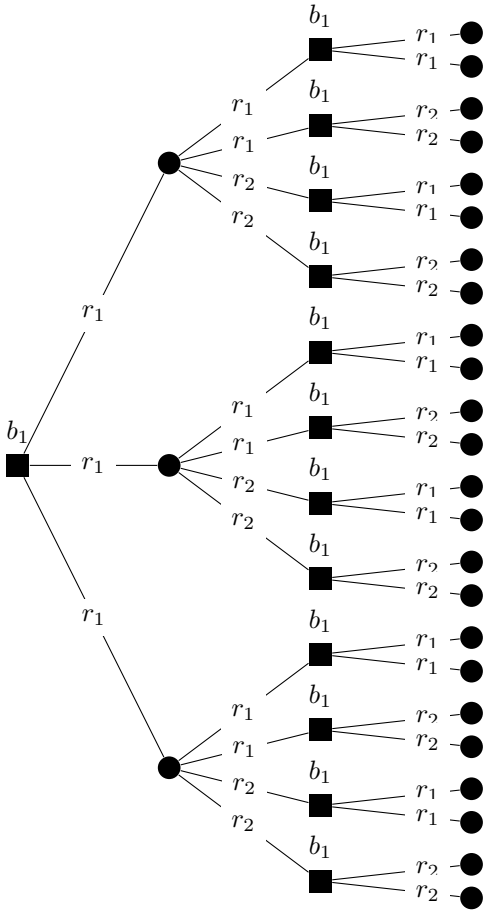


Figure 4.3: T_1

We assign weights to T_1 by first assigning weights b_1 and r_1 to the constraint vertex and edges of T_0, as shown in 4.3. Then, for each leaf v of T_0, we assign a distinct pair of weights (s, r) to each of the p^2 branches added to v in the construction of T_1, such that the edge incident to v gets weight s and all other edges on the branch get weight r. The kp^2 additional constraint in T_1 are assigned weight b_1. Figure 4.3 shows the complete weighted graph T_1 with $\text{supp}(P) = \{r_1, r_2\}$ and with the added constraint vertices assigned weight b_1. To see that this weighted graph in an obstruction for (1.1), suppose a variable assignment x satisfies all kp^2 of the additional constraints in T_1 (the twelve constraints on the right in figure 4.3). Since $b_1 \in \bigcap_{r \in \text{supp}(P)}(r)$, x must assign a value in $\bigcap_{s,r \in \text{supp}(P)}\{t \in R : st \in (r)\}$ to each of the k variables in T_0 (the three variables in the middle of Figure 4.3). Since b_1 satisfies (4.6), this implies that x does not satisfy the one constraint in T_0 (the left most constraint in Figure 4.3).

We will generalize the obstruction described above to yield the following lemma, whose proof is delayed to Section 4.3.

Lemma 4.2.2. *Suppose $\ell(P,Q) < \infty$. Then there is an assignment of weights*

$$w : (E(T_{\ell(P,Q)}), V(T_{\ell(P,Q)})) \to (\mathrm{supp}(P), \mathrm{supp}(Q))$$

such that the factor graph $(T_{\ell(P,Q)}, w)$ is unsatisfiable.

This shows that $\mathcal{T}(P,Q) \neq \emptyset$ when $\ell(P,Q) < \infty$, and therefore gives one direction of Theorem 1.3.7. However, obstructions of this form do not generally give a tight bound on the infimum in (4.5). For example, consider a system of equations over $\mathbb{Z}_4$ with $\mathrm{supp}(P) = \{1,2\}$ and $\mathrm{supp}(Q) = \{0,2\}$. We then have $I_0(P) = (1) \cap (2) = (2)$, and

$$I_1(P) = \bigcap_{r \in \{1,2\}} r \cdot (((2) : (2)) \cap ((2) : (1))) = 0$$

Thus $\ell(P,Q) = 1$, and Lemma 4.2.2 shows that there is a weighting of T_1 that is unsatisfiable. However, we can also construct an unsatisfiable weighting of the smaller graph T_1^+ depicted in Figure 4.2: in Figure 4.2, assign $b_1 = b_2 = 0$ and $x_0 = 2$, $x_1 = 1$, $x_2 = 2$. Then any satisfying assignment of the three variable vertices on the left of Figure 4.2 must be in the ideal (2). But now $b_1 \notin x_0 \cdot (2) = 0$, so the left-most constraint vertex cannot be satisfied. Therefore this weighting of T_1^+ forms an obstruction that is smaller than the one provided by Lemma 4.2.2.

We will generalize the example above to construct weight functions w on T_ℓ^+ such that (T_ℓ^+, w) is unsatisfiable. We start by defining a sequence of ideals $I_\ell^+(P)$ in a similar manner as $I_\ell(P)$ in Definition 1.3.1. The weights we will assign to T_ℓ^+ are derived from $I_\ell^+(P)$ similarly to how the weights of T_ℓ are derived from $I_\ell(P)$ in Lemma 4.2.2. The difference between the sequences $I_\ell(P)$ and $I_\ell^+(P)$ comes from the structure of T_ℓ and T_ℓ^+: $I_\ell(P)$ is the largest ideal such that T_ℓ is satisfiable for any assignment of edge weights in $\mathrm{supp}(P)$ and constraint vertex weights in $I_\ell(P)$; the sequence $I_\ell^+(P)$ is defined so that a similar statement holds for T_ℓ^+. These two sequences of ideals are in general different–see an illustration in Example 4.2.7 below–but they agree when R is a principal ideal ring.

Definition 4.2.3. *Given $x \in R^{\mathbb{N}}$, define a sequence of ideals $I_0^+(x), I_1^+(x), \ldots$ inductively by setting $I_0^+(x) = (x_0)$ and*

$$I_\ell^+(x) = \{x_{2\ell} t : x_{2\ell-1} t \in I_{\ell-1}^+(x), t \in R\}$$

for each positive integer ℓ. Define

$$I_\ell^+(P) = \bigcap_{x \in \mathrm{supp}(P)^{\mathbb{N}}} I_\ell^+(x)$$

and

$$\ell^+(P,Q) = \inf\{\ell \geq 0 : \mathrm{supp}(Q) \not\subseteq I_\ell^+(P)\},$$

with the convention that $\inf(\emptyset) = \infty$.

The proofs of the following three lemmas will be deferred to Section 4.3. The first lemma relates $I_\ell^+(P)$ to $I_\ell(P)$, and establishes that $I_\ell(P) = I_\ell^+(P)$ when R is a principal ideal ring.

Lemma 4.2.4. *For any ring R and distribution P on R, $I_\ell(P) \subseteq I_\ell^+(P)$ for all $\ell \geq 0$. If R is a principal ring, then $I_\ell(P) = I_\ell^+(P)$ for all $\ell \geq 0$.*

The lemma below, an analog of Lemma 4.2.2, shows that $(T^+_{\ell^+(P,Q)}, w)$ is an obstruction, and thus serves to establish another upper bound for the left hand side of (4.5).

Lemma 4.2.5. *Suppose $\ell^+(P,Q) < \infty$. Then there is an assignment of weights*

$$w : (E(T^+_{\ell^+(P,Q)}), V(T^+_{\ell^+(P,Q)})) \to (\operatorname{supp}(P), \operatorname{supp}(Q))$$

such that the factor graph $(T^+_{\ell^+(P,Q)}, w)$ is unsatisfiable.

The lemma below allows us to establish the lower bound for the left hand side of (4.5) that we desire.

Lemma 4.2.6. *Suppose F is an acyclic k-uniform factor graph, and w is a weight function on F with edge weights in $\operatorname{supp}(P)$ and constraint-vertex weights in $\operatorname{supp}(Q)$. Suppose that either $\ell(P,Q) = \infty$, or that $\ell(P,Q) < \infty$ and F does not contain $T^+_{\ell(P,Q)}$ as a subgraph. Then (F,w) is satisfiable.*

Note that Lemmas 4.2.6 and 4.2.5 hold for general finite commutative rings. They establish bounds on the left side of (4.5) for general rings, but these bounds agree only if $\ell(P,Q) = \ell^+(P,Q)$, which is guaranteed to hold for principal ideal rings by Lemma 4.2.4.

We complete the proof of Theorems 1.3.7, 1.3.6, and 1.3.2 assuming Lemmas 4.2.2, 4.2.4, 4.2.5, and 4.2.6, whose proof is delayed to the next section.

Proof of Theorem 1.3.7. By definition of $\alpha_{P,Q}$, we have $\alpha_{P,Q} = 1$ if and only if $\mathcal{T}(P,Q) = \emptyset$. If $\ell(P,Q) = \infty$, then $\mathcal{T}(P,Q) = \emptyset$ by Lemma 4.2.6. Conversely, if $\ell(P,Q) < \infty$, then $\mathcal{T}(P,Q) \neq \emptyset$ by Lemma 4.2.2. Therefore $\mathcal{T}(P,Q) = \emptyset$ if and only if $\ell(P,Q) = \infty$. □

Proof of Theorem 1.3.6. Let $p = |\operatorname{supp}(P)|$ and suppose $\ell(P,Q) < \infty$. For any $T \in \mathcal{T}(P,Q)$ Lemma 4.2.6 shows that T contains $T^+_{\ell(P,Q)}$ as a subgraph. Thus

$$\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\} \geq |C(T^+_{\ell(P,Q)})|$$

By definition of $T^+_{\ell(P,Q)}$, we have

$$|C(T^+_{\ell(P,Q)})| = 1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1},$$

and hence

$$\alpha_{P,Q} = 1 - \frac{1}{\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\}} \geq \frac{\sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}}{1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}}.$$

This proves the lower bound in Theorem 1.3.6. For the upper bound in Theorem 1.3.6, Lemma 4.2.2 shows that $(T_\ell(P,Q), w) \in \mathcal{T}(P,Q)$ for some weight function w. Hence

$$\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\} \leq |C(T_{\ell(P,Q)})|.$$

By definition of $T_{\ell(P,Q)}$, we have

$$|C(T_{\ell(P,Q)})| = 1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}p^{2i},$$

and hence

$$\alpha_{P,Q} \leq \frac{\sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}p^{2i}}{1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}p^{2i}}.$$

□

Proof of Theorem 1.3.2. The case $\ell(P,Q) = \infty$ follows immediately from Theorems 1.3.5 and 1.3.7. Suppose that $\ell(P,Q) < \infty$. By Theorem 1.3.7, $\mathcal{T}(P,Q) \neq \emptyset$. Since R is principal, Lemma 4.2.4 shows that $\ell^+(P,Q) = \ell(P,Q) < \infty$. By the definition of $\{T_\ell^+\}_{\ell \geq 0}$, we have

$$|C(T^+_{\ell(P,Q)})| = 1 + \sum_{i=1}^{\ell(P,Q)} k(k-1)^{i-1}.$$

In light of (4.5), it suffices to show that

$$\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\} = |C(T^+_{\ell(P,Q)})|.$$

By Lemma 4.2.5, there is a weight function w such that $(T^+_{\ell(P,Q)}, w)$ is an obstruction to (1.1), which implies that $\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\} \leq |C(T^+_{\ell(P,Q)})|$. For the reverse inequality, suppose that $(T,w) \in \mathcal{T}(P,Q)$. By Lemma 4.2.6, T contains $T^+_{\ell(P,Q)}$ as a subgraph, and in particular $|C(T)| \geq |C(T^+_{\ell(P,Q)})|$. Therefore $\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\} \geq |C(T^+_{\ell(P,Q)})|$. This completes the verification of (4.5) and thus confirms Theorem 1.3.2. □

We close this section with an example that demonstrates that $I_\ell(P)$ is in general not equal to $I_\ell^+(P)$, and that the bounds on $\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\}$ given by Lemmas 4.2.5, 4.2.2, and 4.2.6 are not always tight for non-principal ideal rings.

Example 4.2.7. *Let* $R = \mathbb{F}_q[X,Y,Z]/(XY - ZY, XY - XZ, X^3, Y^3, Z^3)$. *Let* P *and* Q *be such that* $\operatorname{supp}(P) = \{\bar{X}, \bar{Y}, \bar{Z}\}$ *and* $\operatorname{supp}(Q) = (\bar{X}\bar{Y})$. *Then*

$$I_0(P) = (\bar{X}\bar{Y}), \quad I_\ell(P) = (\bar{X}\bar{Y}\bar{Z}), \quad \ell \geq 1,$$
$$I_\ell^+(P) = (\bar{X}\bar{Y}), \quad \ell \geq 0.$$

Therefore $\ell(P,Q) = 1$ *and* $\ell^+(P,Q) = \infty$.

Proof. We have

$$I_0(P) = (\bar{X}) \cap (\bar{Y}) \cap (\bar{Z}) = (\bar{X}\bar{Y})$$

and

$$I_1(P) = (\bar{X}A_1(P)) \cap (\bar{Y}A_1(P)) \cap (\bar{Z}A_1(P))$$

where

$$A_1(P) = \{t \in R : \bar{X}t, \bar{Y}t, \bar{Z}t \in (\bar{X}\bar{Y})\} = (\bar{X}\bar{Y}, \bar{X}^2, \bar{Y}^2, \bar{Z}^2).$$

Therefore $I_1(P) = (\bar{X}\bar{Y}\bar{Z})$. For $I_2(P)$, we have

$$I_2(P) = (\bar{X}A_2(P)) \cap (\bar{Y}A_2(P)) \cap (\bar{Z}A_2(P)),$$

where

$$A_2(P) = \{t \in R : \bar{X}t, \bar{Y}t, \bar{Z}t \in (\bar{X}\bar{Y}\bar{Z})\} = (\bar{X}\bar{Y}, \bar{X}^2, \bar{Y}^2, \bar{Z}^2)$$

Therefore $I_2(P) = (\bar{X}\bar{Y}\bar{Z})$. Since $I_2(P) = I_1(P) = (\bar{X}\bar{Y}\bar{Z})$, we have $I_\ell(P) = (\bar{X}\bar{Y}\bar{Z})$ for all $\ell \geq 1$.

Now consider $I_\ell^+(P)$. Given $x \in \operatorname{supp}(P)^{\mathbb{N}}$, we show that $I_\ell^+(x) \supseteq (\bar{X}\bar{Y})$ by induction on $\ell \geq 0$. For $\ell = 0$, we have $I_0^+(x) = (x_0) \supseteq (\bar{X}\bar{Y})$. Supposing that $\ell > 0$ and $I_{\ell-1}(x) \supseteq (\bar{X}\bar{Y})$, we have

$$I_\ell^+(x) \supseteq \{x_{2\ell}t : x_{2\ell-1}t \in (\bar{X}\bar{Y})\}.$$

Since $\operatorname{supp}(P) = \{\bar{X}, \bar{Y}, \bar{Z}\}$, there is a $y \in \operatorname{supp}(P) \setminus \{x_{2,\ell}, x_{2\ell-1}\}$. Then $x_{2\ell-1}y \in (\bar{X}\bar{Y})$, so $I_\ell^+(x) \supseteq (x_{2\ell}y) = (\bar{X}\bar{Y})$. This completes the induction step. Taking an intersection over $x \in \operatorname{supp}(P)^{\mathbb{N}}$ then gives $I_\ell^+(P) \supseteq (\bar{X}\bar{Y})$ for all $\ell \geq 0$. For the reverse inclusion, we have $I_0^+(P) = (\bar{X}) \cap (\bar{Y}) \cap (\bar{Z}) = (\bar{X}\bar{Y})$, so $I_\ell^+(P) \subseteq (\bar{X}\bar{Y})$ for all $\ell \geq 0$ by Observation 4.3.4 in the following section. Therefore $I_\ell^+(P) = (\bar{X}\bar{Y})$ for $\ell \geq 0$. □

Since $\ell(P,Q) \neq \ell^+(P,Q)$ in Example 4.2.7, Lemmas 4.2.5 and 4.2.6 do not give tight bounds on $\inf\{|C(T)| : T \in \mathcal{T}(P,Q)\}$. In fact, the smallest obstruction in the case $k = 3$ is given in Figure 4.4, and it is not of the form (T_ℓ, w) or (T_ℓ^+, w) for any $\ell \geq 0$.

For $k \geq 3$ the graph in Figure 4.4 can be generalized to an obstruction with $2k+1$ constraint vertices. Thus, by the definition of $\alpha_{P,Q}$ in Theorem 1.3.5,

$$\alpha_{P,Q} = \frac{2k}{2k+1}.$$

This shows that neither bound in Theorem 1.3.6 is tight, as Theorem 1.3.6 gives

$$\frac{k}{k+1} \leq \alpha_{P,Q} \leq \frac{9k}{9k+1}.$$

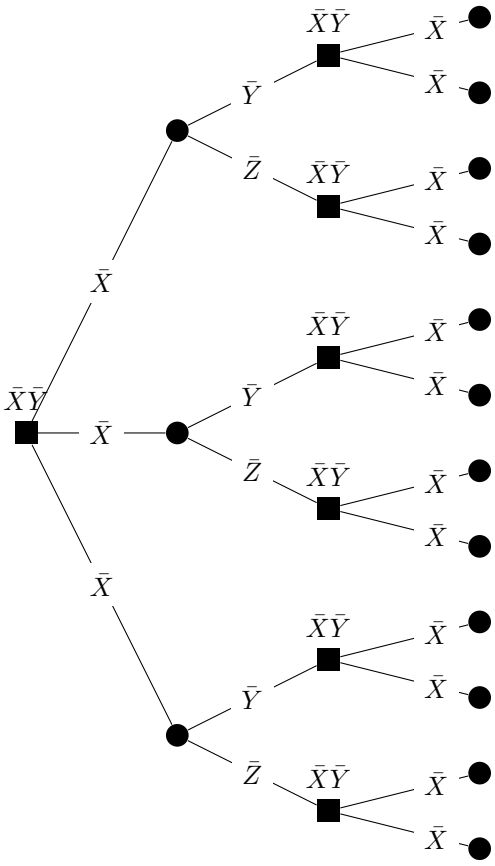


Figure 4.4: A smallest obstruction for (1.1) in Example 4.2.7

4.3 Proofs of Lemmas 4.2.2, 4.2.6, 4.2.5, and 4.2.4

We begin by proving Lemma 4.2.4 and some related facts about the sequences $\{I_\ell(P)\}_{\ell\geq 0}$ and $\{I_\ell^+(P)\}_{\ell\geq 0}$. We will need the following basic facts about finite commutative rings, which will also be useful in Chapter 5. Recall from Definition 3.2.1 that a local ring is a ring with a unique maximal ideal.

Lemma 4.3.1. *Let R be a finite commutative ring. Then there are finite commutative local rings $R_1, \ldots, R_h$ such that*

$$R \cong R_1 \times R_2 \times \cdots \times R_h.$$

If R is principal ideal ring then each R_j is principal ideal ring.

Lemma 4.3.2. *Let R be a finite commutative local principal ideal ring with maximal ideal M. Then every ideal of R is of the form M^i for some $i \geq 0$.*

Proofs of Lemmas 4.3.2 and 4.3.1 can be found in [9]. The following key fact about principal rings allows us to prove the equality in Lemma 4.2.4.

Lemma 4.3.3. *Let $\{I_a\}_{a\in C}$ be a collection of ideals of a finite commutative ring R, and let $r \in R$. Then*

$$\begin{aligned} r\bigcap_{a\in C} I_a &\subseteq \bigcap_{a\in C} rI_a, \\ r\bigcap_{a\in C} I_a &= \bigcap_{a\in C} rI_a, \quad \text{if } R \text{ is a principal ideal ring.} \end{aligned}$$

Proof. If $s \in \bigcap_{a\in C} I_a$, then $rs \in rI_a$ for all $a \in C$. Therefore the first assertion holds.

For the second assertion, by Lemma 4.3.1 it suffices to consider local principal rings. Suppose R is a local ring with maximal ideal M. By Lemma 4.3.2, each I_a is of the form M^i for some $i \geq 0$. Therefore there is a smallest ideal $I^* \in \{I_a\}_{a\in C}$ by inclusion, and $\bigcap_{a\in C} I_i = I^*$. We also have $rI^* \subseteq rI_a$ for all $a \in C$, and so $\bigcap_{a\in C} rI_a = rI^*$. Therefore $r\bigcap_{a\in C} I_a = rI^* = \bigcap_{a\in C} rI_a$. □

Proof of Lemma 4.2.4. Both claims can be proved by induction on $\ell \geq 0$. We first show that $I_\ell(P) \subseteq I_\ell^+(P)$ for $\ell \geq 0$ without assuming that R is principal. For $\ell = 0$, we have

$$I_0(P) = \bigcap_{r\in \mathrm{supp}(P)} rR = \bigcap_{x\in \mathrm{supp}(P)^{\mathbb{N}}} I_0^+(x) = I_0^+(P).$$

Suppose $\ell > 0$ and that the claim holds for $\ell - 1$. We have

$$\begin{aligned} \bigcap_{x\in \mathrm{supp}(P)^{\mathbb{N}}} I_\ell^+(x) &= \bigcap_{x\in \mathrm{supp}(P)^{\mathbb{N}}} \{x_{2\ell}t : x_{2\ell-1}t \in I_{\ell-1}^+(x), t \in R\} \\ &= \bigcap_{r\in \mathrm{supp}(P)} \bigcap_{s\in \mathrm{supp}(P)} \bigcap_{x\in \mathrm{supp}(P)^{\mathbb{N}}} \{rt : st \in I_{\ell-1}^+(x), t \in R\} \\ &\supseteq \bigcap_{r\in \mathrm{supp}(P)} r\left(\bigcap_{s\in \mathrm{supp}(P)} \bigcap_{x\in \mathrm{supp}(P)^{\mathbb{N}}} \{t \in R : st \in I_{\ell-1}^+(x)\} \right) \qquad (4.7) \\ &= \bigcap_{r\in \mathrm{supp}(P)} r\left(\bigcap_{s\in \mathrm{supp}(P)} \{t \in R : st \in I_{\ell-1}^+(P)\} \right) \qquad (4.8) \end{aligned}$$

Using the induction hypothesis,

$$\bigcap_{r\in \mathrm{supp}(P)} r\left(\bigcap_{s\in \mathrm{supp}(P)} \{t \in R : st \in I_{\ell-1}^+(P)\} \right) \supseteq \bigcap_{r\in \mathrm{supp}(P)} r\left(\bigcap_{s\in \mathrm{supp}(P)} \{t \in R : st \in I_{\ell-1}(P)\} \right) \qquad (4.9)$$

$$= I_\ell(P).$$

Therefore $I_\ell(P) \subseteq I_\ell^+(P)$ for all $\ell \geq 0$. If R is a principal ideal ring, then by Lemma 4.3.3, (4.7) holds with equality. Consequently, $I_\ell(P) = I_\ell^+(P)$ for $\ell \geq 0$ follows by induction, with (4.9) assumed as an equality in the inductive hypothesis. □

The following further observation about $\{I_\ell(P)\}_{\ell\geq 0}$ and $\{I^+_\ell(P)\}_{\ell\geq 0}$ will be useful in the proofs of Lemmas 4.2.6, 4.2.5, and 4.2.2.

Observation 4.3.4. *Let P be a distribution on a finite commutative ring R. Then for all $\ell \geq 0$, $I_\ell(P) \supseteq I_{\ell+1}(P)$, and $I^+_\ell(P) \supseteq I^+_{\ell+1}(P)$.*

Proof. Let $\ell \geq 0$. For any $r \in \mathrm{supp}(P)$, we have

$$I_{\ell+1}(P) \subseteq \{rt : rt \in I_\ell(P)\} \subseteq I_\ell(P).$$

On the other hand,

$$I^+_{\ell+1}(P) = \bigcap_{x\in \mathrm{supp}(P)^{\mathbb{N}}} I^+_{\ell+1}(x) = \bigcap_{x\in \mathrm{supp}(P)^{\mathbb{N}}} \{x_{2\ell+2}t : x_{2\ell+1}t \in I^+_\ell(x), t \in R\} \subseteq I^+_\ell(P).$$

□

We now introduce a notion of height for an acyclic factor graph F, and show that it corresponds to the largest ℓ such that T^+_ℓ is a subgraph of F.

Definition 4.3.5. *Let F be a k-uniform factor graph. A 2-subtree of F is a connected acyclic subgraph of F in which every constraint vertex has degree k and every variable vertex has degree at most 2. Let $\mathcal{T}_2(F)$ denote the set of 2-subtrees of F.*

Recall that for a tree T, $L(T)$ denotes the set of leaves of T.

Definition 4.3.6. *Let F be an acyclic factor graph. For $T \in \mathcal{T}_2(F)$ and $a \in C(T)$, define*

$$h_T(a) = \min\{(d_T(a, v) - 1)/2 : v \in L(T)\},$$

where d_T is the usual graph distance metric in the subgraph T. For $a \in C(F)$, define

$$\bar{h}_F(a) = \max\{h_T(a) : T \in \mathcal{T}_2(F)\}.$$

Lemma 4.3.7. *Let F be a connected acyclic factor graph and let $\ell^* \geq 0$. If F does not contain $T^+_{\ell^*}$, then $\bar{h}_F(a) < \ell^*$ for all $a \in C(F)$.*

Proof. We prove the contrapositive. Suppose there is an $a \in C(F)$ such that $\bar{h}_F(a) \geq \ell^*$. Let $T^* \in \mathcal{T}_2(F)$ be such that $h_{T^*}(a) = \bar{h}_F(a)$, and for each $0 \leq \ell \leq \ell^*$ let T^*_ℓ be the subgraph of T^* induced by the vertices $\{u \in V(T^*) \cup C(T^*) : d_{T^*}(a, u) \leq 2\ell + 1\}$. We show that $T^*_\ell \cong T^+_\ell$ for all $0 \leq \ell \leq \ell^*$ by induction on ℓ. Note that $T^*_{\ell^*} \cong T^+_{\ell^*}$ implies the contrapositive of the lemma.

The graph T^*_0 is the subgraph induced by a and its k adjacent variable vertices, and so trivially $T^*_0 \cong T^+_0$.

Suppose $1 \leq \ell \leq \ell^*$ and that $T^*_{\ell-1} \cong T^+_{\ell-1}$. Then by definition of $T^+_{\ell-1}$, $d_{T^*}(a, u) = 2(\ell - 1) + 1$ for all $u \in L(T^*_{\ell-1})$. If there is a $u \in L(T^*) \cap L(T^*_{\ell-1})$ (i.e. if some leaf in $T^*_{\ell-1}$

happens to be a leaf in T^*), then $h_{T^*}(a) \leq \ell - 1 < \bar{h}_F(a)$, which contradicts the choice of T^*. Therefore, since T^* is a 2-subtree of F, all $u \in L(T^*_{\ell-1})$ must have degree 2 in T^*. Let a_u be the neighbour of $u \in L(T^*_{\ell-1})$ that is in $T^* \setminus T^*_{\ell-1}$. Then, T^*_ℓ is obtained from $T^*_{\ell-1}$ by adding a_u for each $u \in L(T^*_{\ell-1})$ along with the $k-1$ neighbours of a_u that are not contained in $T^*_{\ell-1}$. Therefore $T^*_\ell \cong T^+_\ell$. □

Our proof of 4.2.6 will prove that $F(B, b) = (F, w)$ is satisfiable by showing that $\mathrm{supp}(Q) \subseteq \bigcap_{a \in C_m} I_a$, where $\{I_a\}_{a \in C_m}$ is the unique maximal collection of ideals of R such that $\bigoplus_{a \in C_m} I_a \subseteq \{Bx : x \in R^{V_n}\}$. The following lemma is used to relate the sequence $\{I_\ell(P)\}_{\ell \in \mathbb{N}}$ to $\{I_a\}_{a \in C_m}$. For any set S and $s \in S$, we let e_s denote the element of R^S with a one in the s-coordinate and zeros everywhere else.

Lemma 4.3.8. *Given $B \in R^{C_m \times V_n}$, let $\{I_a\}_{a \in C_m}$ be the unique maximal collection of ideals such that*

$$\bigoplus_{a \in C_m} I_a \subseteq \{Bx : x \in R^V\}.$$

Then $\{I_a\}_{a \in C_m}$ satisfies

$$I_a \supseteq B_{a,v} \bigcap_{a' \in C_m \setminus \{a\}} \{t \in R : B_{a',v} t \in I_{a'}\}$$

for all $a, \in C_m$, $v \in V_n$.

Proof. Let $a \in C_m$ and $v \in V_n$. Let $t \in R$ be such that $B_{a',v} t \in I_{a'}$ for all $a' \in C_m \setminus \{a\}$. We show that $B_{a,v} t e_a \in \{Bx : x \in R^{V_n}\}$, which implies that $B_{a,v} t \in I_a$ by maximality of $\{I_a\}_{a \in C_m}$.

For each $a' \in C_m \setminus \{a\}$, since $B_{a',v} t \in I_{a'}$, there is an $x^{(a')} \in R^V$ such that $Bx^{(a')} = B_{a',v} t e_{a'}$. Letting

$$x = t e_v - \sum_{a' \in C_m \setminus \{a\}} x^{(a')}$$

we then have

$$Bx = Bt e_v - \sum_{a' \in C_m \setminus \{a\}} B_{a',v} t e_{a'} = B_{a,v} t e_a.$$

Therefore $B_{a,v} t e_a \in \{Bx : x \in R^{V_n}\}$, as desired. □

Proof of Lemma 4.2.6. Suppose that $\ell(P, Q) = \infty$, or that $\ell(P, Q) < \infty$ and F does not contain $T^+_{\ell(P,Q)}$ as a subgraph. Then there is a finite $\ell^* \leq \ell(P, Q)$ such that F does not contain $T^+_{\ell^*}$ as a subgraph. By Lemma 4.3.7, $\bar{h}_F(a) < \ell^*$ for all $a \in C(F)$. Let $B \in R^{C_m \times V_n}$ be the matrix corresponding to F, so that $F(B) = B$, and let $\{I_a\}_{a \in C_m}$ be the unique maximal collection of ideals such that $\bigoplus_{a \in C_m} I_a \subseteq \{Bx : x \in R^{V_n}\}$. To show that F is satisfiable, it suffices to show that $I_a \supseteq \mathrm{supp}(Q)$ for all $a \in C(F)$. We show that

$$I_a \supseteq I_{\bar{h}_F(a)}(P), \quad \text{for all } a \in C(F). \tag{4.10}$$

By the definition of $\ell(P,Q)$ and the fact that $\bar{h}_F(a) < \ell(P,Q)$ for every $a \in C(F)$, it follows that $I_{\bar{h}_F(a)}(P) \supseteq \operatorname{supp} Q$ for every $a \in C(F)$, which implies the satisfiability of F.

We prove (4.10) by induction on the value of $\bar{h}_F(a)$. If $\bar{h}_F(a) = 0$, then there is a $v \in L(F)$ adjacent to a. Since v is a leaf, $I_a \supseteq (B_{a,v})$ by maximality of $\{I_a\}_{a \in C_m}$. Since $B_{a,v} \in \operatorname{supp}(P)$, we have $(B_{a,v}) \supseteq \bigcap_{r \in \operatorname{supp}(P)} rR = I_0(P)$. This verifies (4.10) for all $a \in C(F)$ for which $\bar{h}_F(a) = 0$.

Now suppose $\ell \geq 1$ and suppose that (4.10) holds for all $a \in C(F)$ for which $\bar{h}_F(a) \leq \ell - 1$. Let $a \in C(F)$ such that $\bar{h}_F(a) = \ell$. Let $v_1, \ldots, v_k$ be the neighbours of a. Since $\bar{h}_F(a) = \ell \geq 1$, the variable vertices $v_1, \ldots, v_k$ are not leaves of F. We claim that there exists $1 \leq i \leq k$ for which the following holds:

$$\bar{h}_F(a') \leq \ell - 1, \quad \text{for all } a' \neq a \text{ adjacent to } v_i \text{ in } F \tag{4.11}$$

Suppose for a contradiction that each v_i is adjacent to some $a_i \neq a$ with $\bar{h}_F(a_i) \geq \ell$. For each a_i, let $T_i \in \mathcal{T}_2(F)$ be such that $a_i \in T_i$ and

$$h_{T_i}(a_i) = \bar{h}_F(a_i), \tag{4.12}$$

recalling that $h_{T_i}(a_i) = \min_{v \in L(T_i)} \{(d_{T_i}(a_i, v) - 1)/2\}$.

For each $1 \leq i \leq k$, let T_i' be the component of $T_i - av_i$ that contains a_i (that is, T_i' is obtained by deleting the edge av_i and taking the component that contains a_i). Note that $T_1', \ldots, T_k'$ are disjoint, as otherwise F would contain a cycle. Let $T = \bigcup_{i \in [k]} (T_i' + av_i)$. We first show that $T \in \mathcal{T}_2(F)$. The connected subgraphs $T_1', \ldots, T_k'$ are connected in T by paths through a, and so T is connected. It is obvious that every $a' \neq a$ in T has degree k in T since they each have degree k in the subtrees $T_1', \ldots, T_k'$. Immediately, all the constraint vertices in T, including a, have degree k in T. For the variable vertices in T, they all have degree two or one in the subtrees $T_1', \ldots, T_k'$. Moreover, $v_1, \ldots, v_k$ were leaves in the subtrees, who become degree two vertices in T. Therefore $T \in \mathcal{T}_2(F)$. Let π be a shortest a, v-path in T among $v \in L(T)$. Since $v_1, \ldots, v_k$ are not leaves of T, π must contain a path π' in T_i from a_i to v for some $i \in [k]$ such that $|\pi| = |\pi'| + 2$. But now, by (4.12),

$$\ell = \bar{h}_F(a) \geq (|\pi| - 1)/2 = (|\pi'| - 1)/2 + 1 \geq h_{T_i}(a_i) + 1 = \bar{h}_F(a_i) + 1 \geq \ell + 1.$$

This is a contradiction, which completes the proof for (4.11).

Let $v' \in \{v_1, \ldots, v_k\}$ be a neighbour of a that satisfies (4.11). By Lemma 4.3.8,

$$I_a \supseteq B_{a,v'} \bigcap_{a' \in N(v') \setminus \{a\}} \{t \in R : B_{a',v'} t \in I_{a'}\}.$$

By the induction hypothesis and the choice of v' so that it satisfies (4.11), we conclude that $I_{a'} \supseteq I_{\bar{h}_F(a')}(P)$ for all $a' \in N(v') \setminus \{a\}$. By Observation 4.3.4, $I_{\bar{h}_F(a')}(P) \supseteq I_{\ell-1}(P)$ since

$\bar{h}_F(a') \le \ell - 1$. Thus, $I_{a'} \supseteq I_{\ell-1}(P)$ and so

$$\begin{aligned} B_{a,v} \bigcap_{a' \in N(v') \setminus \{a\}} \{t \in R : B_{a',v} t \in I_{a'}\} &\supseteq B_{a,v} \bigcap_{a' \in N(v') \setminus \{a\}} \{t \in R : B_{a',v} t \in I_{\ell-1}(P)\} \\ &\supseteq \bigcap_{r \in \operatorname{supp}(P)} r \left(\bigcap_{s \in \operatorname{supp}(P)} \{t \in R : st \in I_{\ell-1}(P)\} \right) \\ &= I_\ell(P). \end{aligned}$$

Therefore $I_a \supseteq I_\ell(P)$, which confirms (4.10) for all $a \in C(F)$ for which $\bar{h}_F(a) = \ell$. By induction, (4.10) holds for every $a \in C(F)$. □

Proof of Lemma 4.2.5. Let $\ell^+ = \ell^+(P, Q)$. Since $\ell^+ < \infty$, there is an $x \in \operatorname{supp}(P)^{\mathbb{N}}$ and a $b \in \operatorname{supp}(Q)$ such that $b \notin I^+_{\ell^+}(x)$.

Let $D_0 = C(T^+_0)$ and $D_\ell = C(T^+_\ell) \setminus C(T^+_{\ell-1})$ for $1 \le \ell \le \ell^+$. Then, $D_0, \dots, D_{\ell^+}$ is a partition of $C(T^+_{\ell^+})$. Similarly, let $U_0 = V(T^+_0)$ and $U_\ell = V(T^+_\ell) \setminus V(T^+_{\ell-1})$ for $1 \le \ell \le \ell^+$. Let a_0 be the unique constraint in D_0. For each $\ell = 0, \dots, \ell^+(P, Q)$ and edge $av \in E(T^+_{\ell^+})$ with $a \in D_\ell$, set

$$w(av) = \begin{cases} x_{2(\ell^+ - \ell)} & v \in U_\ell \\ x_{2(\ell^+ - \ell)+1} & v \notin U_\ell. \end{cases}$$

Define constraint vertex weights on $T^+_{\ell^+}$ by setting $w(a_0) = b$ and setting $w(a) \in \operatorname{supp}(Q)$ arbitrarily for $a \in C(T^+_{\ell^+}) \setminus \{a_0\}$. To show that $(T^+_{\ell^+}, w)$ is unsatisfiable, suppose $y \in R^{V(T^+_{\ell^+})}$ satisfies every constraint in $C(T^+_{\ell^+}) \setminus \{a_0\}$, i.e.

$$\sum_{v \in N(a)} w(av) y_v = w(a), \quad a \in C(T^+_{\ell^+}) \setminus \{a_0\},$$

where $N(a)$ denotes the set of neighbours of a in $T^+_{\ell^+}$. We show that for all $0 \le \ell \le \ell^+$

$$w(av) y_v \in I^+_{\ell^+ - \ell}(x) \quad \text{for all } a \in D_\ell,\ v \in N(a) \cap U_\ell \tag{4.13}$$

by reverse induction on ℓ. Note that (4.13) implies that $\sum_{v \in N(a_0)} w(a_0 v) y_v \in I^+_{\ell^+}(x)$. Since $w(a_0) = b \notin I^+_{\ell^+}(x)$, this implies that y does not satisfy $(T^+_{\ell^+}, w)$.

It remains to prove (4.13). For the base case $\ell = \ell^+$, suppose $a \in D_{\ell^+}$ and $v \in N(a) \cap U_{\ell^+}$. Then as $w(av) = x_0$ by definition of x, $w(av) y_v \in (x_0) = I_0(x)$.

Now suppose $0 \le \ell < \ell^+$ and suppose that (4.13) holds for all greater values of ℓ. Let $a \in D_\ell$ and $v \in N(a) \cap U_\ell$. Since $\ell < \ell^+$, there is an $a' \in N(v) \cap D_{\ell+1}$. Let $v_1, \dots, v_{k-1}$ be the $k-1$ neighbours of a' other than v. In other words, $\{v_1, \dots, v_{k-1}\} = N(a') \cap U_{\ell+1}$. By the induction hypothesis, $w(a' v_i) y_{v_i} \in I^+_{\ell^+ - \ell - 1}(x)$ for every $i \in [k-1]$. Since $a' \ne a_0$, y satisfies a', and so we have

$$w(a'v) y_v = w(a') - \sum_{i \in [k-1]} w(a' v_i) y_{v_i} \in w(a') + I_{\ell^+ - \ell - 1}(x).$$

Since $\ell^+ - \ell - 1 < \ell^+$, by definiton of $\ell^+ = \ell^+(P,Q)$, $w(a') \in \mathrm{supp}(Q) \subseteq I^+_{\ell^+-\ell-1}(x)$. Therefore $w(a'v)y_v \in I^+_{\ell^+-\ell-1}(x)$. Since $a \in D_\ell$, $v \in U_\ell$, we have $w(av) = x_{2(\ell^+-\ell)}$, and since $a' \in D_{\ell+1}$, $v \notin U_{\ell+1}$, we have $w(a'v) = x_{2(\ell^+-\ell)-1}$. Therefore

$$w(av)y_v = x_{2(\ell^+-\ell)}y_v \in \{x_{2(\ell^+-\ell)}t : x_{2(\ell^+-\ell)-1}t \in I^+_{\ell^+-\ell-1}(x)\} = I^+_{\ell^+-\ell}(x).$$

This completes the induction step and confirms that (4.13) holds for all $0 \le \ell \le \ell^+$. □

Proof of Lemma 4.2.2. Let $\bar{\ell} = \ell(P,Q)$. Since $\bar{\ell} < \infty$, there is a $b \in \mathrm{supp}(Q)$ such that

$$b \notin I_{\bar{\ell}}(P) = \bigcap_{r \in \mathrm{supp}(P)} r\left(\bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-1}(P)\}\right).$$

Let $r_0 \in \mathrm{supp}(P)$ be such that

$$b \notin r_0\left(\bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-1}(P)\}\right). \tag{4.14}$$

Let $D_0 = C(T_0), U_0 = V(T_0)$ and let $D_\ell = C(T_\ell) \setminus C(T_{\ell-1}), U_\ell = V(T_\ell) \setminus V(T_{\ell-1})$ for $1 \le \ell \le \bar{\ell}$.

Assign weights w to the edges of $T_{\ell(P,Q)}(P)$ as follows.

- $w(a_0v) = r_0$ for all $v \in N(a_0)$.
- For all $v \in U_\ell$ with $0 \le \ell < \bar{\ell}$ and all $(r,s) \in \mathrm{supp}(P)^2$, there is a constraint $a \in N(v) \cap D_{\ell+1}$ such that $w(av) = s$ and $w(av') = r$ for all $v' \in N(a) \cap U_{\ell+1}$.

Since each variable vertex $v \in U_\ell$ with $0 \le \ell < \bar{\ell}$ is adjacent to precisely $|\mathrm{supp}(P)|^2$ constraint vertices in $D_{\ell+1}$ by definition of $T_{\bar{\ell}}$, the edge weights w are well defined. Define constraint vertex weights by setting $w(a_0) = b$ and setting $w(a) \in \mathrm{supp}(Q)$ arbitrarily for $a \in C(T_{\bar{\ell}}) \setminus \{a_0\}$.

To show that $(T_{\bar{\ell}}, w)$ is unsatisfiable, suppose $y \in R^{V(T_{\bar{\ell}})}$ satisfies

$$\sum_{v \in N(a)} w(av)y_v = w(a), \quad a \in C(T_{\bar{\ell}}) \setminus \{a_0\}.$$

We show by reverse induction on ℓ that

$$y_v \in \bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-\ell-1}(P)\}, \quad \text{for all } v \in U_\ell \tag{4.15}$$

for all $0 \le \ell \le \bar{\ell}$, where $I_{-1}(P)$ is defined to be R. We show that (4.15) implies that y does not satisfy a_0. We have $N(a_0) \subseteq U_0$. By (4.15) with $\ell = 0$,

$$\sum_{v \in N(a_0)} w(a_0 v) y_v = r_0 \sum_{v \in N(a_0)} y_y \in r_0 \left(\bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-1}(P)\} \right).$$

By (4.14), this shows that y does not satisfy a_0. Therefore $(T_{\bar{\ell}}, w)$ is not satisfiable.

Now we prove (4.15). For the base case $\ell = \bar{\ell}$, we have

$$\bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in R\} = R,$$

and so (4.15) holds trivially.

Now suppose $0 \le \ell < \bar{\ell}$ and that (4.15) holds for $\ell + 1$. Let $v \in U_\ell$ and let $(r', s') \in \mathrm{supp}(P)^2$. Then there is an $a \in N(v) \cap D_{\ell+1}$ such that $w(av) = s'$ and $w(av') = r'$ for all $v' \in N(a) \cap U_{\ell+1}$. By the induction hypothesis

$$y_{v'} \in \bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-\ell-2}(P)\}$$

for all $v' \in N(a) \cap U_{\ell+1}$. Since y satisfies $a \ne a_0$, we have

$$w(av) y_v = w(a) - r' \sum_{v' \in N(a) \setminus \{v\}} y_{v'} \in w(a) + r' \bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-\ell-2}(P)\}.$$

Since $w(av) = s'$ and by definition of $\bar{\ell} = \ell(P, Q)$ that

$$w(a) \in \mathrm{supp}(Q) \subseteq I_{\ell-1}(P) \subseteq r' \bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-\ell-2}(P)\},$$

we obtain that

$$s' y_v \in r' \bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-\ell-2}(P)\}.$$

Since this holds for all $r' \in \mathrm{supp}(P)$, we have $s' y_v \in I_{\bar{\ell}-\ell-1}(P)$ for all $s' \in \mathrm{supp}(P)$. Therefore $y_v \in \bigcap_{s \in \mathrm{supp}(P)} \{t \in R : st \in I_{\bar{\ell}-\ell-1}(P)\}$. This completes the induction step, and so (4.15) holds for $0 \le \ell \le \bar{\ell}$. □

Chapter 5

Linear SAT threshold for principal rings: proof of Theorem 1.3.3

We prove Theorem 1.3.3 by a sequence of reductions to Theorem 3.2.2. First, we reduce the problem to considering equations over finite local rings.

Recall from Lemma 4.3.1 that any finite ring R can be writen as a product $R \cong R_1 \times \cdots \times R_h$ where $R_1, \ldots, R_h$ are local rings (not necessarily principal). Let $Bx = b$ be a system of linear equations over R. For each $i \in [h]$, let $\pi_i : R \to R_i$ be the natural projection map onto the i-th coordinate. For each $i \in [h]$, applying π_i entrywise to both B and b defines a system of linear equations over R_i, which we denote by $B^{(i)}\, x = b^{(i)}$. It follows immediately that

$$Bx = b \text{ is satisfiable over } R \iff B^{(i)}\, x = b^{(i)} \text{ is satisfiable over } R_i, \text{ for every } i \in [h]. \quad (5.1)$$

Consider the linear system $\mathbf{B}x = \mathbf{b}$ given in (1.1). By (5.1), it suffices to determine the critical m around which one of the systems $\mathbf{B}^{(i)}x = \mathbf{b}^{(i)}$ becomes unsatisfiable. This leads us to study the distribution of $(\mathbf{B}^{(i)}, \mathbf{b}^{(i)})$. Given P and Q, both of which are distributions over $R = R_1 \times \cdots \times R_h$, let P_i and Q_i be the marginal distributions of P and Q respectively on R_i. In other words, for each $i \in [h]$, $P_i := P \circ \pi_i^{-1}$ and $Q_i := Q \circ \pi_i^{-1}$. Clearly, $(\mathbf{B}^{(i)}, \mathbf{b}^{(i)})$ has the same distribution as $(\mathbf{B}, \mathbf{b})$ with (P, Q) replaced by (P_i, Q_i).

Our next lemma shows that the marginal distributions of P on each local ring $R_1, \ldots, R_h$ have a very simple form if $\alpha_{P,Q} = 1$.

Lemma 5.0.1. *Suppose $R = R_1 \times \cdots \times R_h$, where $R_1, \ldots, R_h$ are local principal rings. Suppose Q, P are distributions on R such that $\alpha_{P,Q} = 1$. Let $P_i := P \circ \pi_i^{-1}$ and $Q_i := Q \circ \pi_i^{-1}$. If $i \in [h]$ is such that $(P_i^{\otimes k}, Q_i)$ satisfies* **(A2)**, *then $|\{rR_i : r \in \mathrm{supp}(P_i)\}| = 1$.*

Proof of Lemma 5.0.1. Let M_i be the unique maximal ideal of R_i, and suppose for a contradiction that $|rR_i : r \in \mathrm{supp}(P_i)\}| \geq 2$ and $(P_i^{\otimes k}, Q)$ satisfies **(A2)**. Then there are $r_1, r_2 \in \mathrm{supp}(P)$ such that $\pi_i(r_1)R_i \neq \pi_i(r_2)R_i$, and some $b \in \mathrm{supp}(Q)$ such that $\pi_i(b) \neq 0$.

Since R_i is a local principal ring, we have without loss of generality that $\pi_i(r_1)R_i \subsetneq \pi_i(r_2)R_i$. It follows then that there is an $s \in R_i$ such that $\pi_i(r_1) = \pi_i(r_2)s$. Moreover, s cannot be a unit as otherwise it contradicts with $\pi_i(r_1)R_i \neq \pi_i(r_2)R_i$. Hence, $s \in M_i$. By Observation 4.3.4, there exists an $\ell^* \geq 0$ such that $I_{\ell^*}(P) = I_{\ell^*-1}(P)$. Then we have

$$I_{\ell^*}(P) \subseteq \{r_1 t : r_2 t \in I_{\ell^*-1}(P),\ t \in R\},$$

and so

$$\pi_i(I_{\ell^*}(P)) \subseteq \{\pi_i(r_1)t : \pi_i(r_2)t \in \pi_i(I_{\ell^*-1}(P)),\ t \in R_i\} \subseteq s\pi_i(I_{\ell^*-1}(P)) = s\pi_i(I_{\ell^*}(P)).$$

By Lemma 4.3.2, there is a $\nu \geq 0$ such that $M^\nu = 0$. Since $s \in M$, we then have $\pi_i(I_{\ell^*}(P)) = s^\nu \pi_i(I_{\ell^*}(P)) = 0$. Since $\pi_i(b) \neq 0$, this implies $\ell(P, Q) \leq \ell^*$, which contradicts the assumption that $\alpha_{P,Q} = 1$. □

Finally, we complete the proof of Theorem 1.3.3 by reducing it to Theorem 3.2.2 via Lemma 5.0.1.

Proof of Theorem 1.3.3. By Lemma 4.3.1, there exist principal local rings $R_1, \ldots, R_h$ such that $R \cong R_1 \times \cdots \times R_h$. By (5.1) and the discussions below it, it suffices to show that the set of systems $\mathbf{B}^{(i)}x = \mathbf{b}^{(i)}$ is a.a.s. simultanously satisfiable if $m < (d_k - \varepsilon)n/k$, and there exists some $i \in [h]$ such that $\mathbf{B}^{(i)}x = \mathbf{b}^{(i)}$ is a.a.s. unsatisfiable if $m > (d_k + \varepsilon)n/k$.

We first consider the subcritical regime where $m < (d_k - \varepsilon)n/k$. Let $i \in [h]$. If $(P_i^{\otimes k}, Q_i)$ admits a constant solution $r_i \in R_i$ then $\mathbf{B}^{(i)}x = \mathbf{b}^{(i)}$ is satisfiable trivially. Thus we may assume that $(P^{\otimes k}, Q_i)$ satisfies **(A2)**. Let $g \in R_i$ be a generator of the maximal ideal M_i. Since R_i is local and principal, by Lemma 4.3.2(b), there is an $\ell \geq 0$ such that

$$rR_i = M_i^\ell = (g^\ell) \text{ for all } r \in \operatorname{supp}(P_i), \tag{5.2}$$

where the uniformity of ℓ for every $r \in \operatorname{supp}(P_i)$ is a consequence of the fact that $|\{rR_i : r \in \operatorname{supp}(P_i)\}| = 1$ which follows by Lemma 5.0.1. Define a map $\phi : M_i^\ell \to R_i/\operatorname{Ann}(M_i^\ell)$ by $\phi(sg^\ell) = s + \operatorname{Ann}(M_i^\ell)$ for all $s \in R_i$. For a matrix B and vector b over R_i, let $\phi(B)$ and $\phi(b)$ be the matrix and vector obtained by applying ϕ entrywise. We use the following facts about ϕ.

Claim 5.0.2. *(a) $\phi(r)$ is a unit of $R_i/\operatorname{Ann}(M_i^\ell)$ for all $r \in \operatorname{supp}(P_i)$.*

(b) Suppose $B \in R_i^{C_m \times V_n}$ and $b \in R_i^{C_m}$ both have entries in M_i^ℓ. Then $Bx = b$ is satisfied by some $x \in R^V$ if and only if $\phi(B)y = \phi(b)$ is satisfied by some $y \in (R_i/M_i^\ell)^V$.

Proof of Claim 5.0.2. If $r \in \operatorname{supp}(P)$, then $rR = M_i^\ell \supsetneq M_i^{\ell+1}$, and so $s \notin M_i$ for any $s \in R_i$ such that $r = sg^\ell$. Therefore $\phi(r) \notin M_i/\operatorname{Ann}(M_i^\ell)$ for all $r \in \operatorname{supp}(P)$, and so $\phi(r)$ is a unit of $R/\operatorname{Ann}(M_i^\ell)$ for all $r \in \operatorname{supp}(P)$. This confirms part (a).

For part (b), note that ϕ is an isomorphsim of R_i-modules. Applying ϕ coordinate wise gives an isomorphism between $(M_i^\ell)^{C_m}$ and $(R_i/\operatorname{Ann}(M_i^\ell))^{C_m}$. Therefore $\phi(b)$ is in the column space of $\phi(B)$ if and only if b is in the column space of B. Moreover, for any $y \in R_i^V$ and $\bar{y} = y + \bigoplus_{v\in V_n} \operatorname{Ann}(M_i^\ell)$ we have $\phi(B)y = \phi(B)\bar{y}$. Therefore $Bx = b$ is satisfied by some $x \in R_i^{V_n}$ if and only if $\phi(B)\bar{y} = \phi(b)$ is satisfied by some $\bar{y} \in \bigoplus_{v\in V_n} R_i/\operatorname{Ann}(M_i^\ell)$. □

By (5.2), we have $\operatorname{supp}(P_i) \subseteq M_i^\ell$. We also have $\operatorname{supp}(Q_i) \subseteq M_i^\ell$, as otherwise $\operatorname{supp}(Q_i) \not\subseteq M_i^\ell = \pi_i(I_0(P))$, which would imply that $\ell(P, Q) = 0$, contradicting with the assumption that $\alpha_{P,Q} = 1$. Thus, Claim 5.0.2(b) implies that the system $\mathbf{B}^{(i)}\bar{x} = \mathbf{b}^{(i)}$ is satisfiable if and only if $\phi(\mathbf{B}^{(i)})\bar{x} = \phi(\mathbf{b}^{(i)})$ is satisfiable. By part (a) of the claim and Theorem 3.2.2, $\phi(\mathbf{B}^{(i)})\bar{x} = \phi(\mathbf{b}^{(i)})$ is a.a.s. satisfiable. This completes the proof for the subcritical regime.

The proof for the supercirical regime $m > (d_k + \varepsilon)n/k$ is similar. Since $(P^{\otimes k}, Q)$ satisfies **(A2)**, there must exists $i \in [h]$ such that $(P_i^{\otimes k}, Q_i)$ satisfies **(A2)**. By Claim 5.0.2(a) and Theorem 3.2.2, a.a.s. $\phi(\mathbf{B}^{(i)})\bar{x} = \phi(\mathbf{b}^{(i)})$ is unsatisfiable. Thus, a.a.s. $\mathbf{B}^{(i)}\bar{x} = \mathbf{b}^{(i)}$ is unsatisfiable by Claim 5.0.2(b), which, together with (5.1), further implies that a.a.s. (1.1) is unsatisfiable, completing the proof of Theorem 1.3.3.

□

Chapter 6

Linear SAT threshold for nonprincipal rings: proof of Theorem 1.3.9

Throughout this section, we let $R = \mathbb{F}_q[X,Y]/(X^2,Y^2)$. Given any $B \in \{\bar{X},\bar{Y}\}^{C_m \times V_n}$ and $b \in \{\bar{X}\bar{Y},0\}^{C_m}$, define $B_X, B_Y \in \mathbb{F}_q^{C_m \times V_n}$ by

$$(B_X)_{a,v} = \mathbf{1}_{\{B_{a,v}=\bar{X}\}}, \quad (B_Y)_{a,v} = \mathbf{1}_{\{B_{a,v}=\bar{Y}\}}, \quad b_{XY} = \mathbf{1}_{\{b_a=\bar{X}\bar{Y}\}},$$

where $\mathbf{1}_{\{B_{a,v}=p\}}$ is the indicator function on $(a,v) \in C_m \times V_n$, ie. $\mathbf{1}_{\{B_{a,v}=p\}} = 1$ if $B_{a,v} = p$ and $\mathbf{1}_{\{B_{a,v}=p\}} = 0$ otherwise. Let $B_{XY} = [B_X \mid B_Y]$. To prove Theorem 1.3.9, we first show that the system $Bx = b$ over R is equivalent to $B_{XY}x = b_{XY}$ as a system of equations over $\mathbb{F}_q$. This holds deterministically for any $B \in \{\bar{X},\bar{Y}\}^{C_m \times V_n}$ and $b \in \{\bar{X}\bar{Y},0\}^{C_m}$. We then use a coupling argument together with Theorem 7.3.5 to apply the results of [14], which gives the satisfiability threshold.

Lemma 6.0.1. *Suppose that $B \in \{\bar{X},\bar{Y}\}^{C_m \times V_n}$ and $b \in \{\bar{X}\bar{Y},0\}^{C_m}$. If $Bx = b$ is satisfiable then there are $c, c' \in \mathbb{F}_q^{V_n}$ such that $B(c\bar{X} + c'\bar{Y}) = b$.*

Proof. Let $x \in R^{V_n}$ be such that $Bx = b$. Let $c^{(1)}, c^{(2)}, c^{(3)}, c^{(4)} \in \mathbb{F}_q^{V_n}$ be such that $x = c^{(1)}\bar{X}\bar{Y} + c^{(2)}\bar{X} + c^{(3)}\bar{Y} + c^{(4)}$. We then have

$$\bar{X}\bar{Y}Bc^{(1)} + \bar{X}Bc^{(2)} + \bar{Y}Bc^{(3)} + B_{a,v}c^{(4)} = b \in (\bar{X}\bar{Y})^{C_m}$$

Since $B_{a,v} \in \{\bar{X},\bar{Y}\}$ for all $a \in C_m$, $v \in V_n$, we have $\bar{X}\bar{Y}Bc^{(1)}, \bar{X}Bc^{(2)}, \bar{Y}Bc^{(3)} \in (\bar{X}\bar{Y})^{C_m}$. Thus, for any $a \in C_m$,

$$(Bc^{(4)})_a = \bar{X} \sum_{v \in V_n : a(v) = \bar{X}} c_v^{(4)} + \bar{Y} \sum_{v \in V_n : a(v) = \bar{Y}} c_v^{(4)} \in (\bar{X}\bar{Y})$$

But the only $u, u' \in \mathbb{F}_q$ such that $\bar{X}u + \bar{Y}u' \in (\bar{X}\bar{Y})$ are $u = u' = 0$. Therefore $(Bc^{(4)})_a = 0$ for all $a \in C_m$. We also have $\bar{X}\bar{Y}Bc^{(1)} = 0$, as $\bar{X}\bar{Y}B_{a,v} = 0$ for all $a \in C_m$ and $v \in V_n$. Therefore $B(c^{(2)}\bar{X} + c^{(3)}\bar{Y}) = b$. □

Lemma 6.0.2. *Suppose that $B \in \{\bar{X}, \bar{Y}\}^{C_m \times V_n}$ and $b \in \{\bar{X}\bar{Y}, 0\}^{C_m}$. Then, the system $B_{XY}\, x = b_{XY}$ is satisfiable over $\mathbb{F}_q^V \times \mathbb{F}_q^V$ if and only if $Bx = b$ is satisfiable over R^{V_n}.*

Proof. Given $z = (x, y) \in \mathbb{F}_q^{V_n} \times \mathbb{F}_q^{V_n}$, define $z' \in R^V$ by

$$z' = x\bar{Y} + y\bar{X}.$$

By Lemma 6.0.1, it suffices to show that $B_{XY}z = b_{XY}$ if and only if $Bz' = b$. For any $a \in C_m$, we have

$$\begin{aligned}
(Bz')_a &= \sum_{v \in V_n} B_{a,v} z'_v \\
&= \sum_{v \in V_n : B_{a,v} = \bar{X}} \bar{X}(x_v\bar{Y} + y_v\bar{X}) + \sum_{v \in V_n : B_{a,v} = \bar{Y}} \bar{Y}(x_v\bar{Y} + y_v\bar{X}) \\
&= \bar{X}\bar{Y} \left(\sum_{v \in V_n : B_{a,v} = \bar{X}} x_v + \sum_{v \in V_n : B_{a,v} = \bar{Y}} y_v \right) \\
&= \bar{X}\bar{Y}(B_{XY}z)_a.
\end{aligned}$$

Since $b = \bar{X}\bar{Y}b_{XY}$, this shows that $B_{XY}z = b_{XY}$ if and only if $Bz' = b$. □

6.1 Satisfiability

In this section we prove satisfiability of $\mathbf{B}_{XY}x = \mathbf{b}_{XY}$ in the subcritical case. Due to Lemma 6.0.2, it suffices to prove the following.

Lemma 6.1.1. *Suppose $m < (d_{k,\rho} - \varepsilon)n/k$ for some fixed $\varepsilon > 0$. Then $\mathbf{B}_{XY}$ has full row rank a.a.s.*

We prove Lemma 6.1.1 using a coupling argument and the full rank condition in [14]. As mentioned in Section 2, the random matrix model in [14] can be constructed using the random graph $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$. The construction is as follows: given integer random variables $\boldsymbol{k}$, $\boldsymbol{d}$ such that $\mathbb{E}[\boldsymbol{d}^r], \mathbb{E}[\boldsymbol{k}^r] < \infty$ for some $r > 2$ and a random variable χ on $\mathbb{F}_q \setminus \{0\}$, choose $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ and $\{\chi_{a_i,v_j}\}_{i,j\in\mathbb{N}}$ independently, such that $\{\chi_{a_i,v_j}\}_{i,j\in\mathbb{N}}$ are i.i.d. copies of χ. Then set

$$\mathbf{A}_{a_i,v_i} = \mathbf{1}_{\{a_iv_i \in E(G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}})\}} \chi_{a_i,v_i}.$$

The following theorem is the main result of [14].

Theorem 6.1.2 ([14])**.** *Let $K(z)$ and $D(z)$ be the probability generating functions of $\boldsymbol{k}$ and $\boldsymbol{d}$ respectively. Define*

$$\Phi_{\boldsymbol{k},\boldsymbol{d}}(z) = D(1 - K'(z)/\mathbb{E}[k]) - \frac{\mathbb{E}[d]}{\mathbb{E}[k]}(1 - K(z) - (1-z)K'(z)). \tag{6.1}$$

If q and $\gcd(\boldsymbol{d})$ are relatively prime and $\Phi_{\boldsymbol{k},\boldsymbol{d}}(z) < \Phi_{\boldsymbol{k},\boldsymbol{d}}(0)$ for all $0 < z \leq 1$, then $\mathbf{A}$ has full row rank a.a.s.

We deduce Lemma 6.1.1 from Theorem 6.1.2 via a coupling argument. In particular, we construct a random $\{0,1\}$-matrix $\mathbf{C}$ with two submatrices $\mathbf{C}', \mathbf{C}''$ such that

- every row of $\mathbf{C}$ is in $\mathbf{C}'$ a.a.s,
- every column of $\mathbf{C}$ is in $\mathbf{C}''$,
- $\mathbf{C}'$ is contiguous with $\mathbf{A}$ for some choice of $\boldsymbol{k}, \boldsymbol{d}, \chi$,
- $\mathbf{C}''$ is contiguous with $\mathbf{B}_{XY}$ up to permuting columns.

Since $\mathbf{C}$, $\mathbf{C}''$, and $\mathbf{C}'''$ are $\{0,1\}$-matrices, it suffices to define their factor graphs $F(\mathbf{C})$, $F(\mathbf{C}')$, $F(\mathbf{C}'')$. We first define three coupled multigraphs $M_{n,m,k,\rho,d^*\varepsilon}$, $M'_{n,k,\rho,d^*,\varepsilon}$, $M''_{n,m,k,\rho,\varepsilon}$. Fix $d^* > 0$ such that $(d^* - \delta)n/k > m$ for some $\delta > 0$ and fix $\varepsilon > 0$.

- Let $\boldsymbol{m} \sim \mathrm{Po}(d^*n/k)$, and let $\hat{\boldsymbol{m}} = \max\{m, \boldsymbol{m}\}$. The vertex set of $M_{n,m,k,\rho,d^*\varepsilon}$ is $V_{2n} \cup C_{\hat{\boldsymbol{m}}}$.

  To define the edge set of $M_{n,m,k,\rho,d^*\varepsilon}$, we first choose a random partition (A, A', B, B') of V_{2n} independent of $\hat{\boldsymbol{m}}$. Choose a set S uniformly from the subsets of V_{2n} of size $(2-2\varepsilon)n$. Conditioned on S, let A be a uniformly chosen random subset of S, and let $B = S \setminus A$. Conditioned on A and S, choose A' uniformly from the set of subsets $A' \subseteq V_{2n} \setminus S$ that minimize $||A \cup A'| - n|$. Finally, let $B' = V_{2n} \setminus (S \cup A')$.

  Now, conditioned on $\hat{\boldsymbol{m}}$ and (A, A', B, B'), let $\{b_{a,i}\}_{a \in C_{\hat{\boldsymbol{m}}}, i \in [k]}$ be i.i.d. random variables on V_{2n}, where $\mathbb{P}[b_{a,i} = v] = \rho/|A \cup A'|$ for $v \in A \cup A'$ and $\mathbb{P}[b_{a,i} = v] = (1-\rho)/|B \cup B'|$ for $v \in B \cup B'$. The edge set of $M_{n,m,k,\rho,d^*\varepsilon}$ is given by the multiset $\{ab_{a,i} : i \in [k], a \in C_{\hat{\boldsymbol{m}}}\}$, so that there is an edge from $a \in C_{\hat{\boldsymbol{m}}}$ to $v \in V_{2n}$ for each $i \in [k]$ such that $b_{a,i} = v$.

- Let $M'_{n,k,\rho,d^*,\varepsilon}$ be the subgraph of $M_{n,m,k,\rho,d^*,\varepsilon}$ induced by $C_{\boldsymbol{m}} \cup S$, and let $\boldsymbol{\kappa}'_a$ be the degree of a in $M'_{n,k,\rho,d^*,\varepsilon}$.
- Let $M''_{n,m,k,\rho,\varepsilon}$ be the subgraph of $M_{n,m,k,\rho,d^*,\varepsilon}$ induced by $C_m \cup V_{2n}$.

Note that $b_{a,1}, \ldots, b_{a,k} \in V_{2n}$ may not be distinct, and so $M_{n,m,k,\rho,\varepsilon}$ is random bipartite multigraph. Let $\mathcal{S}$ be the event that $M_{n,m,k,\rho,\varepsilon}$ has no parallel edges. We have

$$\mathbb{P}[\mathcal{S} \mid \hat{\boldsymbol{m}}] \geq \left(1 - \frac{k}{n}\right)^{\hat{\boldsymbol{m}}} \tag{6.2}$$

and $\hat{\boldsymbol{m}} = O(n)$ a.a.s., and hence $\mathbb{P}[\mathcal{S}] = \Omega(1)$. Choose the random graphs $(F(\mathbf{C}), F(\mathbf{C}'), F(\mathbf{C}''))$ from the conditional joint distribution of $(M_{n,m,k,\rho,d^*\varepsilon}, M'_{n,k,\rho,d^*,\varepsilon}, M''_{n,m,k,\rho,\varepsilon})$ given $\mathcal{S}$.

The following lemma shows that $F(\mathbf{C}'')$ is contiguous with $F(\mathbf{B}_{XY})$ up to permuting columns. Let π be a random permutation of V_{2n}, and let $\tilde{\mathbf{B}}_{XY}$ be the matrix obtained by permuting columns of $\mathbf{B}_{XY}$ according to π.

Lemma 6.1.3. *For any $\varepsilon > 0$ and $m = m(n)$, we have $F(\tilde{\mathbf{B}}_{XY}) \triangleleft\triangleright F(\mathbf{C}'')$*

We also show that $F(\mathbf{C}')$ is contiguous with $G^{(1)}_{(2-2\varepsilon)n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ for an appropriate choice of parameters. For each $T \subseteq V_{2n}$ of size $(2-2\varepsilon)$, fix a bijection $\sigma_T : T \to V_{(2-2\varepsilon)n}$. Let $\tilde{\mathbf{C}}' \in \{0,1\}^{C_{\boldsymbol{m}} \times V_{(2-2\varepsilon)n}}$ be the matrix obtained mapping the columns of $\mathbf{C}' \in \{0,1\}^{C_{\boldsymbol{m}} \times S}$ onto $V_{(2-2\varepsilon)n}$ via σ.

Lemma 6.1.4. *Let $\varepsilon > 0$ and let $d^* > 0$. Let $\boldsymbol{k} \sim \mathrm{Bin}(k, 1-\varepsilon)$, and let $\mathbf{d}$ be a convex combination of Poisson random variables $\boldsymbol{d} \sim \mathrm{Po}(\rho d^*)/2 + \mathrm{Po}((1-\rho)d^*)/2$. Then $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \triangleright F(\tilde{\mathbf{C}}')$.*

We defer the proofs of Lemmas 6.1.3 and 6.1.4 to Chapter 7, which also contains the proofs of the contiguity results discussed in Chapter 2. Using the lemmas, we now prove Lemma 6.1.1.

Proof of Lemma 6.1.1. Suppose $m < (d_{k,\rho} - \varepsilon')n/k$ for some fixed $\varepsilon' > 0$. Set $d^* = (1 - \varepsilon'/2)d_{k,\rho}$, and set $\boldsymbol{d} \sim \mathrm{Po}(\rho d^*)/2 + \mathrm{Po}((1-\rho)d^*)$. Setting $\boldsymbol{k}_\varepsilon \sim \mathrm{Bin}(k, 1-\varepsilon)$ for $\varepsilon > 0$, we have

$$\lim_{\varepsilon \to 0} \Phi_{\boldsymbol{k}_\varepsilon,\boldsymbol{d}}(z) = e^{-\rho d^* z^{k-1}} + e^{-(1-\rho)d^* z^{k-1}} - d^*(1-1/k)z^k + d^* z^{k-1} - d^*/k = \Phi_{d^*,k,\rho}(z)$$

Since $d^* < d_{k,\rho}$, we have $\Phi_{d^*,k,\rho}(z) < \Phi_{d^*,k,\rho}(0)$ for all $z \in (0,1]$. Taking derivatives shows that $\Phi^{(i)}_{d^*,k,\rho}(0) = 0$ for $i \in [k-1]$ and $\Phi^{(k)}_{d^*,k,\rho}(0) < 0$, so it follows that $\Phi_{\boldsymbol{k},\boldsymbol{d}}(z) < \Phi_{\boldsymbol{k},\boldsymbol{d}}(0)$ for all $z \in (0,1]$ for some sufficiently small $\varepsilon > 0$. Set $\boldsymbol{k} = \boldsymbol{k}_\varepsilon$, and take this choice of d^*, ε in the definition of $\mathbf{C}$, $\mathbf{C}'$, and $\mathbf{C}''$.

Theorem 6.1.2 shows that the $\{0,1\}$-matrix $\mathbf{A}$ with $F(\mathbf{A}) = G^{(1)}_{(2-2\varepsilon)n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ has full row rank a.a.s. for this choice of $\boldsymbol{k}, \boldsymbol{d}$. By Lemma 6.1.4, it follows that $\mathbf{C}'$ has full row rank a.a.s. On the event that $\boldsymbol{m} > m$ and $\mathbf{C}'$ has full row rank, $\mathbf{C}'$ contains every row of $\mathbf{C}$, and so $\mathbf{C}$ also has full row rank. Since $\boldsymbol{m} \sim \mathrm{Po}(d^* n/k)$ and $d^*/k > (d_{k,\rho} - \varepsilon')/k > m/n$, we have $\boldsymbol{m} > m$ a.a.s. Therefore $\mathbf{C}$ has full row rank a.a.s., and the submatrix $\mathbf{C}''$ has full row rank a.a.s. Lemma 6.1.3 then shows that $\mathbf{B}_{XY}$ has full row rank a.a.s. □

6.2 Unsatisfiability

In this section we prove that $\mathbf{B}_{XY}x = \mathbf{b}_{XY}$ is a.a.s. unsatisfiable in the supercritical case. Recall from the proof Theorem 3.2.2 that the 2-core a matrix B is the largest submatrix of B where every column contains at least 2 non-zero entries, and every row contains the same number of nonzero entries as it does in B. To show that $\mathbf{B}_{XY}x = \mathbf{b}$ is unsatisfiable, is suffices to show that the subsystem $(\mathbf{B}_{XY})_{\text{2-core}}x = (\mathbf{b})_{\text{2-core}}$ is unsatisfiable.

We begin by determining the dimensions of $(\mathbf{B}_{XY})_{\text{2-core}}$. For that, we use the following theorem, which is essentially Theorem 1.2 of [13], restated in a form that is easy to apply in our setting.

Suppose $\boldsymbol{d} = (d_1, \dots, d_n)$ and $\boldsymbol{k} = (k_1, \dots, k_{m(n)})$ are sequences of tuples, where $m(n)$ is a function of n. Suppose $(\boldsymbol{d}_n, \boldsymbol{k}_n)_{n\in\mathbb{N}}$ satisfies the following condition

$$\exists r > 2, C > 0 : \sum_{i=1}^{n} d_i^r \le Cn, \quad \sum_{i=1}^{m(n)} k_i^r \le Cm(n) \quad \text{for all } n. \tag{6.3}$$

Let π_d and π_k be defined by

$$\pi_d(j) = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{\mathbf{1}_{\{d_i=j\}}}{n}, \quad \pi_k(j) = \lim_{n\to\infty} \sum_{i=1}^{m} \frac{\mathbf{1}_{\{k_i=j\}}}{m}. \tag{6.4}$$

Then it is easy to see that (6.3) implies that π_d and π_k are well-defined and that both of them are probability distribution functions. Let $\bar{d} = \mathbb{E}[\pi_{\boldsymbol{d}}]$ and let $\bar{k} = \mathbb{E}[\pi_{\boldsymbol{k}}]$. Let D be the probability generating function of $\pi_{\boldsymbol{d}}$ and let K be the probability generating function of $\pi_{\boldsymbol{k}}$. Let

$$\phi(\alpha) = 1 - \alpha - D'(1 - K'(\alpha)/k)/d$$

and let

$$\eta = \eta_{d,k} = \max\{x \in [0,1] : \phi(x) = 0\}$$

For for each tuple $\boldsymbol{d}_n = (d_1, \dots, d_n)$ and $\boldsymbol{k}_n = (k_1, \dots, k_m)$, let $B_{\boldsymbol{k}_n,\boldsymbol{d}_n}$ be a random binary matrix chosen uniformly from the set of matrices with k_i nonzero entries in row i and d_j nonzero entries in column j for each $i \in [m(n)], j \in [n]$.

Theorem 6.2.1. *Suppose that $(\boldsymbol{d}, \boldsymbol{k})$ satisfies (6.3). Let $B \sim B_{\boldsymbol{d},\boldsymbol{k}}$. Let n^* and m^* be the number of columns and rows of $B_{2\text{-core}}$. Then,*

$$\frac{n^*}{n} \to 1 - D\left(1 - \frac{K'(\eta)}{\bar{k}}\right) - \frac{K'(\eta)}{\bar{k}} D'\left(1 - \frac{K'(\eta)}{\bar{k}}\right), \quad \frac{m*}{n} \to \frac{\bar{d}}{\bar{k}} K(\eta)$$

Proof. The proof is identical to the proof of [13, Theorem 1.2].

In our case, every row of $\mathbf{B}_{XY}$ has k nonzero entries, and the number of nonzero entries in each column of $\mathbf{B}_{XY}$ is close to a sequence of i.i.d. copies of $\boldsymbol{d}$, where $\mathbf{d}$ is a convex combination of Poissons $\boldsymbol{d} \sim \mathrm{Po}(\rho d^*)/2 + \mathrm{Po}((1-\rho)d^*)/2$ with expectation $\mathbb{E}[\boldsymbol{d}] = d^*/2$. Therefore, if n^* and m^* are the number of rows and columns in the 2-core of $\mathbf{B}_{XY}$, we get

$$\frac{n^*}{n} \to 2 - e^{-\eta^{k-1}} - e^{-(1-\rho)d^*\eta^{k-1}} - d^*\eta^{k-1} + d^*\eta^k, \quad \frac{m^*}{n} \to \frac{d^*}{k}\eta^k \tag{6.5}$$

Proof of Theorem 1.3.9. By Lemma 6.0.2, it suffices to show that $\mathbf{B}_{XY}x = \mathbf{b}_{XY}$ is satisfiable a.a.s. for $m < (d_{k,\rho} - \varepsilon)n/k$, and unsatisfiable a.a.s. for $m > (d_{k,\rho} + \varepsilon)n/k$.

The system $\mathbf{B}_{XY} = \mathbf{b}_{XY}$ is satisfiable a.a.s. in the subcritical case $m < (d_{k,\rho} - \varepsilon)n/k$ by Lemma 6.1.1.

For the supercritical case, $m > (d_{k,\rho} + \varepsilon)n/k$, consider the subsystem $\mathbf{B}'_{XY}x = \mathbf{b}'_{XY}$ consisting of the first $(d_{k,\rho} - \varepsilon_n)n/k$ rows of $\mathbf{B}_{XY}x = \mathbf{b}_{XY}$, where $\varepsilon_n = o(1)$. Let n' and m' be the number of rows and columns of in the 2-core of $\mathbf{B}'_{XY}$. By choosing ε_n so that $\varepsilon_n \to 0$ sufficiently slowly, we can ensure that $\mathbf{B}'_{XY}$ has full row rank by Lemma 6.1.1. By (6.5),

$$\begin{aligned}\frac{n' - m'}{n} &\to 1 - e^{-\rho d_{k,\rho}\eta_{d_{k,\rho},k}^{k-1}} - e^{-(1-\rho)d_{k,\rho}\eta_{d_{k,\rho},k}^{k-1}} - d_{k,\rho}\eta_{d_{k,\rho},k}^{k-1} + d_{k,\rho}\eta_{d_{k,\rho},k}^{k} - \frac{d_{k,\rho}}{k}\eta_{d_{k,\rho},k}^{k} \\ &= \Phi_{d,k,\rho}(0) - \Phi_{d,k,\rho}(\eta_{d_{k,\rho},k}) = 0\end{aligned}$$

in probability. The remainder of the proof follows a similar argument to the proof of Theorem 3.2.2, which we outline below.

Since $\mathbf{B}'_{XY}$ has full row rank a.a.s., $(\mathbf{B}'_{XY})_{2\text{-core}}$ also has full row rank a.a.s. Therefore the number of solutions to $(\mathbf{B}'_{XY})_{2\text{-core}}x = (\mathbf{b}'_{XY})_{2\text{-core}}$ is $q^{(n'-m')} = q^{o(n)}$ a.a.s. Let R' be the set of rows in $\mathbf{B}_{XY}$ that are not in $\mathbf{B}'_{XY}$ and have all of their non-zero entries in the columns of $(\mathbf{B}'_{XY})_{2\text{-core}}$. Since $n' = \Theta(n)$ and $m - (d_{k,\rho} - \varepsilon_n)n/k = \Omega(n)$ a.a.s., we can show that $|R'| = \Omega(n)$ a.a.s. Since $\text{supp}(Q) = \{\bar{X}\bar{Y}, 0\}$, the assumption **(A2)** is satisfied for $(P^{\otimes k}, Q)$, and we can show that $\mathbb{P}[(\mathbf{B}_{XY})_a x = (\mathbf{b}_{XY})_a] < 1 - \delta$ for all $x \in R^{V_{2n}}$ and $a \in R'$, where $\delta > 0$ does not depend on x or a. Therefore, the event that one of the $q^{o(n)}$ solutions to $(\mathbf{B}'_{XY})_{2\text{-core}}x = (\mathbf{b}'_{XY})_{2\text{-core}}$ also satisfies $(\mathbf{B}_{XY})_a x = (\mathbf{b}_{XY})_a$ for all $a \in R'$ has probability $q^{o(n)}(1-\delta)^{\Omega(n)}$. Hence,

$$\mathbb{P}[\exists x \in R^n : (\mathbf{B}_{XY})_{2\text{-core}}x = (\mathbf{b}_{XY})_{2\text{-core}}] \leq q^{o(n)}(1-\delta)^{\Omega(n)} = o(1).$$

Since the subsystem $(\mathbf{B}_{XY})_{2\text{-core}}x = (\mathbf{b}_{XY})_{2\text{-core}}$ is unsatisfiable a.a.s., it follows that $\mathbf{B}_{XY}x = \mathbf{b}_{XY}$ is unsatisfiable a.a.s. □

Chapter 7

Proofs of Theorems 2.0.3 and 2.0.4 and Lemmas 6.1.3 and 6.1.4

We begin this chapter by proving the first statement of Theorem 2.0.4, as stated in the following lemma. Let $\mathcal{S}_n$ be the set of simple factor graphs on n variable vertices.

Lemma 7.0.1. *Let $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ be a sequence of sets of factor graphs. Then*

$$\lim_{n\to\infty}\left|\mathbb{P}[G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}}\in\mathcal{E}_n]-\mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}\in\mathcal{E}_n]\right|=0.$$

The distributions of $G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ and $G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ are the same when conditioned on having a fixed degree sequence, so the only difficulty lies in showing that the distributions of the degree sequences of $G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ and $G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ are similar. The degree sequence of $G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ has the same distribution as that of $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$, but conditioning on the event $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}\in\mathcal{S}_n$ in the definition of $G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ may change the distribution slightly. We use a method of moments argument similar to [13, Lemma 4.3] to calculate the probability that $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}\in\mathcal{S}_n$ given that $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ has a typical degree sequence, and then use this to show that conditioning on $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}\in\mathcal{S}_n$ does not significantly change the degree sequence distribution.

Given Lemma 7.0.1, the following lemma completes the proof of the first statement in Theorem 2.0.3.

Lemma 7.0.2. *Suppose $\boldsymbol{d}$ is a Poisson random variable. Then $G_{n,m,\boldsymbol{k}}\triangleleft\triangleright G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$.*

The statements in Theorems 2.0.4 and 2.0.3 about graphs with a random number of constraint vertices will be proved in Section 7.3. In Section 7.4, we use Theorem 7.3.5 and adaptations of arguments from Section 7.2 to prove Lemmas 6.1.3 and 6.1.4. The following local limit theorem will be the main tool used in all of the following subsections.

Theorem 7.0.3 ([23]). *Suppose that $\{X_i\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. integer valued random variables such that the greatest common divisor of the support of X_i is one, and such that $\mathbb{E}[X_i^2] = \sigma^2 \in (0,\infty)$. Then*

$$\lim_{n\to\infty} \sup_{j\in\mathbb{Z}} \left| \sqrt{n}\mathbb{P}\left[\sum_{i=1}^{n} X_i = j\right] - \frac{e^{\frac{-(j-n\mathbb{E}[X_i])^2}{2n\sigma^2}}}{\sqrt{2\pi\sigma^2}} \right| = 0.$$

7.1 Proof of Lemma 7.0.1

Recall from (2.1) that $\mathcal{D}_n$ is the event

$$\sum_{i=1}^{n} \boldsymbol{d}'_{v_i} = \sum_{i=1}^{\boldsymbol{m}'} \boldsymbol{k}'_{a_i},$$

where $\{\boldsymbol{k}_{a_i}\}_{i\in\mathbb{N}}$, $\{\boldsymbol{d}_{v_i}\}_{i\in\mathbb{N}}$, $\boldsymbol{m}'$ are mutually independent, such that $\{\boldsymbol{k}_{a_i}\}_{i\in\mathbb{N}}$ are i.i.d. copies of $\boldsymbol{k}$, $\{\boldsymbol{d}_{v_i}\}_{i\in\mathbb{N}}$ are i.i.d. copies of $\boldsymbol{d}$, and $\boldsymbol{m}'$ is Possion with mean dn/k. We begin by estimating the probability of the event $\mathcal{D}_n$.

Lemma 7.1.1. *If* $\mathrm{Var}(\boldsymbol{k}) = \mathrm{Var}(\boldsymbol{d}) = 0$, *then* $\mathbb{P}[\mathcal{D}_n \mid \boldsymbol{m}' = m] = 1$. *Otherwise* $\mathbb{P}[\mathcal{D}_n \mid \boldsymbol{m}' = m] = \Theta(n^{-1/2})$.

Proof. If $\mathrm{Var}(\boldsymbol{k}) = \mathrm{Var}(\boldsymbol{d}) = 0$, then $nd = m(n)k$ and we have $\mathbb{P}[\mathcal{D}_n \mid \boldsymbol{m}' = m] = 1$.

If $\mathrm{Var}(\boldsymbol{k}) > 0$, then by Theorem 7.0.3,

$$\begin{aligned}
\mathbb{P}[\mathcal{D}_n \mid \boldsymbol{m}' = m] &\geq \sum_{j:|nd-j|\leq\sqrt{n}} \mathbb{P}[\sum_{a\in C_m} \boldsymbol{k}'_a = j]\mathbb{P}[\sum_{v\in V_n} \boldsymbol{d}'_v = j] \\
&= \Omega(1/\sqrt{n})\mathbb{P}[|\sum_{v\in V_n} \boldsymbol{d}'_v - dn| < \sqrt{n}] \\
&= \Omega(1/\sqrt{n}),
\end{aligned}$$

and the uniform upper bound from Theorem 7.0.3 gives

$$\sum_{j=0}^{\infty} \mathbb{P}[\sum_{a\in C_m} \boldsymbol{k}'_a = j]\mathbb{P}[\sum_{v\in V_n} \boldsymbol{d}'_v = j] = O(1/\sqrt{n})$$

Therefore $\mathbb{P}[\mathcal{D}] = \Theta(1/\sqrt{n})$. A similar argument applies in the case where $\mathrm{Var}(\boldsymbol{d}) > 0$. □

Let $\mathcal{M}'_n$ be the event that $\{\boldsymbol{k}'_a\}_{a\in C_{\boldsymbol{m}'}}$, $\{\boldsymbol{d}'_v\}_{v\in V_n}$ satisfies

$$\max_{a\in C_{\boldsymbol{m}'}} \boldsymbol{k}'_a + \max_{v\in V_n} \boldsymbol{d}'_v \leq \sqrt{n}/\log n. \tag{7.1}$$

Since $\mathbb{E}[\boldsymbol{k}^{2+\varepsilon}], \mathbb{E}[\boldsymbol{d}^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$, we have $\mathbb{P}[\mathcal{M}'_n \mid \boldsymbol{m}' = m] = 1 - o(1)$. We also need the following conditional probability.

Lemma 7.1.2. $\mathbb{P}[\mathcal{M}'_n \mid \mathcal{D}_n, \boldsymbol{m}' = m] = 1 - o(1)$

Proof. First note that, since $\mathbb{E}[\boldsymbol{k}^{2+\varepsilon}], \mathbb{E}[\boldsymbol{d}^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$, we have $\mathbb{P}[\boldsymbol{k} > \sqrt{n}/(2\log n)] = o(1/n)$ and $\mathbb{P}[\boldsymbol{d} > \sqrt{n}/(2\log n)] = o(1/n)$. We show that

$$\mathbb{P}[\boldsymbol{k}'_a > \sqrt{n}/(2\log n) \mid \mathcal{D}_n, \boldsymbol{m}' = m] = o(1/n) \tag{7.2}$$

for all $a \in C_m$, and a similar argument shows that $\mathbb{P}[\boldsymbol{d}'_v > \sqrt{n}/(2\log n) \mid \mathcal{D}_n, \boldsymbol{m}' = m] = o(1/n)$ for all $v \in V_n$. The claim then follows by a union bound.

If $\mathrm{Var}(\boldsymbol{k}) = 0$, then clearly $\mathbb{P}[\boldsymbol{k}'_a > \sqrt{n}/(2\log n) \mid \mathcal{D}_n, \boldsymbol{m}' = m] = o(1/n)$. Suppose $\mathrm{Var}(\boldsymbol{k}) > 0$. If $\mathrm{Var}(\boldsymbol{d}) = 0$, then by Theorem 7.0.3

$$\begin{aligned}\mathbb{P}[\{\boldsymbol{k}'_a > \sqrt{n}/(2\log n)]\} \cap \mathcal{D}_n \mid \boldsymbol{m}' = m] &= \sum_{j=\sqrt{n}/(2\log n)}^{\infty} \mathbb{P}[\boldsymbol{k}'_a = j]\mathbb{P}\left[\sum_{a' \in C_m \setminus \{a\}} \boldsymbol{k}'_a = dn - j\right] \\ &= O(1/\sqrt{n})\mathbb{P}[\boldsymbol{k} > \sqrt{n}/(2\log n)].\end{aligned}$$

Similarly, if $\mathrm{Var}(\boldsymbol{d}) > 0$, then by Theorem 7.0.3

$$\begin{aligned}\mathbb{P}[\{\boldsymbol{k}'_a > \sqrt{n}/(2\log n)\} \cap \mathcal{D}_n \mid \boldsymbol{m}' = m] &= \sum_{j=0}^{\infty} \mathbb{P}\left[\boldsymbol{k}'_a > \sqrt{n}/(2\log n), \sum_{a \in C_m} \boldsymbol{k}'_a = j\right] \mathbb{P}[\sum_{v \in V_n} \boldsymbol{d}'_v = j] \\ &= O(1/\sqrt{n})\mathbb{P}[\boldsymbol{k} > \sqrt{n}/(2\log n)]\end{aligned}$$

By Lemma 7.1.1, $\mathbb{P}[\mathcal{D}_n] = \Theta(1/\sqrt{n})$, and so in either case

$$\mathbb{P}[\boldsymbol{k}'_a > \sqrt{n}/(2\log n) \mid \mathcal{D}_n, \boldsymbol{m}' = m] = O(1)\mathbb{P}[\boldsymbol{k} > \sqrt{n}/(2\log n)] = o(1/n),$$

which verifies 7.2. □

As a consequence of Lemmas 7.1.1 and 7.1.2, we have

$$\mathbb{P}[\mathcal{D}_n \mid \mathcal{M}'_n, \boldsymbol{m}' = m] = \begin{cases} 1 - o(1) & \mathrm{Var}(\boldsymbol{k}) = \mathrm{Var}(\boldsymbol{d}) = 0 \\ \Theta(1/\sqrt{n}) & \text{Otherwise.} \end{cases} \tag{7.3}$$

The proof of the following lemma is the same as [13, Claim 4.4], except that (7.3) is used in place of [13, Lemma 4.2].

Lemma 7.1.3. *Conditioned on $\mathcal{M}'_n \cap \mathcal{D}_n \cap \{\boldsymbol{m}' = m\}$, we have $\frac{1}{n}\sum_{v \in V_n} \boldsymbol{d}'^2_v \to \mathbb{E}[\boldsymbol{d}^2]$ and $\frac{1}{n}\sum_{a \in C_m} \boldsymbol{k}'^2_a \to d\mathbb{E}[\boldsymbol{k}^2]/k$ in probability.*

Theorem 7.0.3 also shows that $\frac{1}{n}\sum_{v \in V_n} \boldsymbol{d}'_v \to \mathbb{E}[\boldsymbol{d}]$ in probability conditioned on $\mathcal{D}_n \cap \{\boldsymbol{m}' = m\}$. Therefore there is a function $r(n) = o(1)$ such that the degree sequence

$\{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}$ of $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ satisfies

$$\max_{a\in C_{\boldsymbol{m}}} \boldsymbol{k}_a + \max_{v\in V_n} \boldsymbol{d}_v \le \sqrt{n}/\log n, \tag{7.4}$$

$$\left|\frac{1}{n}\sum_{v\in V_n} \boldsymbol{d}_v - \mathbb{E}[\boldsymbol{d}]\right|, \left|\frac{1}{n}\sum_{v\in V_n} \boldsymbol{d}_v^2 - \mathbb{E}[\boldsymbol{d}^2]\right|, \left|\frac{1}{n}\sum_{a\in C_m} \boldsymbol{k}_a^2 - d\mathbb{E}[\boldsymbol{k}^2]/k\right| \le r(n), \tag{7.5}$$

a.a.s. Let $\mathcal{M}_n$ be this event.

To calculate the probability that $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ is simple, we show that the number of parallel edges of $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ is asymptotically Poisson on $\mathcal{M}_n$. We use a method of moments argument similar to the one used in [13].

Lemma 7.1.4. *Let Y be the number of parallel edges in $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ and let*

$$\lambda = \frac{(\mathbb{E}[\boldsymbol{d}^2]-d)(\mathbb{E}[\boldsymbol{k}^2]-k)}{2dk}.$$

Then for all $\ell \ge 1$, on $\mathcal{M}_n$ we have

$$\mathbb{E}\left[\prod_{i=0}^{\ell-1}(Y-i) \,\middle|\, \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}\right] = \lambda^\ell + o(1). \tag{7.6}$$

Proof. Let U_ℓ be the set of ℓ-tuples $(c_i,u_i)_{i\in[\ell]} \in (C_m\times V_n)^\ell$ such that $k_{c_i}, d_{u_i} \ge 2$ for all $i\in[\ell]$ and such that $(c_1,u_1),\dots,(c_\ell,u_\ell)$ are distinct. For $(c_i,u_i)_{i\in[\ell]}\in U_\ell$, let $Y[(c_i,u_i)_{i\in[\ell]}]$ be the indicator of the event that there are exactly two edges between c_i and u_i in $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ for all $i\in[\ell]$. Let Z be the number of ℓ-tuples of distinct multiedges in $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ where at least one multiedge consists of more than two edges. We then have

$$\prod_{i=0}^{\ell-1}(Y-i) = \sum_{(c_i,u_i)_{i\in[\ell]}\in U_\ell} Y[(c_i,u_i)_{i\in[\ell]}] + Z.$$

Standard arguments show that $\mathbb{E}[Z \mid \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}] = o(1)$. For $(c_i,u_i)_{i\in[\ell]}\in U_\ell$, we have

$$\mathbb{E}[Y[(c_i,u_i)_{i\in[\ell]}] \mid \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}] = \frac{\left(\sum_{v\in V_n}\boldsymbol{d}_v - 2\ell\right)!}{\left(\sum_{v\in V_n}\boldsymbol{d}_v\right)!}\prod_{i\in[\ell]} 2\binom{\boldsymbol{d}_{u_i}}{2}\binom{\boldsymbol{k}_{c_i}}{2}$$

where $\prod_{i\in[\ell]} 2\binom{\boldsymbol{d}_{u_i}}{2}\binom{\boldsymbol{k}_{c_i}}{2}$ counts the number of ways to match 2 vertex copies in $\{c_i\}\times[k_{c_i}]$ with 2 vertex copies in $\{u_i\}\times[\boldsymbol{d}_{v_i}]$ for $i\in[\ell]$, and $\frac{\left(\sum_{v\in V_n}\boldsymbol{d}_v - 2\ell\right)!}{\left(\sum_{v\in V_n}\boldsymbol{d}_v\right)!}$ is the probability that all of these matchings are contained in a uniform random matching Γ. On the event (7.4), summing over $(c_i,u_i)_{i\in[\ell]}\in U_\ell$ gives

$$\sum_{(c_i,u_i)_{i\in[\ell]}\in U_\ell}\prod_{i\in[\ell]} 2\binom{\boldsymbol{d}_{u_i}}{2}\binom{\boldsymbol{k}_{c_i}}{2} = \left(\frac{1}{2}\sum_{a\in C_m, v\in V_n}\boldsymbol{d}_v(\boldsymbol{d}_v-1)\boldsymbol{k}_a(\boldsymbol{k}_a-1)\right)^\ell + o(n^{2\ell}),$$

and on the event (7.5),

$$\frac{\left(\sum_{v\in V_n} \boldsymbol{d}_v - 2\ell\right)!}{\left(\sum_{v\in V_n} \boldsymbol{d}_v\right)!} \sim \frac{1}{(dn)^{2\ell}}, \qquad \sum_{a\in C_m, v\in V_n} \boldsymbol{d}_v(\boldsymbol{d}_v-1)\boldsymbol{k}_a(\boldsymbol{k}_a-1) \sim \frac{dn^2}{k}(\mathbb{E}[\boldsymbol{d}^2]-d)(\mathbb{E}[\boldsymbol{k}^2]-k).$$

Therefore

$$\mathbb{E}\left[\sum_{(c_i,u_i)_{\in[\ell]}\in U_\ell} Y[(c_i,u_i)_{\in[\ell]}] \mid \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}\right] \sim \left(\frac{(\mathbb{E}[\boldsymbol{d}^2]-d)(\mathbb{E}[\boldsymbol{k}^2]-k)}{2dk}\right)^\ell = \lambda^\ell.$$

Since $\mathbb{E}[Z \mid \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}] = o(1)$, this proves (7.6). □

Lemma 7.1.5. *On $\mathcal{M}_n$, we have $\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n \mid \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}] \sim e^{-\lambda}$. As a consequence, $\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] \sim e^{-\lambda}$.*

Proof. By Lemma 7.1.4 and the method of moments (eg. [10, Theorem 1.22]),

$$\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n \mid \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}] = \mathbb{P}[Y=0 \mid \{\boldsymbol{k}_a\}_{a\in C_m}, \{\boldsymbol{d}_v\}_{v\in V_n}] \sim e^{-\lambda}$$

on $\mathcal{M}_n$. Since $\mathbb{P}[\mathcal{M}_n] = 1-o(1)$, we also get $\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] \sim e^{-\lambda}$. □

We can now prove Lemma 7.0.1.

Proof of Lemma 7.0.1. Let G be a simple factor graph with degree sequence $\{k_a\}_{a\in C_m}$, $\{d_v\}_{v\in V_n}$. Given a fixed degree sequence, the conditional distributions of $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \mid M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n$ and $G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ are both uniform on simple factor graphs with the given degree sequence. Therefore

$$\begin{aligned}\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} = G \mid M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] \\ = \frac{\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n \mid \{\boldsymbol{k}_a\} = \{k_a\}, \{\boldsymbol{d}_v\} = \{d_v\}]}{\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]} \mathbb{P}[G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}} = G].\end{aligned}$$

Therefore Lemma 7.1.5 shows that on $\mathcal{M}_n$,

$$\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n \mid \{\boldsymbol{k}_a\}, \{\boldsymbol{d}_v\}] \sim \mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]$$

Therefore, for any sequence of sets of graphs $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$,

$$\left|\mathbb{P}[\{M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n\} \cap \mathcal{M}_n \mid M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] - \mathbb{P}[\{G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n\} \cap \mathcal{M}_n]\right| = o(1).$$

By Lemma 7.1.5, $\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] = \Omega(1)$, and so $\mathbb{P}[\mathcal{M}_n \mid M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] = 1-o(1)$. Therefore $\mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] = \mathbb{P}[G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] + o(1)$. □

7.2 Proof of Lemma 7.0.2

We first prove the following asymptotic equivalence statement.

Lemma 7.2.1. *For any sequence of sets of factor graphs $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$, we have*

$$\lim_{n\to\infty} |\mathbb{P}[G_{n,m,\boldsymbol{k}} \in \mathcal{E}_n] - \mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{E}_n \mid M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n]| = 0. \tag{7.7}$$

Moreover, $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n] = \Omega(1)$.

Proof. Let $\hat{k} = \mathbb{E}[(\boldsymbol{k}-1)\mathbf{1}_{\{\boldsymbol{k}\geq 1\}}]$, and let $\boldsymbol{m}(j) = |\{a \in C_m : \boldsymbol{k}'_a = j\}|$ for $j \in \mathbb{N}$. Let $\mathcal{T}_n$ be the event

$$|\sum_{a\in C_m} (\boldsymbol{k}'_a - 1)\mathbf{1}_{\{\boldsymbol{k}'_a\geq 1\}} - \hat{k}dn/k| < n^{2/3}, \quad \max_{a\in C_m} \boldsymbol{k}'_a \leq \sqrt{n}/\log n.$$

Since $m = dn/k + O(\sqrt{n})$ and $\mathbb{E}[\boldsymbol{k}^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$, we have $\mathbb{P}[\mathcal{T}_n] = 1 - o(1)$. For each $a \in C_m$, the probability that a is not incident to any parallel edges in $M_{n,m,\boldsymbol{k}}$ is $\prod_{j=1}^{\boldsymbol{k}_a-1}(1-\frac{j}{n})$. Hence, on $\mathcal{T}_n$,

$$\begin{aligned}\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n \mid \{\boldsymbol{k}'_a\}_{a\in C_m}] &= \prod_{a\in C_m}\prod_{j=2}^{\boldsymbol{k}'_a}(1-\frac{j-1}{n}) \\ &= \prod_{j=2}^{\sqrt{n}/\log n}(1-\frac{j-1}{n})^{\boldsymbol{m}(j)}.\end{aligned}$$

Using the inequality $(1-1/n)^n < e^{-1} < (1-1/n)^{n-1}$, for $2 \leq j \leq \sqrt{n}/\log n$,

$$\left(1-\frac{1}{\sqrt{n}\log n}\right)^{\frac{\boldsymbol{m}(j)}{\sqrt{n}\log n}} e^{-\frac{(j-1)\boldsymbol{m}(j)}{n}} < \left(1-\frac{j-1}{n}\right)^{\boldsymbol{m}(j)} < e^{-\frac{(j-1)\boldsymbol{m}(j)}{n}},$$

and hence

$$\left(1-\frac{1}{\sqrt{n}\log n}\right)^{\frac{m}{\sqrt{n}\log n}} e^{-\frac{\sum_{a\in C_m}(\boldsymbol{k}'_a-1)\mathbf{1}_{\{\boldsymbol{k}'_a\geq 1\}}}{n}} < \prod_{j=1}^{\sqrt{n}/\log n}(1-\frac{j}{n})^{m(j+1)} < e^{-\frac{\sum_{a\in C_m}(\boldsymbol{k}'_a-1)\mathbf{1}_{\{\boldsymbol{k}'_a\geq 1\}}}{n}}.$$

Since $\mathbb{E}[\boldsymbol{k}^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$, we have

$$\frac{\sum_{a\in C_m}(\boldsymbol{k}'_a-1)\mathbf{1}_{\{\boldsymbol{k}'_a\geq 1\}}}{n} \to d\hat{k}/k$$

in probability. Therefore $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n \mid \{\boldsymbol{k}'_a\}_{a\in C_m}] = (1-o(1))e^{-d\hat{k}/k}$ on $\mathcal{T}_n$. Since $\mathbb{P}[\mathcal{T}_n] = 1-o(1)$, we also have $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n] = (1-o(1))e^{-d\hat{k}/k}$.

We now prove (7.7). Note that conditioned on a fixed degree sequence, $G_{n,m,\boldsymbol{k}}$ and $M_{n,m,\boldsymbol{k}} \mid M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n$ are identically distributed. Therefore, for any simple graph G with degree sequence $\{k_a\}_{a\in C_m}$, we have

$$\mathbb{P}[M_{n,m,\boldsymbol{k}} = G \mid M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n] = \frac{\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n \mid \{\boldsymbol{k}'_a\}_{a\in C_m} = \{k_a\}_{a\in C_m}]}{\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n]}\mathbb{P}[G_{n,m,\boldsymbol{k}} = G],$$

and on $\mathcal{T}_n$,

$$\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n \mid \{\boldsymbol{k}'_a\}_{a\in C_m}] \sim \mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n].$$

Therefore for all sequences of sets of graphs $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$,

$$\mathbb{P}[\{M_{n,m,\boldsymbol{k}} \in \mathcal{E}_n\} \cap \mathcal{T}_n \mid M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n] = (1 - o(1))\mathbb{P}[\{G_{n,m,\boldsymbol{k}} \in \mathcal{E}_n\} \cap \mathcal{T}_n].$$

We have $\mathbb{P}[\mathcal{T}_n] = 1 - o(1)$, and since $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n] = \Theta(1)$ we also have $\mathbb{P}[\mathcal{T}_n \mid M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n] = 1 - o(1)$. Therefore $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{E}_n \mid M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n] = \mathbb{P}[G_{n,m,\boldsymbol{k}} \in \mathcal{E}_n] + o(1)$. □

Lemma 7.2.2. *Let $Z_n \sim \mathrm{Po}(dn)$ be independent of $M_{n,m,\boldsymbol{k}}$. Then $\{M_{n,m,\boldsymbol{k}}\}_{n\in\mathbb{N}} \triangleleft\triangleright \{M_{n,m,\boldsymbol{k}} \mid e(M_{n,m,\boldsymbol{k}}) = Z_n\}_{n\in\mathbb{N}}$.*

Proof. Let $\mathcal{T}_n(C)$ be the event $\{|e(M_{n,m,\boldsymbol{k}}) - dn| \le C\sqrt{n}\}$. We prove the following two claims.

Claim 7.2.3.

$$\lim_{C\to\infty} \liminf_{n\to\infty} \mathbb{P}[\mathcal{T}_n(C)] = \lim_{C\to\infty} \liminf_{n\to\infty} \mathbb{P}[\mathcal{T}_n(C) \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = 1.$$

Claim 7.2.4. *Suppose G_n is a sequence of factor graphs such that $e(G_n) = dn + O(\sqrt{n})$. Then $\mathbb{P}[M_{n,m,\boldsymbol{k}} = G_n \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = \Theta(1)\mathbb{P}[M_{n,m,\boldsymbol{k}} = G_n]$.*

Assuming Claim 7.2.4 holds, for any sequence of sets of graphs $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$,

$$\mathbb{P}[\{M_{n,m,\boldsymbol{k}} \in \mathcal{E}_n\} \cap \mathcal{T}_n(C) \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = \Theta(1)\mathbb{P}[\{M_{n,m,\boldsymbol{k}} \in \mathcal{E}_n\} \cap \mathcal{T}_n(C)].$$

Assuming Claim 7.2.3 holds, it follows that $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{E}_n \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = o(1)$ if and only if $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{E}_n] = o(1)$.

Proof of Claim 7.2.3. If $\mathrm{Var}(\boldsymbol{k}) = 0$, then $e(M_{n,m,\boldsymbol{k}}) = km = dn + O(\sqrt{n})$, and so

$$\mathbb{P}[T_n(C)] = \mathbb{P}[\mathcal{T}_n(C) \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = 1$$

for sufficiently large n and C.

Now suppose $\mathrm{Var}(k) > 0$. Since $e(M_{n,m,\boldsymbol{k}})$ is a sum of $m = dn/k + O(\sqrt{n})$ i.i.d. random variables, we have $\lim_{C\to\infty} \liminf_{n\to\infty} \mathbb{P}[T_n(C)] = 1$ by the CLT.

For the conditional probability, Lemma 7.1.1 with $\boldsymbol{d} \sim \mathrm{Po}(d)$ gives $\mathbb{P}[\sum_{a\in C_m} \boldsymbol{k}'_a = \sum_{v\in V_n} \boldsymbol{d}'_v] = \Theta(1/\sqrt{n})$. Since $Z_n \sim \mathrm{Po}(dn)$ is independent of $M_{n,m,\boldsymbol{k}}$, we have $(e(M_{n,m,\boldsymbol{k}}), Z_n) \sim (\sum_{a\in C_m} \boldsymbol{k}'_a, \sum_{v\in V_n} \boldsymbol{d}'_v)$, and so $\mathbb{P}[e(M_{n,m,\boldsymbol{k}}) = Z_n] = \Theta(1/\sqrt{n})$. Therefore

$$\mathbb{P}[\mathcal{T}_n(C) \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = 1 - \Theta(\sqrt{n}) \sum_{\substack{j\in\mathbb{N} \\ |j-dn|\geq C\sqrt{n}}} \mathbb{P}[e(M_{n,m,\boldsymbol{k}}) = j]\mathbb{P}[j = Z_n].$$

By Theorem 7.0.3 applied to the sum $\sum_{a\in C_m} \boldsymbol{k}'_a = e(M_{n,m,\boldsymbol{k}})$, we have

$$\sup\{\mathbb{P}[e(M_{n,m,\boldsymbol{k}}) = j] : j \in \mathbb{N}\} = \Theta(1/\sqrt{n}),$$

and so

$$\mathbb{P}[\mathcal{T}_n(C) \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = 1 - \Theta(1)\mathbb{P}[|Z_n - dn| \geq C\sqrt{n}].$$

Since $Z_n \sim \mathrm{Po}(dn)$, we have $\lim_{C\to\infty} \mathbb{P}[|Z_n - dn| \geq C\sqrt{n}] = 0$. Therefore

$$\lim_{C\to\infty} \liminf_{n\to\infty} \mathbb{P}[\mathcal{T}_n(C) \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = 1.$$

□

Proof of Claim 7.2.4. If $e(G_n) = dn + O(\sqrt{n})$, then $\mathbb{P}[e(G_n) = Z_n] = \Theta(1/\sqrt{n})$ by Theorem 7.0.3. Since $\mathbb{P}[e(M_{n,m,\boldsymbol{k}}) = Z_n] = \Theta(1/\sqrt{n})$, we then have

$$\begin{aligned}\mathbb{P}[M_{n,m,\boldsymbol{k}} = G_n \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] &= \frac{\mathbb{P}[M_{n,m,\boldsymbol{k}} = G_n]\mathbb{P}[e(G_n) = Z_n]}{\mathbb{P}[e(M_{n,m,\boldsymbol{k}}) = Z_n]} \\ &= \Theta(1)\mathbb{P}[M_{n,m,\boldsymbol{k}} = G_n].\end{aligned}$$

□

□

Nest, we relate the random graphs $M_{n,m,\boldsymbol{k}}$ and $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$. We need the following definition, which relates the distributions of the degree sequence of $M_{n,m,\boldsymbol{k}}$.

Definition 7.2.5. *Let S be a finite set and fix $Z \geq 0$. Let $\{X_i\}_{i\in[Z]}$ be i.i.d. uniform random variables on S. Define the distribution* $\mathrm{Mult}(S, Z)$ *on $\mathbb{Z}^S$ to be the distribution of $\{|\{i \in [Z] : X_i = s\}|\}_{s\in S}$.*

The following claim is used to compare the distributions of the degree sequences of $M_{n,m,\boldsymbol{k}}$ and $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$.

Claim 7.2.6. *Let S be a finite set and let $Z \geq 0$. Let $\{\boldsymbol{g}_s\}_{s\in S}$ be i.i.d* $\mathrm{Po}(\lambda)$ *random variables for some $\lambda > 0$. Then the conditional distribution $\{\boldsymbol{g}_s\}_{s\in S} \mid \{\sum_{s\in S} \boldsymbol{g}_s = Z\}$ is* $\mathrm{Mult}(S, Z)$.

Proof. Let $\{h_s\}_{s\in S} \in \mathbb{N}^S$ be such that $\sum_{s\in S} h_s = Z$. Then

$$\mathbb{P}[\mathrm{Mult}(S,Z) = \{h_s\}_{s\in S}] = \binom{Z}{h_s, s\in S} n^{-Z} = \frac{Z!}{\prod_{s\in S} h_s!}|S|^{-Z}.$$

On the other hand, since $\sum_{s\in S} \boldsymbol{g}_s \sim \mathrm{Po}(\lambda|S|)$, we have

$$\mathbb{P}[\{\boldsymbol{g}_s\}_{s\in S} = \{h_s\}_{s\in S} \mid \sum_{s\in S} \boldsymbol{g}_s = Z] = \prod_{s\in S} \frac{\lambda^{h_s} e^{-\lambda}}{h_s!} \cdot \frac{Z!}{(\lambda|S|)^Z e^{-\lambda|S|}} = \frac{Z!}{\prod_{s\in S} h_s!}|S|^{-Z}.$$

Therefore $\{\boldsymbol{g}_s\}_{s\in S} \mid \{\sum_{s\in S} \boldsymbol{g}_s = Z\} \sim \mathrm{Mult}(S,Z)$. □

Recall that we denote the degree sequence of a factor graph G by $(d(G), k(G)) = (\{d_v\}_{v\in V_n}, \{k_a\}_{a\in C_m})$ and the number of edges by $e(G)$. For any $a \in C_m$, $v \in V_n$, we denote by $e(a,v)$ the number of edges between a and v.

Lemma 7.2.7. *Suppose $\boldsymbol{d}$ is a Poisson random variable, and let $Z_n \sim \mathrm{Po}(dn)$ be independent of $M_{n,m,\boldsymbol{k}}$. Then $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$ and $M_{n,m,\boldsymbol{k}} \mid e(M_{n,m,\boldsymbol{k}}) = Z_n$ are identically distributed.*

Proof. Let G be a bipartite graph on $C_m \cup V_n$. Since Z_n is independent of $M_{n,m,\boldsymbol{k}}$, we have

$$\begin{aligned}\mathbb{P}[M_{n,m,\boldsymbol{k}} = G \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = \mathbb{P}[M_{n,m,\boldsymbol{k}} = G \mid k(M_{n,m,\boldsymbol{k}}) = k(G), d(M_{n,m,\boldsymbol{k}}) = d(G)] \\ \times \mathbb{P}[d(M_{n,m,\boldsymbol{k}}) = d(G) \mid k(M_{n,m,\boldsymbol{k}}) = k(G)] \\ \times \mathbb{P}[k(M_{n,m,\boldsymbol{k}}) = k(G) \mid e(M_{n,m,\boldsymbol{k}}) = Z_n].\end{aligned}$$

Therefore it suffices to prove the following three claims.

Claim 7.2.8.

$$\begin{aligned}&\mathbb{P}[M_{n,m,\boldsymbol{k}} = G \mid k(M_{n,m,\boldsymbol{k}}) = k(G), d(M_{n,m,\boldsymbol{k}}) = d(G)] \\ &\quad = \mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} = G \mid k(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) = k(G), d(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) = d(G)]\end{aligned}$$

Proof. We first calculate $\mathbb{P}[M_{n,m,\boldsymbol{k}} = G \mid k(M_{n,m,\boldsymbol{k}}) = k(G)]$. The number of ways to choose $(u_1, \dots, u_{\deg_G(a)})$ so that $\{au_i : i \in [\deg_G(a)]\}$ is the multi set of edges incident to a in G is $\deg_G(a)! / \prod_{v\in V_n} e(a,v)!$, and the probability of making each of these choices is $n^{-\deg_G(a)}$. Since $\{(u_1, \dots, u_{\deg_G(a)})\}_{a\in C_m}$ are chosen independently, we get

$$\mathbb{P}[M_{n,m,\boldsymbol{k}} = G \mid k(M_{n,m,\boldsymbol{k}}) = k(G)] = \frac{\prod_{a\in C_m} \deg_G(a)!}{n^{e(G)} \prod_{a\in C_m, v\in V_n} e(a,v)!}.$$

For $\mathbb{P}[d(M_{n,m,\boldsymbol{k}}) = d(G) \mid k(M_{n,m,\boldsymbol{k}}) = k(G)]$, the number of ways to choose $\{(v_1, \dots, v_{\deg_G(a)})\}_{a\in C_m}$ so that $d(M_{n,m,\boldsymbol{k}}) = d(G)$ is $e(G)! / \prod_{v\in V_n} \deg_G(v)!$, so we have

$$\mathbb{P}[d(M_{n,m,\boldsymbol{k}}) = d(G) \mid k(M_{n,m,\boldsymbol{k}}) = k(G)] = \frac{e(G)!}{n^{e(G)} \prod_{v\in V_n} \deg_G(v)!}.$$

Therefore

$$\mathbb{P}[M_{n,m,\boldsymbol{k}} = G \mid k(M_{n,m,\boldsymbol{k}}) = k(G), d(M_{n,m,\boldsymbol{k}}) = d(G)] = \frac{\prod_{a\in C_m} \deg_G(a)! \prod_{v\in V_n} \deg_G(v)!}{e(G)! \prod_{a\in C_m, v\in V_n} e(a,v)!}.$$

Now consider $M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}$, which is chosen according to the pairing model. Let $c(a) = \{a\} \times [\deg_G(a)]$ and $c(v) = \{v\} \times [\deg_G(v)]$ be the sets of vertex copies of $a \in C_m$ and $v \in V_n$. Let $\mathcal{G}$ be the set of matchings Γ between $\bigcup_{a\in C_m} c(a)$ and $\bigcup_{v\in V_n} c(v)$ that contain exactly $e(a,v)$ edges between $c(v)$ and $c(v)$ for each $a \in C_m$, $v \in V_n$. Then the set of $\prod_{a\in C_m} \deg_G(a)! \prod_{v\in V_n} \deg_G(v)!$ of permutations of vertex copies acts transitively on $\mathcal{G}$, and every matching in $\mathcal{G}$ is stabilized by the $\prod_{a\in C_m, v\in V_n} e(a,v)!$ permutations that map edges between $c(a)$ and $c(v)$ to other edges between $c(a)$ and $c(v)$ for each $a \in C_m, v \in V_n$. Therefore

$$|\mathcal{G}| = \frac{\prod_{a\in C_m} \deg_G(a)! \prod_{v\in V_n} \deg_G(v)!}{\prod_{a\in C_m, v\in V_n} e(a,v)!}.$$

Since there are $e(G)!$ matchings in total, the claim follows. □

Claim 7.2.9.

$$\mathbb{P}[d(M_{n,m,\boldsymbol{k}}) = d(G) \mid k(M_{n,m,\boldsymbol{k}}) = k(G)] = \mathbb{P}[d(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) = d(G) \mid k(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) = k(G)]$$

Proof. Since the endpoints $\{(u_1, \dots, u_{\boldsymbol{k}'_a})\}_{a\in C_m}$ are chosen independently $M_{n,m,\boldsymbol{k}}$, we have $d(M_{n,m,\boldsymbol{k}}) \mid k(M_{n,m,\boldsymbol{k}}) = k(G) \sim \mathrm{Mult}(V_n, e(G))$. On the other hand, since $\{\boldsymbol{k}'_a\}_{a\in C_m}$ and $\{\boldsymbol{d}'_v\}_{v\in V_n}$ are independent,

$$\begin{aligned} d(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) \mid \{k(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) = k(G)\} &\sim d(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) \mid \{e(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) = e(G)\} \\ &\sim \{\boldsymbol{d}'_v\}_{v\in V_n} \mid \left\{ \sum_{v\in V_n} \boldsymbol{d}'_v = e(G) \right\} \end{aligned}$$

and $\{\boldsymbol{d}'_v\}_{v\in V_n} \mid \sum_{v\in V_n} \boldsymbol{d}'_v = e(G) \sim \mathrm{Mult}(V_n, e(G))$ by Claim 7.2.6. □

Claim 7.2.10.

$$\mathbb{P}[k(M_{n,m,\boldsymbol{k}}) = k(G) \mid e(M_{n,m,\boldsymbol{k}}) = Z_n] = \mathbb{P}[k(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}) = k(G)]$$

Proof. Since $\{\boldsymbol{d}'_v\}_{v\in V_n}$ are i.i.d. $\mathrm{Po}(d)$ random variables independent of $\{\boldsymbol{k}'_a\}_{a\in C_m}$, we have

$$(\{\boldsymbol{k}'_a\}_{c\in C_m}, \sum_{v\in V_n} \boldsymbol{d}'_v) \sim (\{\boldsymbol{k}'_a\}_{a\in C_m}, Z_n).$$

Therefore

$$k(M_{n,m,\boldsymbol{k}}) \mid \{e(M_{n,m,\boldsymbol{k}}) = Z_n\} \sim \{\boldsymbol{k}'_a\}_{a\in C_m} \mid \left\{ \sum_{a\in C_m} \boldsymbol{k}'_a = \sum_{v\in V_n} \boldsymbol{d}'_v \right\} \sim k(M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}}).$$

□

□

We can now complete the proof of Lemma 7.0.2.

Proof of Lemma 7.0.2. Let $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ be a sequence of sets of factor graphs, and let $\mathcal{E}'_n = \mathcal{E}_n \cap \mathcal{S}_n$. We have $\mathbb{P}[G_{n,m,\boldsymbol{k}} \in \mathcal{E}'_n] = \mathbb{P}[G_{n,m,\boldsymbol{k}} \in \mathcal{E}_n]$ and $\mathbb{P}[G'_{n,m,\boldsymbol{k}} \in \mathcal{E}'_n] = \mathbb{P}[G'_{n,m,\boldsymbol{k}} \in \mathcal{E}_n]$. By Lemma 7.2.1,

$$\mathbb{P}[G_{n,m,\boldsymbol{k}} \in \mathcal{E}'_n] = \frac{\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{E}'_n]}{\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n]} + o(1)$$

and $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{S}_n] = \Theta(1)$. Hence $\mathbb{P}[G_{n,m,\boldsymbol{k}} \in \mathcal{E}'_n] = o(1)$ if and only if $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{E}'_n] = o(1)$. By Lemmas 7.2.2 and 7.2.7, $\mathbb{P}[M_{n,m,\boldsymbol{k}} \in \mathcal{E}'_n] = o(1)$ if and only if $\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}'_n] = o(1)$. Finally, by Lemma 7.1.5,

$$\mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}'_n] = \frac{\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}'_n]}{\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]} = \Theta(1)\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}'_n].$$

Hence $\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}'_n] = o(1)$ if and only if $\mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}'_n] = o(1)$. □

7.3 Factor graphs with a random number of constraint vertices

In this section we prove the statements of Theorems 2.0.3 and 2.0.4 about $G_{n,\boldsymbol{m}',\boldsymbol{k}}$, $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$, and $G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$. Note that $G_{n,\boldsymbol{m}',\boldsymbol{k}}$ and $M_{n,\boldsymbol{m}',\boldsymbol{k}}$ have $\boldsymbol{m}'$ constraint vertices, $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ and $G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ have $\boldsymbol{m}$ constraint vertices, but contitioning $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ on having no parallel edges may change the distribution of the number of constraint vertices slightly. Let $\hat{\boldsymbol{m}}$ be the number of constraint vertices in $G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$.

Lemma 7.3.1. $\lim_{C\to\infty} \limsup_{n\to\infty} \mathbb{P}[|\boldsymbol{m} - dn/k| > C\sqrt{n}] = 0$.

Proof. We consider multiple cases based on the variance of $\boldsymbol{k}$ and $\boldsymbol{d}$. If $\mathrm{Var}(\boldsymbol{k}) = \mathrm{Var}(\boldsymbol{d}) = 0$, then $\mathbb{P}[\sum_{i=1}^{m} \boldsymbol{k}'_{a_i} = \sum_{i=1}^{n} d'_{v_i}] = 0$ for $m \neq dn/k$, and so $\mathbb{P}[\boldsymbol{m} = dn/k] = 1$.

Now suppose $\mathrm{Var}(\boldsymbol{d}) = 0$ and $\mathrm{Var}(\boldsymbol{k}) > 0$. Then

$$\mathbb{P}[\{|\boldsymbol{m}' - dn/k| > C\sqrt{n}\} \cap \mathcal{D}] = \sum_{m:|m-dn/k|>C\sqrt{n}} \mathbb{P}[\sum_{i=1}^{m} \boldsymbol{k}'_{a_i} = dn]\mathbb{P}[\boldsymbol{m}' = m],$$

and Theorem 7.0.3 gives

$$\sum_{m:|m-dn/k|>C\sqrt{n}} \mathbb{P}[\sum_{i=1}^{m} \boldsymbol{k}'_{a_i} = dn]\mathbb{P}[\boldsymbol{m}' = m] = O(1/\sqrt{n})\mathbb{P}[|\boldsymbol{m}' - dn/k| > C\sqrt{n}],$$

where the $O(1/\sqrt{n})$ term does not depend on C. Since $\mathrm{Var}(\boldsymbol{k}) > 0$, we also have $\mathbb{P}[\mathcal{D}] = \Omega(1/\sqrt{n})$ by conditioning on $\boldsymbol{m}'$ and applying Lemma 7.1.1. Hence $\mathbb{P}[|\boldsymbol{m} - dn/k| > C\sqrt{n}] = O(1)\mathbb{P}[|\boldsymbol{m}' - dn/k| > C\sqrt{n}]$. Since $\boldsymbol{m}' \sim \mathrm{Po}(dn/k)$, this proves the claim.

Similarly, in the case where $\mathrm{Var}(\boldsymbol{d}) > 0$, Theorem 7.0.3 gives

$$\begin{aligned}\mathbb{P}[\{|\boldsymbol{m}' - dn/k| > C\sqrt{n}\} \cap \mathcal{D}] &= \sum_{j\in\mathbb{N}} \mathbb{P}[\{|\boldsymbol{m}' - dn/k| > C\sqrt{n}\} \cap \{\sum_{i=1}^{\boldsymbol{m}'} \boldsymbol{k}'_{a_i} = j\}]\mathbb{P}[\sum_{i=1}^{n} \boldsymbol{d}'_{v_i} = j] \\ &= O(1/\sqrt{n})\mathbb{P}[\{|\boldsymbol{m}' - dn/k| > C\sqrt{n}\}],\end{aligned}$$

and the claim again follows by Lemma 7.1.1. □

Lemma 7.3.2. *Let $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ be a sequence of sets of integers. Then*

$$\lim_{n\to\infty} \left|\mathbb{P}[\boldsymbol{m} \in \mathcal{E}_n] - \mathbb{P}[\boldsymbol{m} \in \mathcal{E}_n \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]\right| = 0$$

Proof. By conditioning on $\boldsymbol{m}$ and applying Lemma 7.1.5, we have for any $C > 0$,

$$\begin{aligned}|\mathbb{P}[M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] - e^{-\lambda}| &\le \mathbb{P}[|\boldsymbol{m} - dn/k| > C\sqrt{n}] \\ &\quad + \sum_{m:|m-dn/k|\le C\sqrt{n}} \left|\mathbb{P}[M^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] - e^{-\lambda}\right| \mathbb{P}[\boldsymbol{m} = m] \\ &= \mathbb{P}[|\boldsymbol{m} - dn/k| > C\sqrt{n}] + o(1)\mathbb{P}[|\boldsymbol{m} - dn/k| \le C\sqrt{n}].\end{aligned}$$

Hence, by Lemma 7.3.1, $\mathbb{P}[M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] \sim e^{-\lambda}$.

Now let $\varepsilon > 0$. Since $\mathbb{P}[M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] = \Theta(1)$, by Lemma 7.3.1 there is a $C > 0$ such that $\mathbb{P}[|\boldsymbol{m} - dn/k| > C\sqrt{n}] + \mathbb{P}[|\boldsymbol{m} - dn/k| > C\sqrt{n} \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] < \varepsilon$. Hence

$$\begin{aligned}&\left|\mathbb{P}[\boldsymbol{m} \in \mathcal{E}_n] - \mathbb{P}[\boldsymbol{m} \in \mathcal{E}_n \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]\right| \\ &\quad \le \varepsilon + \sum_{m\in\mathcal{E}_n:|m-dn/k|\le C\sqrt{n}} \left(1 - \frac{\mathbb{P}[M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n \mid \boldsymbol{m} = m]}{\mathbb{P}[M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]}\right) \mathbb{P}[\boldsymbol{m} = m]\end{aligned}$$

By Lemma 7.1.5, we have $\sup_{m:|m-dn/k|\le C\sqrt{n}} |\mathbb{P}[M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n \mid \boldsymbol{m} = m] - e^{-\lambda}| = o(1)$. Since $\mathbb{P}[M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] \sim e^{-\lambda}$, this shows that $\left|\mathbb{P}[\boldsymbol{m} \in \mathcal{E}_n] - \mathbb{P}[\boldsymbol{m} \in \mathcal{E}_n \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]\right| = \varepsilon + o(1)$, which proves the claim. □

Lemma 7.3.3. *Suppose $\boldsymbol{d} \sim \mathrm{Po}(d)$ and let $\{\mathcal{C}_n\}$ be a sequence of sets of integers. Then $\mathbb{P}[\boldsymbol{m}' \in \mathcal{C}_n] = o(1)$ if and only if $\mathbb{P}[\boldsymbol{m} \in \mathcal{C}_n]$.*

Proof. By Lemma 7.3.1, it suffices to show that, for all sequences $m(n) = dn/k + O(\sqrt{n})$,

$$\mathbb{P}[\boldsymbol{m} = m(n)] = \frac{\mathbb{P}[\mathcal{D} \mid \boldsymbol{m}' = m(n)]}{\mathbb{P}[\mathcal{D}]}\mathbb{P}[\boldsymbol{m}' = m(n)] = \Theta(1)\mathbb{P}[\boldsymbol{m}' = m(n)]. \tag{7.8}$$

Since $\operatorname{Var}(\boldsymbol{d}) > 0$ and $\gcd(\boldsymbol{d}) = 1$, by Lemma 7.1.1 we have $\mathbb{P}[\mathcal{D} \mid \boldsymbol{m}' = m(n)] = \Theta(1/\sqrt{n})$ for all sequences $m(n) = dn/k + O(\sqrt{n})$. Moreover, by conditioning on $\boldsymbol{m}'$, Lemma 7.1.1 shows that $\mathbb{P}[\mathcal{D}] = \Theta(1/\sqrt{n})$. Therefore (7.8) holds. □

We prove the statements in Theorems 2.0.3 and 2.0.4 related to the models $G_{n,\boldsymbol{m}',\boldsymbol{k}}$, $G^{(1)}_{n,\boldsymbol{k},\boldsymbol{d}}$, and $G^{(2)}_{n,\boldsymbol{k},\boldsymbol{d}}$ in the following two lemmas. This completes the proofs of Theorems 2.0.3 and 2.0.4.

Lemma 7.3.4. *Suppose $\boldsymbol{d}$ is a Poisson random variable. Then $G_{n,\boldsymbol{m}',\boldsymbol{k}} \triangleleft\triangleright G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$.*

Proof. Suppose $\{\mathcal{E}_n\}$ is a sequence of factor graphs such that $\mathbb{P}[G_{n,\boldsymbol{m}',\boldsymbol{k}} \in \mathcal{E}_n] = o(1)$ and let $\varepsilon > 0$. Then there is a sequence of sets of integers $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ such that $\mathbb{P}[\boldsymbol{m}' \in \{\mathcal{C}_n\}] = 1 - o(1)$ and $\mathbb{P}[G_{n,m(n),\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] = o(1)$ for all sequences $m(n)$ such that $m(n) \in \mathcal{C}_n$. By Lemmas 7.3.1 and 7.3.2, there is a $C > 0$ such that $\mathbb{P}[|\boldsymbol{m}' - dn/k| > C\sqrt{n}] < \varepsilon + o(1)$ and $\mathbb{P}[|\boldsymbol{m} - dn/k| > C\sqrt{n} \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{d},\boldsymbol{k}} \in \mathcal{S}_n] < \varepsilon + o(1)$. By Theorem 2.0.4, $\mathbb{P}[G^{(2)}_{n,m(n),\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] = o(1)$ for all sequences $m(n)$ with $m(n) \in \mathcal{C}_n \cap \{m : |m - dn/k| \le C\sqrt{n}\}$ and by Lemmas 7.3.3 and 7.3.2, we have $\mathbb{P}[\boldsymbol{m} \in C_n \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] = 1 - o(1)$. Therefore

$$\begin{aligned}\mathbb{P}[G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] &\le \varepsilon + o(1) + \sum_{m\in\mathcal{C}_n : |m-dn/k|\le C\sqrt{n}} \mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n]\mathbb{P}[\boldsymbol{m} = m \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n] \\ &= \varepsilon + o(1) + o(1)\mathbb{P}[|\boldsymbol{m} - dn/k| \le C\sqrt{n} \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n].\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this shows that $\{G_{n,\boldsymbol{m}',\boldsymbol{k},\boldsymbol{d}}\} \triangleright \{G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}\}$. A similar argument shows that $\{G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}\} \triangleright \{G_{n,\boldsymbol{m}',\boldsymbol{k},\boldsymbol{d}}\}$. □

Lemma 7.3.5. *Let $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ be a sequence of sets of factor graphs. Then*

$$\lim_{n\to\infty} \left|\mathbb{P}[G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] - \mathbb{P}[G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n]\right| = 0.$$

Proof. Let $\varepsilon > 0$. By Lemma 7.3.1, there is a $C > 0$ be such that $\mathbb{P}[|\boldsymbol{m} - dn/k| > C\sqrt{n}] + \mathbb{P}[|\boldsymbol{m}' - dn/k| > C\sqrt{n}] < \varepsilon$. We have

$$\begin{aligned}\left|\mathbb{P}[G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] - \mathbb{P}[G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n]\right| \le &\sum_{\substack{m\in\mathbb{N}\\ |m-dn/k|\le C\sqrt{n}}} \left|\mathbb{P}[G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] - \mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n]\right| \mathbb{P}[\boldsymbol{m} = m] \\ &+ \sum_{\substack{m\in\mathbb{N}\\ |m-dn/k|\le C\sqrt{n}}} \mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] \left|\mathbb{P}[\boldsymbol{m} = m] - \mathbb{P}[\boldsymbol{m} = m \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]\right| + \varepsilon.\end{aligned}$$

By Theorem 2.0.4, we have

$$\sum_{m:|m-dn/k|\le C\sqrt{n}} \left|\mathbb{P}[G^{(1)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] - \mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n]\right| \mathbb{P}[\boldsymbol{m} = m] = o(1),$$

and by Lemma 7.3.2,

$$\sum_{m:|m-dn/k|\le C\sqrt{n}} \mathbb{P}[G^{(2)}_{n,m,\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] \left|\mathbb{P}[\boldsymbol{m} = m] - \mathbb{P}[\boldsymbol{m} = m \mid M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{S}_n]\right| = o(1).$$

Since $\varepsilon > 0$ was arbitrary, this shows that $\left|\mathbb{P}[G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] - \mathbb{P}[G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n]\right| = o(1)$. □

7.4 Proof of Lemmas 6.1.3 and 6.1.4

In this section, we use Theorem 7.3.5 and adaptations of arguments from Section 7.2 to prove Lemmas 6.1.3 and 6.1.4. In the following, a matching of S with T is a set of ordered pairs $\Gamma \subseteq S \times T$ such that $|\{(s,t) \in \Gamma, t \in T\}| \le 1$ for all $s \in S$ and $|\{(s,t) \in \Gamma, s \in S\}| \le 1$ for all $t \in T$. A matching $\Gamma \subseteq S \times T$ is perfect if $|\Gamma| = |S| = |T|$.

Lemma 7.4.1. *Let $\Gamma \subseteq (A \cup A') \times (B \cup B')$ be a uniformly chosen maximum size matching of $A \cup A'$ with $B \cup B'$. Let $\mathcal{S}'$ be the event that $e_{M''_{n,m,k,\rho,\varepsilon}}(a, \{v, v'\}) \le 1$ for all $a \in C_m$ and $(v, v') \in \Gamma$, and let $\mathcal{C}$ be the event that $|A \cup A'| = |B \cup B'| = n$. Let π be a uniform random permutation of V_{2n} chosen independently of $\mathbf{B}_{XY}$, and let $\tilde{\mathbf{B}}_{XY}$ be the matrix obtained from $\mathbf{B}_{XY}$ by permuting columns by π. Then $\mathbb{P}[\mathcal{S}' \cap \mathcal{C}] = \Omega(1)$ and $F(\tilde{\mathbf{B}}_{XY}) \sim M''_{n,m,k,\rho,\varepsilon} \mid \mathcal{S}' \cap \mathcal{C}$.*

Proof. Let $\Gamma_\pi \subseteq \pi(V_n) \times \pi(V_{2n} \setminus V_n)$ be the matching given by $\Gamma_\pi = \{(\pi(v_i), \pi(v_{i+n}) : i \in [n]\}$. Conditioned on $\mathcal{C}$, $A \cup A'$ is a uniformly chosen subset of V_{2n} of size n. Hence, conditioned on $\mathcal{C}$, the distributions of Γ and Γ_π are identical. Therefore it suffices to show that

$$\mathbb{P}[M''_{n,m,k,\rho,\varepsilon} = G \mid \mathcal{S}', \mathcal{C}, \Gamma = \gamma] = \mathbb{P}[F(\tilde{\mathbf{B}}_{XY}) = G \mid \Gamma_\pi = \gamma] \tag{7.9}$$

for any perfect matching $\gamma \subseteq T \times (V_{2n} \setminus T)$ with $T \subseteq V_n$, and any factor graph G with $e_G(a, \{v, v'\}) \le 1$ for all $(v, v') \in \gamma$, $a \in C_m$. For any such γ, G, we have

$$\begin{aligned}\mathbb{P}[M''_{n,m,k,\rho,\varepsilon} = G \mid \Gamma = \gamma] &= \prod_{a \in C_m} \binom{k}{e_G(a,v), v \in V_{2n}} \rho^{e_G(a,T)} (1-\rho)^{e_G(a, V_{2n} \setminus T)} n^{-k} \\ &= \prod_{a \in C_m} k! \rho^{e_G(a,T)} (1-\rho)^{e_G(a, V_{2n} \setminus T)} n^{-k} \end{aligned} \tag{7.10}$$

where $\binom{k}{e_G(a,v), v \in V_{2n}}$ is the number of ways to assign $b_{a,i}$, $i \in [k]$ so that $e_G(a,v) = |\{i \in [k] : b_{a,i} = v\}|$ for all $v \in V_{2n}$, and $\rho^{e_G(a,T)}(1-\rho)^{e_G(a, V_{2n} \setminus T)} n^{-k}$ is the probability of each

assignment. The last line holds by the assumption that $e_G(a, \{v, v'\}) \le 1$ for all $(v, v') \in \gamma$, $a \in C_m$. Moreover, by considering each $a \in C_m$ independently, we have

$$\mathbb{P}[\mathcal{S}' \mid \Gamma = \gamma] = \prod_{a \in C_m} \prod_{i=1}^{k} \left(1 - \frac{\rho(i-1)}{n} - \frac{(1-\rho)(i-1)}{n}\right) = \left(\frac{n!}{(n-k)!n^k}\right)^m = \Omega(1). \tag{7.11}$$

Since γ is assumed to be a perfect matching, we have $\{\Gamma = \gamma\} \cap \mathcal{C} = \{\Gamma = \gamma\}$ as an equality of events. Thus, by combining 7.10 and 7.11,

$$\begin{aligned}
\mathbb{P}[M''_{n,m,k,\rho,\varepsilon} = G \mid \mathcal{S}', \mathcal{C}, \Gamma = \gamma] &= \mathbb{P}[M''_{n,m,k,\rho,\varepsilon} = G \mid \mathcal{S}', \Gamma = \gamma] \\
&= \left(\frac{n!}{(n-k)!n^k}\right)^{-m} \prod_{a \in C_m} k! \rho^{e_G(a,T)} (1-\rho)^{e_G(a, V_{2n} \setminus T)} n^{-k} \\
&= \prod_{a \in C_m} \binom{n}{k}^{-1} \rho^{e_G(a,T)} (1-\rho)^{e_G(a, V_{2n} \setminus T)}. \qquad (7.12)
\end{aligned}$$

On the other hand, for any permutation σ such that $\Gamma_\sigma = \gamma$,

$$\mathbb{P}[F(\tilde{\mathbf{B}}_{XY}) = G \mid \pi = \sigma] = \prod_{a \in C_m} \binom{n}{k}^{-1} \rho^{e_G(a,T)} (1-\rho)^{e_G(a, V_{2n} \setminus T)} \tag{7.13}$$

where $\binom{n}{k}^{-1}$ is the probability of choosing the neighbourhood of a in $F(\mathbf{B})$ so that

$$\{i \in [n] : v_i \in N_{F(\mathbf{B})}(a)\} = \{i \in [n] : \sigma(v_i) \in N_G(a) \text{ or } \sigma(v_{i+n}) \in N_G(a)\},$$

and $\rho^{e_G(a,T)}(1-\rho)^{e_G(a, V_{2n} \setminus T)}$ is the probability of then assigning the coefficients $\mathbf{B}_{a,v}$ for $v \in N_{F(\mathbf{B})}(a)$ so that $\sigma(\{v_i : \mathbf{B}_{a,v_i} = \bar{X}\}) = N_G(a) \cap T$ and $\sigma(\{v_{i+n} : \mathbf{B}_{a,v_i} = \bar{Y}\}) = N_G(a) \setminus T$. Summing (7.13) over all σ with $\Gamma_\sigma = \gamma$ gives

$$\mathbb{P}[F(\tilde{\mathbf{B}}_{XY}) = G \mid \Gamma_\pi = \gamma] = \prod_{a \in C_m} \binom{n}{k}^{-1} \rho^{e_G(a,T)} (1-\rho)^{e_G(a, V_{2n} \setminus T)}.$$

Combined with 7.12, this verifies 7.9.

To show that $\mathbb{P}[\mathcal{S}'] = \Omega(1)$, first note that summing over γ in (7.11) gives $\mathbb{P}[\mathcal{S}' \mid \mathcal{C}] = \Omega(1)$. Since A' is chosen to minimize $||A \cup A'| - n|$, we have $|A \cup A'| = |B \cup B'| = n$ whenever $|A|, |B| \le n$. Since A is a uniformly chosen subset of S, where $|S| = (2-2\varepsilon)n$, this occurs with probability $1 - o(1)$. Therefore $\mathbb{P}[\mathcal{C}] = 1 - o(1)$, and hence $\mathbb{P}[\mathcal{S}' \cap \mathcal{C}] = \Omega(1)$. □

We now show that $M'_{n,m,k,\rho,\varepsilon}$ is contiguous with the random graph $M^{(2)}_{(2-2\varepsilon)n, \boldsymbol{m}, \boldsymbol{k}, \boldsymbol{d}}$, where $\boldsymbol{k} \sim \mathrm{Bin}(k, (1-\varepsilon))$ and $\boldsymbol{d}$ is a convex combination of Poisson random variables $\boldsymbol{d} \sim \mathrm{Po}(\rho d^*)/2 + \mathrm{Po}((1-\rho)d^*)/2$. Using the notation of section 2, we have $d = \mathbb{E}[\boldsymbol{d}] = d^*/2$, $k = \mathbb{E}[\boldsymbol{k}] = (1-\varepsilon)k$, and $\boldsymbol{m}' \sim \mathrm{Po}(d(2-2\varepsilon)n/k) = \mathrm{Po}(d^* n/k)$.

Let $\boldsymbol{\delta}''_v \sim Po(d^*\rho)$ and $\boldsymbol{\delta}'''_v \sim Po(d^*(1-\rho))$ for $v \in V_{2n}$ be such that $\{\boldsymbol{\delta}''_v\}_{v\in V_{2n}}$, $\{\boldsymbol{\delta}'''_v\}_{v\in V_{2n}}$, and $M'_{n,k,\rho,d^*,\varepsilon}$ are mutually independent. Let $\boldsymbol{\delta}'_v = \boldsymbol{\delta}''_v$ for $v \in A$ and $\boldsymbol{\delta}'_v = \boldsymbol{\delta}'''_v$ for $v \in B$. To make the distribution of the degree sequence of $M'_{n,k,\rho,d^*,\varepsilon}$ match that of $M^{(2)}_{(2-2\varepsilon)n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$, we condition on the event

$$\mathcal{D}'_n = \left\{ \sum_{a\in C_{\boldsymbol{m}}} \boldsymbol{\kappa}'_a = \sum_{v\in A} \boldsymbol{\delta}''_v + \sum_{v\in B} \boldsymbol{\delta}'''_v \right\}.$$

where $\boldsymbol{\kappa}_a$ is the degree of a in $M'_{n,m,k,\rho,\varepsilon}$.

Lemma 7.4.2. $M'_{n,k,\rho,d^*,\varepsilon} \triangleleft\triangleright M'_{n,k,\rho,d^*,\varepsilon} \mid \mathcal{D}'_n$

Proof. Let $\mathcal{T}_n(C)$ be the event

$$\mathcal{T}_n(C) = \left\{ ||A| - (1-\varepsilon)n| \le C\sqrt{n}, |\boldsymbol{m} - d^*n/k| \le C\sqrt{n}, \left| \sum_{a\in C_{\boldsymbol{m}}} \boldsymbol{\kappa}'_a - (1-\varepsilon)d^*n \right| \le C\sqrt{n} \right\}$$

Similar to Lemma 7.2.2, the lemma follows from the following two claims, which are analogous to Claims 7.2.3 and 7.2.4.

Claim 7.4.3.

$$\lim_{C\to\infty} \liminf_{n\to\infty} \mathbb{P}[\mathcal{T}_n(C)] = \lim_{C\to\infty} \liminf_{n\to\infty} \mathbb{P}[\mathcal{T}_n(C) \mid \mathcal{D}'_n] = 1.$$

Claim 7.4.4. *Suppose G_n is a sequence of factor graphs such that $|e(G_n) - (1-\varepsilon)d^*n| = O(\sqrt{n})$ and $|c(G_n) - d^*n/k| = O(\sqrt{n})$. Suppose a_n is such that $|a_n - (1-\varepsilon)n| = O(\sqrt{n})$. Then*

$$\mathbb{P}[M'_{n,k,\rho,d^*,\varepsilon} = G_n, |A| = a_n \mid \mathcal{D}'_n] = \Theta(1)\mathbb{P}[M'_{n,k,\rho,d^*,\varepsilon}, |A| = a_n].$$

□

Lemma 7.4.5. *For each $T \subseteq V_{2n}$ of size $(2-2\varepsilon)n$, let σ_T be a bijection from T to $V_{(2-2\varepsilon)n}$. Let $\tilde{M}'_{n,k,\rho,d^*,\varepsilon}$ be the factor graph obtained by mapping vertices in $A\cup B$ to $V_{(2-2\varepsilon)n}$ via $\sigma_{A\cup B}$. Then $\tilde{M}'_{n,k,\rho,d^*,\varepsilon} \mid \mathcal{D}'_n \sim M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$.*

Proof. Choose $(\hat{A}, \hat{A}', \hat{B}, \hat{B}', \{\boldsymbol{\kappa_a}\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{\delta_v}\}_{v\in\hat{A}\cup\hat{B}}, \hat{M}'_{n,k,\rho,d^*,\varepsilon})$ from the conditional distribution of $(A, A', B, B', \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{\delta}'_v\}_{v\in A\cup B}, M'_{n,k,\rho,d^*,\varepsilon})$ given $\mathcal{D}'_n$. We first show that the distributions of the degree sequences of $\hat{M}'_{n,k,\rho,d^*,\varepsilon}$ and $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ agree after mapping by $\sigma_{\hat{A}\cup\hat{B}}$.

Claim 7.4.6.

$$(\{\boldsymbol{\kappa}_a\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{\delta}_{\sigma_{\hat{A}\cup\hat{B}}(v)}\}_{v\in\hat{A}\cup\hat{B}}) \sim (\{\deg_{M^{(2)}_{n,\boldsymbol{kd}}}(a)\}_{a\in C_{\boldsymbol{m}}}, \{\deg_{M^{(2)}_{(2-2\varepsilon)n,\boldsymbol{k},\boldsymbol{d}}}(v)\}_{v\in V_{(2-2\varepsilon)n}}).$$

Proof. First note that $\{\boldsymbol{\delta}'_{\sigma_{A\cup B}(v)}\}_{v\in A\cup B}$ is a sequence of i.i.d. $\mathrm{Po}(\rho d^*)/2 + \mathrm{Po}((1-\rho)d^*)/2$ random variables that is independent of $\{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}}$. Moreover, $\{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}}$ is a sequence of i.i.d. $\mathrm{Bin}(k, 1-\varepsilon)$ random variables. Therefore $(\{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{\delta}_{\sigma_{A\cup B}(v)}\}_{v\in A\cup B})$ has the

same joint distribution as $(\{\boldsymbol{k}'_a\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{d}'_v\}_{v\in V_{(2-2\varepsilon)n}})$ in the notation of (2.1). Therefore $(\{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{\delta}'_{\sigma_{A\cup B}(v)}\}_{v\in A\cup B}) \mid \mathcal{D}'_n \sim (\{\boldsymbol{k}'_a\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{d}'\}_{v\in V_{2-2\varepsilon)n}}) \mid \mathcal{D}_n$, which proves the claim. □

Claim 7.4.7. *Conditioned on the event $|\hat{A}\cup\hat{A}'| = n$, we have*

$$(\{\boldsymbol{\kappa}_a\}_{a\in C_{\boldsymbol{m}}}, \{\boldsymbol{\delta}_v\}_{v\in\hat{A}\cup\hat{B}}) \sim (\{\boldsymbol{\kappa}_a\}_{a\in C_{\boldsymbol{m}}}, \{\deg_{\hat{M}'_{n,k,\rho,d^*,\varepsilon}}(v)\}_{v\in\hat{A}\cup\hat{B}}).$$

Proof. Let $\alpha,\beta\subseteq V_{2n}$ be disjoint sets such that $|\alpha\cup\beta| = (2-2\varepsilon)n$ and $|\alpha|,|\beta|\le n$. Let $\{k_a\}_{a\in C_j}\in\mathbb{N}^{C_j}$ for some j and let $\{d_v\}_{v\in\alpha\cup\beta}\in\mathbb{N}^{\alpha\cup\beta}$ such that $\sum_{v\in\alpha\cup\beta} d_v = \sum_{a\in C_j} k_a$. Let $M = \sum_{a\in C_j} k_a$. Then

$$\begin{aligned}
&\mathbb{P}[\{\boldsymbol{\delta}_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \mid \{\boldsymbol{\kappa}_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, \hat{A}=\alpha, \hat{B}=\beta]\\
&= \frac{\mathbb{P}[\{\boldsymbol{\delta}'_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta}, \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A=\alpha, B=\beta, \mathcal{D}'_n]}{\mathbb{P}[\{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A=\alpha, B=\beta, \mathcal{D}'_n]}\\
&= \frac{\mathbb{P}[\{\boldsymbol{\delta}''_v = d_v\}_{v\in\alpha}, \{\boldsymbol{\delta}'''_v = d_v\}_{v\in\beta}, \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A=\alpha, B=\beta]}{\mathbb{P}[\sum_{v\in\alpha}\boldsymbol{\delta}''_v + \sum_{v\in\beta}\boldsymbol{\delta}'''_v = M, \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A=\alpha, B=\beta]}\\
&= \frac{\mathbb{P}[\{\boldsymbol{\delta}''_v = d_v\}_{v\in\alpha}, \{\boldsymbol{\delta}'''_v = d_v\}_{v\in\beta}]}{\mathbb{P}[\sum_{v\in\alpha}\boldsymbol{\delta}''_v + \sum_{v\in\beta}\boldsymbol{\delta}'''_v = M]}
\end{aligned}$$

where the last equality follows from the independence of $\{\boldsymbol{\delta}''_v\}_{v\in V_{2n}}$, $\{\boldsymbol{\delta}'''_v\}_{v\in V_{2n}}$ and $\{\boldsymbol{\kappa}'\}_{a\in C_{\boldsymbol{m}}}, A, B$. Since the $\{\boldsymbol{\delta}''_v\}_{v\in V_{2n}}$, $\{\boldsymbol{\delta}'''_v\}_{v\in V_{2n}}$ are i.i.d. Poisson, we have

$$\mathbb{P}[\{\boldsymbol{\delta}''_v = d_v\}_{v\in\alpha}, \{\boldsymbol{\delta}'''_v = d_v\}_{v\in\beta}] = \prod_{v\in\alpha}\frac{e^{-\rho d^*}(\rho d^*)^{d_v}}{d_v!}\prod_{v\in\beta}\frac{e^{-(1-\rho)d^*}((1-\rho)d^*)^{d_v}}{d_v!}.$$

The sum $\sum_{v\in\alpha}\boldsymbol{\delta}''_v + \sum_{v\in\beta}\boldsymbol{\delta}'''_v$ is Poisson with mean $\rho|\alpha|d^* + (1-\rho)|\beta|d^*$, so we get

$$\begin{aligned}
\frac{\mathbb{P}[\{\boldsymbol{\delta}''_v = d_v\}_{v\in\alpha}, \{\boldsymbol{\delta}'''_v = d_v\}_{v\in\beta}]}{\mathbb{P}[\sum_{v\in\alpha}\boldsymbol{\delta}''_v + \sum_{v\in\beta}\boldsymbol{\delta}'''_v = M]} &= \frac{e^{\rho|\alpha|d^*+(1-\rho)|\beta|d^*}M!}{(\rho|\alpha|d^* + (1-\rho)|\beta|d^*)^M} \qquad (7.14)\\
&\quad\times\prod_{v\in\alpha}\frac{e^{-\rho d^*}(\rho d^*)^{d_v}}{d_v!}\prod_{v\in\beta}\frac{e^{-(1-\rho)d^*}((1-\rho)d^*)^{d_v}}{d_v!}\\
&= \frac{\rho^{\sum_{v\in\alpha}d_v}(1-\rho)^{\sum_{v\in\beta}d_v}M!}{\prod_{v\in\alpha\cup\beta}d_v!(\rho|\alpha| + (1-\rho)|\beta|)^M}. \qquad (7.15)
\end{aligned}$$

Now consider the distribution of $\{\deg_{\hat{M}'_{n,k,\rho,d^*,\varepsilon}}(v)\}_{v\in\hat{A}\cap\hat{B}}$. Let $\hat{\boldsymbol{\delta}}_v = \deg_{\hat{M}'_{n,k,\rho,d^*,\varepsilon}}(v)$ for $v\in\hat{A}\cup\hat{B}$, and let $\hat{\boldsymbol{\delta}}'_v = \deg_{M'_{n,k,\rho,d^*,\varepsilon}}(v)$ for $v\in A\cup B$. Using the independence of $\{\boldsymbol{\delta}''_v\}_{v\in V_{2n}}$, $\{\boldsymbol{\delta}'''_v\}_{v\in V_{2n}}$

and $\{\boldsymbol{\kappa}'\}_{a\in C_{\boldsymbol{m}}}, A, B$, we have

$$\begin{aligned}
&\mathbb{P}[\{\hat{\boldsymbol{\delta}}_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \mid \{\boldsymbol{\kappa}_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, \hat{A} = \alpha, \hat{B} = \beta] \\
&= \frac{\mathbb{P}[\{\hat{\boldsymbol{\delta}}'_v\}_{v\in A\cup B} = \{d_v\}_{v\in\alpha\cup\beta}, \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta, \mathcal{D}'_n]}{\mathbb{P}[\{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta, \mathcal{D}'_n]} \\
&= \mathbb{P}[\{\hat{\boldsymbol{\delta}}'_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \mid \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta].
\end{aligned}$$

Let $\mathcal{B} = \{(a,i) \in C_j \times [k] : b_{a,i} \in A \cup B\}$, and let $\mathscr{B}$ be the set of subsets $D \subseteq C_j \times [k]$ such that $|\{i \in [k] : (a,i) \in D\}| = k_a$ for all $a \in C_j$. Then we can write

$$\begin{aligned}
&\mathbb{P}[\{\hat{\boldsymbol{\delta}}'_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \mid \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta] \\
&= \sum_{D\in\mathscr{B}} \big(\mathbb{P}[\{\hat{\boldsymbol{\delta}}'_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \mid \mathcal{B} = D, A = \alpha, B = \beta] \\
&\qquad\qquad \times \mathbb{P}[\mathcal{B} = D \mid \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta]\big) \qquad (7.16)
\end{aligned}$$

To calculate this probability, first note that the $\{b_{a,i}\}_{a\in C_j, i\in[k]}$ remain independent conditioned on $\mathcal{B}$, A, and B. Then, for any $(a,i) \in D$ and $v \in \alpha$, we have

$$\begin{aligned}
\mathbb{P}[b_{a,i} = v \mid \mathcal{B} = D, A = \alpha, B = \beta] &= \mathbb{P}[b_{a,i} = v \mid b_{a,i} \in \alpha\cup\beta, A = \alpha, B = \beta] \\
&= \frac{\mathbb{P}[b_{a,i} = v \mid A = \alpha, B = \beta]}{\mathbb{P}[b_{a,i} \in \alpha\cup\beta \mid A = \alpha, B = \beta]}
\end{aligned}$$

Since $|\alpha|, |\beta| \le n$, the event $A = \alpha, B = \beta$ implies $|A \cup A'| = |B \cup B'| = n$. Hence,

$$\frac{\mathbb{P}[b_{a,i} = v \mid A = \alpha, B = \beta]}{\mathbb{P}[b_{a,i} \in \alpha\cup\beta \mid A = \alpha, B = \beta]} = \frac{\rho/n}{\rho|\alpha|/n + (1-\rho)|\beta|/n}.$$

Similarly, we have $\mathbb{P}[b_{a,i} = v \mid \mathcal{B} = D, A = \alpha, B = \beta] = (1-\rho)/(\rho|\alpha| + (1-\rho)|\beta|)$ for $v \in \beta$. Therefore, we get

$$\begin{aligned}
\mathbb{P}[\{\hat{\boldsymbol{\delta}}'_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \mid \mathcal{B} = D, A = \alpha, B = \beta] &= \binom{M}{d_v, v\in\alpha\cup\beta} \prod_{v\in\alpha} \left(\frac{\rho}{\rho|A| + (1-\rho)|\beta|}\right)^{d_v} \\
&\quad \times \prod_{v\in\beta} \left(\frac{1-\rho}{\rho|A| + (1-\rho)|\beta|}\right)^{d_v} \\
&= \frac{\rho^{\sum_{v\in\alpha} d_v} (1-\rho)^{\sum_{v\in\beta} d_v} M!}{\prod_{v\in\alpha\cup\beta} d_v! (\rho|\alpha| + (1-\rho)|\beta|)^M}
\end{aligned}$$

Substituting into (7.16) shows that

$$\begin{aligned}
&\mathbb{P}[\{\hat{\boldsymbol{\delta}}'_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \mid \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta] \\
&\qquad = \frac{\rho^{\sum_{v\in\alpha} d_v} (1-\rho)^{\sum_{v\in\beta} d_v} M!}{\prod_{v\in\alpha\cup\beta} d_v! (\rho|\alpha| + (1-\rho)|\beta|)^M} \qquad (7.17)
\end{aligned}$$

which agrees with (7.15). This proves the claim. □

Since $\mathbb{P}[|\hat{A} \cup \hat{A}'| = n] = 1 - o(1)$, Claims 7.4.6 and 7.4.7 show that the degree sequences of $\tilde{M}'_{n,k,\rho,d^*,\varepsilon} \mid \mathcal{D}'_n$ and $M^{(2)}_{a,\boldsymbol{k},\boldsymbol{d}}$ converge in total variation distance as $n \to \infty$. Therefore the following claim completes the proof of the lemma.

Claim 7.4.8. *Let $\{d_v\}_{v\in(2-2\varepsilon)n} \in \mathbb{N}^{V_{(2-2\varepsilon)n}}$ and let $\{k_a\}_{a\in C_j} \in \mathbb{N}^{C_j}$. Then, conditioned on having degree sequences $\{d_v\}_{v\in(2-2\varepsilon)n}$, $\{k_a\}_{a\in C_j}$, the distributions of $\tilde{M}'_{n,k,\rho,d^*,\varepsilon} \mid \mathcal{C}$ and $M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}}$ are identical.*

Proof.

$$\begin{aligned}
&\mathbb{P}[M'_{n,k,\rho,d^*,\varepsilon} = G \mid \{\boldsymbol{\kappa}_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta] \\
&= \prod_{a\in C_j} \binom{k_a}{e_G(a,v), v \in V_{\alpha\cup\beta}} \prod_{v\in\alpha} \left(\frac{\rho}{\rho|\alpha| + (1-\rho)|\beta|}\right)^{e_G(a,v)} \prod_{v\in\beta} \left(\frac{1-\rho}{\rho|\alpha| + (1-\rho)|\beta|}\right)^{e_G(a,v)} \\
&= \frac{\prod_{a\in C_j} k_a! \rho^{\sum_{v\in\alpha} \deg_G(v)} (1-\rho)^{\sum_{v\in\beta} \deg_G(v)}}{\prod_{a\in C_v, v\in\alpha\cup\beta} e(a,v)! (\rho|\alpha| + (1-\rho)|\beta|)^M}
\end{aligned}$$

Using (7.17), we get

$$\begin{aligned}
&\mathbb{P}[M'_{n,k,\rho,d^*,\varepsilon} = G \mid \{\hat{\boldsymbol{\delta}}'_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \{\boldsymbol{\kappa}_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta] \\
&= \frac{\mathbb{P}[M'_{n,k,\rho,d^*,\varepsilon} = G \mid \{\boldsymbol{\kappa}_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta]}{\mathbb{P}[\{\hat{\boldsymbol{\delta}}'_v\}_{v\in\hat{A}\cup\hat{B}} = \{d_v\}_{v\in\alpha\cup\beta} \mid \{\boldsymbol{\kappa}'_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}, A = \alpha, B = \beta]} \\
&= \frac{\prod_{v\in\alpha\cup\beta} d_v! \prod_{a\in C_j} k_a!}{M! \prod_{a\in C_j, v\in\alpha\cup\beta} e_G(a,v)!}
\end{aligned}$$

Since $M^{(2)}_{n,\boldsymbol{k},\boldsymbol{d}}$ is drawn from the paring model, the proof of Claim 7.2.8 shows that

$$\mathbb{P}[M^{(2)}_{n,\boldsymbol{k},\boldsymbol{d}} = G \mid \{\boldsymbol{d}_v\}_{v\in V_{(2-2\varepsilon)n}} = \{d_v\}_{v\in V_{(2-2\varepsilon)n}} \{\boldsymbol{k}_a\}_{a\in C_{\boldsymbol{m}}} = \{k_a\}_{a\in C_j}] = \frac{\prod_{v\in\alpha\cup\beta} d_v! \prod_{a\in C_j} k_a!}{M! \prod_{a\in C_j, v\in\alpha\cup\beta} e_G(a,v)!}$$

which proves the claim. □

□

Proof of Lemma 6.1.3. Let $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ be a sequence of sets of simple factor graphs. By Lemma 7.4.1, we have

$$\mathbb{P}[F(\tilde{B}_{XY}) \in \mathcal{E}_n] = \mathbb{P}[M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n \mid \mathcal{S}' \cap \mathcal{C}]$$

and

$$\mathbb{P}[\{M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n\} \cap \mathcal{C}] \leq \mathbb{P}[M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n \mid \mathcal{S}' \cap \mathcal{C}] \leq \frac{\mathbb{P}[M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n]}{\mathbb{P}[\mathcal{S}' \cap \mathcal{C}]}$$

where $\mathbb{P}[\mathcal{S}' \cap \mathcal{C}] = \Omega(1)$ and $\mathbb{P}[\mathcal{C}] = 1 - o(1)$. Therefore $M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n$ a.a.s. if and only if $F(\tilde{B}_{XY}) \in \mathcal{E}_n$ a.a.s.

By the definition of $F(\mathbf{C}'')$, we have $\mathbb{P}[F(\mathbf{C}'') \in \mathcal{E}_n] = \mathbb{P}[M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n \mid \mathcal{S}]$, where $\mathcal{S}$ is the event that $M_{n,m,k,\rho,d^*,\varepsilon}$ has no parallel edges. By (6.2), $\mathbb{P}[\mathcal{S}] = \Omega(1)$, so we get

$$\mathbb{P}[M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n] \leq \mathbb{P}[F(\mathbf{C}'') \in \mathcal{E}_n] \leq O(1)\mathbb{P}[M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n].$$

Therefore $F(\mathbf{C}'') \in \mathcal{E}_n$ a.a.s. if and only if $M''_{n,m,k,\rho,\varepsilon} \in \mathcal{E}_n$ a.a.s. This completes the proof. □

Proof of Lemma 6.1.4. Let $\{\mathcal{E}_n\}$ be a sequence of sets of simple factor graphs, and suppose $\mathbb{P}[G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] = o(1)$. By Theorem 7.3.5,

$$\mathbb{P}[M^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] \leq \mathbb{P}[G^{(2)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] \leq \mathbb{P}[G^{(1)}_{n,\boldsymbol{m},\boldsymbol{k},\boldsymbol{d}} \in \mathcal{E}_n] + o(1) = o(1).$$

Then by Lemma 7.4.5, $\mathbb{P}[\tilde{M}'_{n,k,\rho,d^*,\varepsilon} \in \mathcal{E}_n \mid \mathcal{D}'_n] = o(1)$, and by Lemma 7.4.2, $\mathbb{P}[\tilde{M}'_{n,k,\rho,d^*,\varepsilon} \in \mathcal{E}_n] = o(1)$. Similar to the proof of Lemma 6.1.3, we have $\mathbb{P}[F(\tilde{\mathbf{C}}') \in \mathcal{E}_n] = \mathbb{P}[\tilde{M}'_{n,k,\rho,d^*,\varepsilon} \in \mathcal{E}_n \mid \mathcal{S}]$ with $\mathbb{P}[\mathcal{S}] = \Omega(1)$, so we get $\mathbb{P}[F(\tilde{\mathbf{C}}') \in \mathcal{E}_n] = o(1)$. This completes the proof. □

References

[1] Dimitris Achlioptas, Arthur Chtcherba, Gabriel Istrate, and Cristopher Moore. The phase transition in 1-in-k SAT and nae 3-SAT. *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*, 05 2001.

[2] Dimitris Achlioptas, Amin Coja-Oghlan, and Federico Ricci-Tersenghi. On the solution-space geometry of random constraint satisfaction problems. *Random Structures & Algorithms*, 38(3):251–268, 2011.

[3] Dimitris Achlioptas and Michael Molloy. The solution space geometry of random linear equations. *Random Structures & Algorithms*, 46(2):197–231, 2015.

[4] Dimitris Achlioptas and Cristopher Moore. The asymptotic order of the random k -SAT threshold. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, FOCS '02, page 779–788, USA, 2002. IEEE Computer Society.

[5] Dimitris Achlioptas and Cristopher Moore. Random k-SAT: Two moments suffice to cross a sharp threshold. *SIAM Journal on Computing*, 36(3):740–762, 2006.

[6] Dimitris Achlioptas and Yuval Peres. The threshold for random k-SAT is 2k (ln 2 - o(k)). In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '03, page 223–231, New York, NY, USA, 2003. Association for Computing Machinery.

[7] Peter Ayre, Amin Coja-Oghlan, Pu Gao, and Noëla Müller. The satisfiability threshold for random linear equations. *Combinatorica*, 40(2):179–235, 2020.

[8] Victor Bapst, Amin Coja-Oghlan, Samuel Hetterich, Felicia Raßmann, and Dan Vilenchik. The condensation phase transition in random graph coloring. *Communications in Mathematical Physics*, 341(2):543–606, Jan 2016.

[9] Gilberto Bini and Flaminio Flamini. *Finite commutative rings and their applications*. Springer, 2002.

[10] Béla Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2001.

[11] Béla Bollobás, Jeong Han Kim, and Jacques Verstraëte. Regular subgraphs of random graphs. *Random Structures & Algorithms*, 29(1):1–13, 2006.

[12] Peter Cheeseman, Bob Kanefsky, and William M. Taylor. Where the really hard problems are. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence - Volume 1*, IJCAI'91, page 331–337, San Francisco, CA, USA, 1991. Morgan Kaufmann Publishers Inc.

[13] Amin Coja-Oghlan, Alperen A Ergür, Pu Gao, Samuel Hetterich, and Maurice Rolvien. The rank of sparse random matrices. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 579–591. SIAM, 2020.

[14] Amin Coja-Oghlan, Pu Gao, Max Hahn-Klimroth, Joon Lee, Noela Müller, and Maurice Rolvien. The full rank condition for sparse random matrices. *Combinatorics, probability and computing*, 33(5):643–707, 2024.

[15] Amin Coja-Oghlan, Tobias Kapetanopoulos, and Noela Müller. The replica symmetric phase of random constraint satisfaction problems. *Combinatorics, Probability and Computing*, 29(3):346–422, 2020.

[16] Amin Coja-Oghlan and Konstantinos Panagiotou. Going after the k-SAT threshold. In *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '13, page 705–714, New York, NY, USA, 2013. Association for Computing Machinery.

[17] Amin Coja-Oghlan and Konstantinos Panagiotou. The asymptotic k-SAT threshold. *Advances in Mathematics*, 288:985–1068, 2016.

[18] Harold Connamacher and Michael Molloy. The satisfiability threshold for a seemingly intractable random constraint satisfaction problem. *SIAM Journal on Discrete Mathematics*, 26(2):768–800, 2012.

[19] Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, page 151–158, New York, NY, USA, 1971. Association for Computing Machinery.

[20] Kevin Costello, Terence Tao, and Van Vu. Random symmetric matrices are almost surely nonsingular. *Duke Mathematical Journal*, 135, 11 2006.

[21] Kevin Costello and Van Vu. On the rank of random sparse matrices. *Duke Mathematical Journal*, 19(3):321–342, May 2010.

[22] James M. Crawford and Larry D. Auton. Experimental results on the crossover point in random 3-SAT. *Artificial Intelligence*, 81(1):31–57, 1996. Frontiers in Problem Solving: Phase Transitions and Complexity.

[23] Burgess Davis and David McDonald. An elementary proof of the local central limit theorem. *Journal of Theoretical Probability*, 8(3):693–701, July 1995.

[24] Martin Dietzfelbinger, Andreas Goerdt, Michael Mitzenmacher, Andrea Montanari, Rasmus Pagh, and Michael Rink. Tight thresholds for cuckoo hashing via XORSAT, 2010.

[25] Jian Ding, Allan Sly, and Nike Sun. Satisfiability threshold for random regular NAE-SAT. *Communications in Mathematical Physics*, 341(2):435–489, Jan 2016.

[26] Jian Ding, Allan Sly, and Nike Sun. Proof of the satisfiability conjecture for large k. *Annals of Mathematics*, 196(1):1 – 388, 2022.

[27] Olivier Dubois and Jacques Mandler. The 3-XORSAT threshold. *Comptes Rendus Mathematique*, 335(11):963–966, 2002.

[28] Ehud Friedgut and Jean Bourgain. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American mathematical Society*, 12(4):1017–1054, 1999.

[29] Pu Gao and Theodore Morrison. The satisfiability threshold and solution space of random uniquely extendable constraint satisfaction problems, 2026.

[30] Pu Gao and Theodore Morrison. The satisfiability threshold of random linear equations over finite commutative rings. Submitted to the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2026.

[31] Andreas Goerdt and Lutz Falke. Satisfiability thresholds beyond k-XORSAT. In *International Computer Science Symposium in Russia*, pages 148–159. Springer, 2012.

[32] Lefteris M. Kirousis, Evangelos Kranakis, Danny Krizanc, and Yannis C. Stamatiou. Approximating the unsatisfiability threshold of random formulas. *Random Structures & Algorithms*, 12(3):253–269, 1998.

[33] F.R. Kschischang, B.J. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*, 47(2):498–519, 2001.

[34] Leonid A. Levin. Universal sequential search problems. *Problems of Information Transmission*, 9(3), 1973.

[35] Nikita Lvov. Universality results for random matrices over finite local rings, 2026.

[36] Kenneth Maples. Cokernels of random matrices satisfy the Cohen-Lenstra heuristics, 2013.

[37] M Mézard, T Mora, and R Zecchina. Clustering of solutions in the random satisfiability problem. *Phys Rev Lett*, 94(19):197205, May 2005.

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

[39] David Mitchell, Bart Selman, and Hector Levesque. Hard and easy distributions of SAT problems. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, AAAI'92, page 459–465. AAAI Press, 1992.

[40] Michael Molloy. The freezing threshold for k-colourings of a random graph. In *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing*, STOC '12, page 921–930, New York, NY, USA, 2012. Association for Computing Machinery.

[41] Rémi Monasson and Riccardo Zecchina. Entropy of the k-satisfiability problem. *Phys. Rev. Lett.*, 76:3881–3885, May 1996.

[42] Andrea Montanari, Ricardo Restrepo, and Prasad Tetali. Reconstruction and clustering in random constraint satisfaction problems. *SIAM Journal on Discrete Mathematics*, 25(2):771–808, 2011.

[43] M. Mézard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297(5582):812–815, 2002.

[44] Hoi Nguyen and Melanie Wood. Random integral matrices: universality of surjectivity and the cokernel. *Inventiones mathematicae*, 228:1–76, 04 2022.

[45] Boris Pittel and Gregory Sorkin. The satisfiability threshold for k-XORSAT. *Combinatorics, Probability and Computing*, 25:236–268, 2016.

[46] Federico Ricci-Tersenghi and Guilhem Semerjian. On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms. *Journal of Statistical Mechanics: Theory and Experiment*, 2009:P09001, 09 2009.

[47] Bart Selman, Hector Levesque, and David Mitchell. A new method for solving hard satisfiability problems. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, AAAI'92, page 440–446. AAAI Press, 1992.

[48] Konstantin Tikhomirov. Singularity of random bernoulli matrices. *Annals of Mathematics*, 191:593, 03 2020.