PARTITION FUNCTION ESTIMATION AND PHASE TRANSITIONS ON RANDOM SATISFIABILITY PROBLEMS

Dissertation

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

der Technischen Universität Dortmund
an der Fakultät für Informatik

von

Arnab Chatterjee

aus *Kharagpur,West Bengal,India*

Dortmund

2025

Dekan:
Prof. Dr. Jens Teubner

Gutachter:
Prof. Dr. Amin Coja-Oghlan
TU Dortmund

Prof. Dr. Dimitris Achlioptas
University of Athens

Datum der mündlichen Prüfung: 25.11.2025

DEDICATION

To my parents
and
all young minds of West Bengal.

Acknowledgements

"So long, and thanks for all the fish."

Looking back at my time as a master student at IIIT Delhi, at the end of my second semester, when many of my friends chose their thesis topics in the booming areas of machine learning and artificial intelligence, I was fortunate to have *Prof. Subhabrata Samajder* as my master thesis advisor who introduced me to the world of "*Random Graphs*". Many of the random ideas that struck me during my thesis took shape in various forms over the next several years and culminated in this thesis. I would say *Prof. Samajder* is one of those responsible person for turning my attention towards theoretical computer science. This thesis is my humble token of appreciation for his effort and sincerity.

Nevertheless, one year and a half working at the industry after completing my masters, I made one of the most important decisions of my life – pursuing a PhD. in the field of random graphs. So, the cruise again resume its journey with a different commander now on the deck. My PhD. supervisor *Prof. Amin Coja-Oghlan* patiently let me explore the depths but gave the screw the crucial half-turn just when needed to prevent me from floating away. I can't thank him enough for his constant encouragement and his confidence in me. Each and every time whenever I came up with a new idea he has been patient with me to stretch my mind and raise me up to think like a mathematician. Throughout the PhD. his guidance and support played a pivotal role in my decision to dedicate myself towards random graphs and probabilistic combinatorics more specifically on random satisfiability problems.

I would also like to extend my sincere appreciation to my dissertation committee members– *Dimitris Achlioptas*, *Jean Christoph Jung* and *Kevin Buchin* for taking the time to read my thesis, participate in my defense, and provide thoughtful feedback..

I am also beholden to my coauthors from world's prestigious institutes — Prof. Catherine Greenhill (UNSW), Prof. Mihyung Kang (TU Graz), Prof. Noela Múller (TU Eindhoven) and Prof. Gregory B. Sorkin (LSE). Collaborating with them broadened my understanding of the field and advanced my learning trajectory. Besides them, I am also thankful to work closely with my colleagues at TU Dortmund. *Kostas, Lena, Maurice, Olga, Pavel* and *Ulrike* created a space that was both supportive and engaging, with conversations covering from technical details to everyday's casual banter. Special thanks go to *Maurice* who helped me not only in finding accommodation, but also in some official works as an interpreter during my earlier days in Germany.

I also thank *Haodong* and *Joon* with whom I spent a very good time at Leiden and our conversation often went late into the evenings reminding me that the mathematics can be serious but also fun.

During my PhD. I also have the valuable opportunity to spend a research visit at *University of California, Irvine* (UCI). I owe my heartiest thanks to *Prof. Asaf Ferber* for hosting and making me feel welcome to his research group. Beside him, I also thank to his group members – *Marcelo, Mason* and *Xiaonan* with whom I shared my california stays, so not just productive discussion but also many memorable days, from working on existing problems to explore new ideas.

The whole PhD, journey is not complete without the people who provide strength outside the academic world. My parents *Soumen Chatterjee* and *Mamata Chatterjee* are above and beyond all thanks that I can ever gather. I am dedicating this thesis to my lovely parents as well as to all the young minds of West Bengal. It is not possible to adequately express in words the encouragement and support they have given me throughout my graduation timeframe. I hope I somewhat succeeded in meeting their expectations. I also owe heartiest thanks to my *bhai* (cousin) and *masi* (aunty), whose encouragement and motivation gave me the strength to keep going even when things seem uncertain.

My deepest appreciation goes to my fiancée, *Susmita* who has been my constant companion and whose belief, love, care on me has carried me through the difficult times.

This thesis, which contains whatever my research output that I could possibly 'write' in words, is a joint fruit of labor, persistence and confidence of a lot of people, spread all around the globe. I took this opportunity to thank all those who helped me turn a possibility into a reality. I surely missed more names than I remembered to mention above. But mentioned or unmentioned, my gratitude transgresses the words I used to express my heartiest thanks.

Arnab,
California, September 2025.

Abstract

This thesis emphasizes on the estimation of partition functions and analyze phase transitions in random satisfiability problems with focuses on random 2-SAT, random k-XORSAT and random k-SAT models. *Partition functions* capture the exponential growth of solution spaces and establish a bridge among combinatorics, probability, and statistical physics. Studying their asymptotics and fluctuations helps us to understand the mechanisms behind sharp phase transitions and the solution space geometry in random constraint satisfaction problems.

Our first contribution establishes a central limit theorem for the number of solutions (also called 'partition function' in physics jargon) of random 2-SAT – first CLT of this type for any random CSPs. Thereby it provides a precise probabilistic characterization of fluctuations on the logarithm of the number of satisfying assignments of order $\sqrt{n}$ with n the number of variables. In addition to this we effectively evaluated the formula for variance on the number of random 2-SAT solutions. The proof techniques relies on the Martingale central limit theorem along with the Gibbs uniqueness property and the local convergence to the Galton-Watson tree combined with a coupling argument called *'Aizenmann-Sims-Starr scheme'*.

The second part of the thesis investigates the performance of a statistical physics inspired message passing algorithm called '*Belief Propagation Guided Decimation*' on the random k-XORSAT problem. Specifically, we derive an explicit threshold upto which the algorithm succeeds with a strictly positive probability between 0 and 1. Additionally, we study a thought experiment called *'Decimation process'* for which we determine different phase transitions such as (non)-reconstruction and condensation phase transition and their connection to BPGD (in which regimes these two processes diverges or converge).

Finally, for random k-SAT, we revisited the Gibbs uniqueness threshold, improving the lower bound over the previous work by Montanari and Shah [83]. More specifically, we count the number of actual satisfying assignments of random k-SAT which is given by the physics inspired *'replica symmetry solution'* upto the Gibbs uniqueness threshold. Mathematically, we find an explicit expression on the logarithm of the number of solutions of random k-SAT in terms of the *Bethe free entropy* which is a function defined for a probability measure in the unit intterval. Moreover, our lower bound in contrast to Montanari-Shah bound is significant particularly for small k.

In a nutshell, this thesis advance the rigorous understanding of random satisfiability problems by combining the algorithmic analysis, probabilistic combinatorics and statistical physics equipment. In light of both the structural properties of random formulas and the effectiveness of different message passing algorithms along with the universal principles governing fluctuations, correlation decay and mathematical foundation for the phenomena predicted by spin glass theory, point toward new directions for the future research on random satisfiability problems.

Contents

List of Figures

1
Introduction

"The theory of probability as mathematical discipline can and should be developed from axioms in exactly the same way as Geometry and Algebra."

– Andrey Kolmogorov

In theoretical computer science, profound insights often arise at the intersection of the discrete and probabilistic paradigms, formally referred to as *probabilistic combinatorics*. In this setting, *random graphs* serve as a basic framework where edges are assigned at random according to a given probability distribution [48]. Many researchers investigate the circumstances under which important global properties emerge namely connectivity, the emergence of a giant component and the threshold in chromatic number [3, 5, 48, 63, 94]. These phase transitions or discontinuities, not only reflect phenomena in statistical physics but also reveal the average-case complexity measures of algorithms [1]. Generalizing this perspective from graphs to higher dimensions naturally allows for a deep investigation of random constraint satisfaction problems (*CSPs*) in which constraints are a generalization of edges and involve assignments rather than colorings. As a consequence, random CSPs form a unifying locus: they capture the combinatorial complexity of random graphs which require the precise algorithmic attention typical in computer science and use probabilistic techniques honed in the mathematics domain [9, 10]. This thesis capitalizes on this ternary relation in studying phase transitions in random CSPs, characterizing the number of satisfying assignments, the constraint density boundaries beyond which satisfiability disappears, analyze the performance of few physics inspired algorithms on random CSPs and detailing the structural mechanisms underlying the phase transition [8].

The Constraint Satisfaction Problem (CSP) is defined as the set of n discrete-valued variables taking values from a finite domain and finds a satisfying assignment subject to m-constraints. Each of the constraints enforces some requirements on a subset of the variables [99]. A solution of the CSP is an assignment of the variables that satisfies simultaneously all the constraints. There are numerous studies were tackled in the field of computer science as well as in combinatorics and statistical mechanics [11, 61, 66, 68]. Its inter-disciplinary research interest comes from a broad spectrum of applications ranging from coding theory and communication engineering to computer architecture, operational research and artificial intelligence [17, 45, 46, 54]. Famous examples of CSPs are the k-satisfiability (k-SAT) problem and the graph q-coloring one(q-COL). In k-SAT the variables are boolean and each constraint is the disjunction (OR) of k-literals (either a variable or its negation). In the second one the variables are placed on the vertices of a graph where they can take q-possible colors and each edge of the graph enforce the constraint that the two end vertices of an edge take different colors.

CSPs can be analyzed from several different perspectives. One fundamental approach comes from computational complexity theory [51, 84, 91] that classifies CSPs based on their worst-case difficulty. Particularly, it determines whether there exists an efficient polynomial time algorithm in the number of variables n and m clauses to determine the existence of a solution for every possible instances. It may come in several forms — beside the search variant which tries to find a satisfying assignment, the decision variant handles the question of whether there exists a solution or not [66]. Once a solution does exist one can might interest in how many such solutions exist [100]. While the general CSPs can be defined over deterministic structures, a complementary insightful perspective turns up when randomness comes into picture above the classical CSPs, in which behaviors are hard to observe in a deterministically generated instance [73], defined as *random* CSPs — the models in which the structure, assignments of variables and constraints, the matrix corresponding to the linear equations in $\mathbb{F}_q$ are chosen according to some probability distribution.

Over the years, a significant developments have been made to illuminate the various behavior of random CSPs. It share a formal mathematical analogy with the models studied in statistical mechanics of disordered systems, particularly in the mean-field spin glasses [16, 87, 88, 90] where the interactions induced by constraints are of vexing nature and due to the randomness in their construction, they don't possess any underlying finite dimensional structure. For instance, take the example of q-coloring problem on random graphs, the variable can be treated as Potts spins, and our goal is to find the ground state in the anti-ferromagnetic Potts model [13, 39]. A straight forward observation is when all the connection edges in the ground state are bi-colored, the problems admits a satisfying assignment as all the constraints are satisfied simultaneously. In the case of satisfiability problem, for instance in the case of random k-SAT (with $k = 3$ for simpler version) each variable inside a clause one can assume as a spin in the Ising model and the clauses can be thought as the interactions between the spins with satisfying clauses refer to a certain configuration of the spins [37, 40, 77, 78, 89]. With the connection to statistical mechanics one could define an energy function that penalizes the unsatisfied clauses. Like previous example here also our goal is to find the ground state of this system which would correspond to a satisfying assignment. Analogously, one can find the low-energy configuration in an Ising model. A particular interesting research topic is to analyze the regime where both the number of variables(n) and clauses (m) tends to infinity (∞) at a fixed ratio $\alpha = m/n$, called as the *constraint density ratio*. The random CSPs exhibit the threshold phenomena in this regime, the probability of some of the properties falling abruptly from 1 to 0 [74, 80] as a function of α which treated as a control parameter. Any of these phase transitions occurs at the satisfiability threshold denoted as α_{SAT} which depends on the parameter k, d, q of the problem (where d refers to average variable degree of a random satisfiability problem and q refers to the number of colorings that can be assigned to a variable in the case of q-coloring problem on random (hyper)-graph). For $\alpha < \alpha_{\text{SAT}}$, the system admits a satisfying assignment that satisfy all the constraint simultaneously, while a random instance is typically unsatisfiable for $\alpha > \alpha_{\text{SAT}}$[1].

To establish phase transitions in different random CSPs, statistical physicists have contributed significantly by analyzing non-rigorous methods and explained the combinatorial structure of CSPs by providing in details the solution space geometry of the problem and its connection with the phase transitions [59, 66]. Thanks to the similarity between random CSPs and spin glasses, the application of these methods first develop in the context of statistical mechanics of disordered systems namely the *replica symmetry* and *cavity method* [15, 59, 65, 72, 79] in the statistical physics jargon. These non-rigorous techniques have provided the predictions of α_{SAT} in many models as well uncovered many other phase transition corresponds to the structure of the set of solutions in the satisfiability phase. Additionally, the *cavity method* has inspired new algorithms, particularly the message passing algorithms as well spectral algorithms and leads to exciting predictions to the information-theoretic and computational nature of different inference problems. These algorithms exploit the detailed picture of the solution space, out of which some of the predictions have been confirmed later as well rigorously [1, 8, 41, 76]. So, on one side statistical physicists are happy with their heuristic predictions or evidence to a given problem, whereas on the other side mathematician/theoretical computer scientists try to verify their heuristic predictions by means of mathematical proofs. For example, in [32], the authors mathematically establish a formula for the mutual information in statistical inference problems induced by random graphs inspired by cavity method, a non-rigorous physics approach.

In this PhD. also we verify and confirm such a heuristic prediction of two physicists Ricci-Tersenghi and Semerjian [96] using a physics inspired message passing algorithm called as *Belief Propagation Guided Decimation*. In particular, we [22] derive an explicit threshold upto which the algorithm succeeds with a strictly positive probability $\Omega(1)$ and beyond which it fails to find a satisfying assignment with high probability [22]. Alongside, we also analyze a thought experiment called the *decimation process* for which we identify a *(non)-reconstruction* and a *condensation* phase transition. There is one more interesting phase transition occurs in the satisfiable phase, name as *clustering* (dynamic) phase transition occurs at a critical density denoted as α_{clus}. From the name it is clear that this transition emphasizes the drastic change of the shape of the set of solutions, which can be treated as a subset of the whole configuration space. Below this threshold (α_{clus}), the set of solutions is rather

[1]Although Friedgut [47] provides the transition from SAT to UNSAT (Theorem 4.1.2, Chapter 4), the existence of satisfiability thresholds (α_{SAT}) still is an open problem in many interesting problems like- random 3-SAT, random k-NAE-SAT (for small k values – say random 3-NAE-SAT), random 3-graph coloring etc. Many other general CSP thresholds besides k-SAT and coloring predicted by physics heuristics, but not proved rigorously.

well-connected that is any solution can be reached from other by a reshuffling of a confined number of variables, whereas above α_{clus} the set of solution chops into a large number of distinct clusters which corresponds to a pure decomposition of the uniform measure over the set of solutions [59]. Internally, the clusters are well-connected but well-separated from the other. This phase transition also marks the emergence of a certain long range correlations among variables which enables the solvability of an information-theoretic problem called tree reconstruction [85]. But when we are talking about the static version of the model, then the properties are not affected by this clustering transition, rather only sensitive to a further transition called *condensation* phase transition denoted as α_{cond} which affects the number of dominant clusters [59]. In random k-XORSAT both (non)-reconstruction and the clustering thresholds coincide (the phase diagram can be found in Chapter 4). Because of the structure of the k-XORSAT (which can be translated into linear system over $\mathbb{F}_2$), the solution space geometry is completely determined by the linear 2-core which marks both (i). the onset of frozen variables – shattering into clusters and (ii). the onset of long-range correlations on the GW tree model which implies reconstruction possible. Further the clustering threshold α_{clus} can be defined as the appearance of a non-trivial solution of the one step of *Replica Symmetry Breaking* (1RSB) equation with Parisi breaking parameter $\mathscr{X} = 1$, in the context of cavity method [69]. For more details in particular about the reconstruction and systematic connection one can refer to 67. Experimentally, for some random CSPs it is verified that the clustering threshold α_{clus} happens at a much smaller constraint density than the satisfiability threshold α_{SAT} in the large k-limit. For instance, the asymptotic bound for the clustering and satisfiability thresholds for the random k-SAT is $\alpha_{\text{clus}} \sim \frac{2^k}{k}(\log k + \log\log k + 1 - \log 2)$ [81] and $\alpha_{\text{SAT}} \sim 2^k \log 2$ [34] (more details can be found in Chapter 4).

Despite the detailed picture of the solution space geometry of the random CSPs, it remains an interesting research question to understand how the algorithms behave when they are trying to find a solution in the satisfiable regime. More specifically, researchers would like to determine the algorithmic threshold denoted as α_{alg} above which no algorithm is able to find a solution for the problem with high probability in polynomial time. Using numerical simulations, for small k-values one can able to design the algorithms efficiently where the clause to variable densities very close to the satisfiability threshold (α_{SAT}), whereas it cannot be possible to solve numerically for large values of k. In this context, Coja-Oghlan in [27] provides a polynomial time algorithm for random k-SAT upto the constraints densities coincide at leading order with the clustering threshold α_{clus}. Although this strands a broad range of the threshold value where typical instances have a non-empty set of solutions, but there is no known algorithm which can be able to find the solutions efficiently. In some cases it is proven that these algorithms fail to find the solutions [30, 50, 56]. Even if it is hard to predict the algorithmic threshold α_{alg} precisely in terms of structural phase transition one can predict the hypothesis that the clustering threshold is upper bound to the algorithmic one, $\alpha_{\text{alg}} \leq \alpha_{\text{clus}}$. The research in structural phase transitions within the satisfiability regime particularly the emergence of the clustering threshold α_{clus} in terms of long-range correlations, relies on the uniform distribution of all satisfying assignments. In this thesis, we also prove that throughout this satisfiable phase, the logarithm of the number of satisfying assignments of a random 2-SAT exhibits fluctuations of order $\sqrt{n}$, where n is the number of variables.

Going back to the statistical physics inspired perspective, one more prominent threshold arise in terms of characterization on the decay of correlation under the Gibbs (or Boltzmann) measure on the solution of the problem is called *"Gibbs Uniqueness threshold"* denoted as d_{uniq} [98]. This threshold is expressed in terms of the average degree of variable (d) and comes into picture when analyzing the infinite tree limit – Galton-Watson process approximation of the problem. Conversely, other two thresholds (satisfiability and clustering) are defined with respect to the constraint density (α). In sparse random graph models, these two parameters are closely related although d concerns with the local geometric structure of the factor graph, whereas α treats as a control parameter in the random CSP ensembles. For instance, in case of random k-SAT or random k-coloring, the constraint density α determines the expected average degree of variables d, making the uniqueness threshold relate with the other structural transitions. Specially, from an algorithmic prospective, the Gibbs uniqueness threshold plays an important role as the local algorithm such as Belief propagation is effective in this regime. In this thesis, we explicitly provide a lower bound on this Gibbs uniqueness threshold (d_{uniq}) that improves over prior work of Montanari and Shah [83]. Particularly, we prove that for any $k \geq 3$ for clause/variable ratios upto this uniqueness threshold of the corresponding Galton-Watson tree, the number of satisfying assignments of random k-SAT is given by the physics method called *'replica symmetry'* predicted in [77].

This manuscript is based on the three papers [21–23] that have been produced during my PhD. The remaining chapters are organized as follows.

- Chapter 2 introduces the definition and representation of constraint satisfaction problems. We also discuss one of the well-known CSP: the satisfiability problem and its importance in computer science. The chapter concludes with a brief discussion of several statistical physics quantities that will be used throughout the thesis.

- Chapter 3 focuses on message passing algorithms, including Belief Propagation and Warning Propagation, along with some variants derived from them. We discuss the applications and characteristics of algorithms such as *Belief Propagation Guided Decimation*, the *Decimation Process* (a statistical physics–inspired thought process), and the purely combinatorial algorithm namely *Unit Clause Propagation*. We also present one of the most useful algorithms employed in this thesis for estimating the logarithm of the partition function on random k-SAT (further details in Chapter 7).

- In Chapter 4 we examine the satisfiability transition and compares short-range and long-range correlations in random satisfiability problems, with a focus on random k-SAT and random k-XORSAT. We study the Gibbs measure on random CSPs and explore various phase transitions, including reconstruction, non-reconstruction, condensation, and Gibbs uniqueness, in the context of both random k-XORSAT and k-SAT. The chapter ends with a comparison of the different phases and the associated solution space geometry of random k-SAT versus random k-XORSAT.

- Chapter 5 addresses the number of satisfying assignments in random 2-SAT. This chapter contains the first main result of the thesis, establishing a central limit theorem for the number of solutions of random 2-SAT formulas – first time CLT on any kind of random satisfiability problems.

- Chapter 6 presents the second main result of the thesis, analyzing the performance of the Belief Propagation Guided Decimation algorithm on random k-XORSAT and comparing it with the statistical physics–inspired thought process namely *decimation process*. We also identify different phase transitions of the decimation process, pinpointing the regimes of d and θ where BPGD succeeds or fails.

- In Chapter 7 we revisit the Gibbs uniqueness threshold for random k-SAT and improves the lower bound established by Montanari and Shah [83]. Towards the proof of the result we introduce *'interpolation method'* and *'Aizenmann-Sims-Starr scheme'* for proving the matching upper and lower bound on the logarithm of number of satisfying assignments on random k-SAT upto Gibbs uniqueness threshold. We explicitly determine the number of satisfying assignments predicted by the statistical physics–inspired *replica symmetric solution*. The result comes in terms of the *Bethe free entropy* $\mathfrak{B}_{d,k}$ which is a function defined for a probability measure $\pi \in \mathscr{P}(0,1)$.

- The final chapter summarizes the results of the thesis, comparing them with existing work in the probabilistic combinatorics as well as statistical physics literature. We conclude by outlining several interesting open problems and possible directions for future research.

2
Models

"...random constraint satisfaction problems (CSPs) is the geometry of the space of satisfying or almost satisfying assignments ... for which a precise landscape of predictions has been made via statistical physics-based heuristics."

–Jun-Ting Hsieh et. al.

2.1 Constraint Satisfaction Problems

2.1.1 Definitions

A constraint satisfaction problem ('CSP') is defined as the set of n variables denoted as $V = \{x_1, x_2, \cdots, x_n\}$ are submitted to a set of a number of constraints $C = \{a_1, a_2, \cdots, a_m\}$ for some $m \in \mathbb{N}$. The variables $x_i, i \in \{1, \cdots, n\}$ take their values in a finite set Ω. Clearly, when $|\Omega| = 2$, the variables will be treated as Boolean variables: $\Omega = \{0, 1\}$. In statistical physics paradigm these are equivalent to spins $\sigma_i \in \{-1, 1\}$, using the change of variable $\sigma_i = 2x_i - 1$. When the set Ω takes an arbitrary integer q, the variables can be served as Potts spins or colors: $\Omega = \{1, 2, \cdots, q\}$. Further, we call $\sigma_x = \{x_1, x_2, \cdots, x_n\} \in \Omega^n$ as a configuration of the variables and for a subset $\omega \subseteq \{1, \cdots, n\}$ of variables, we call σ_{x_ω} its configuration. The clauses a_j with $j \in \{1, \cdots, m\}$ details a subset $\partial j \subseteq \{1, \cdots, n\}$ of variables [99] and put a constraint on the value of their configuration $\sigma_{x_{\partial j}}$. More specifically, when the constraint is satisfied the function $a_j \colon \Omega^{|\partial j|} \to \{0, 1\}$ assesses to value 1, otherwise to value 0.

There are several variations of CSPs. In the optimization version of the problem, the goal is to find an optimal configuration which minimizes the cost function. The cost function $\mathrm{E} : \Omega^n \to \mathbb{R}^+$ is defined as the counts of total number of constraints which are unsatisfied by a configuration σ:

$$E(\sigma_x) = \sum_{j=1}^{m} (1 - a_j(\sigma_{\partial j})) \tag{2.1.1}$$

The second type of the problem referred as decision making where the aim is to find a configuration $\sigma_x^\star$ with cost function $E(\sigma_x^\star) \leq E_0$, where E_0 is given as the threshold value for the cost function. When this threshold value becomes 0, then the goal will be to find such a configuration which satisfies all the constraints simultaneously. Such this configuration will be referred as the solution of the problem. In other way, a solution or, a satisfying assignment is a mapping $\sigma : V \to \Omega$ that satisfies every constraints. As long as computing the cost function is easy to evaluate, the decision making problem will be easier than the optimization version. Because in the decision case, once we have the optimal configuration we just need to compare it with the threshold value of the cost function. The third variant of the CSPs so called counting problem refers to count the number of satisfying assignments (or, solutions) of a given instance. Generally speaking, this version is much more harder as compared to the previous two variants. In this thesis, we compute such a problem on random 2-SAT where we count the number of satisfying assignments of random 2-SAT throughout the satisfiable regime. Moreover, in this thesis we also determine the number of *actual* satisfying assignments of random k-SAT formula for the clause-to-variable densities upto Gibbs Uniqueness threshold discussed in Chapter 4.

2.1.2 The SAT Problem

The satisfiability problem has a long and exciting history in probabilistic combinatorics as well as in computer science. Consider a boolean formula $\boldsymbol{F}$ consists of n boolean variables and m logical clauses $\{a_1, a_2, \cdots, a_m\}$ on the set of literals. Each literal l_i corresponds to a variable x_i can take either the value x_i or its negation $(\neg x_i)$. Each clause present in the SAT formula $\boldsymbol{F}$ is a disjunction (logical OR $(\vee)$) on the literals and are of the form

$$a_j = l_{i(1)} \vee l_{i(2)} \vee l_{i(3)} \vee \cdots \vee l_{i(\partial j)} \tag{2.1.2}$$

where each literals l_i are formed from the variables in ∂j. So, the variables in ∂j can take values from $2^{|\partial j|}$ possible combinations out of which in one case only the clauses get violated when all the literals take the value 0. Then the satisfiability formula $\boldsymbol{F}$ is the conjunction (logical AND$(\wedge)$) over the set of clauses and are of the form

$$\boldsymbol{F} = a_1 \wedge a_2 \wedge a_3 \wedge \cdots \wedge a_m \tag{2.1.3}$$

and is also called a CNF formula (conjunctive normal form). Subsequently, the formula $\boldsymbol{F}$ satisfies i.e., evaluates to 1 if and only if all the clauses present in $\boldsymbol{F}$ are evaluate to 1. Below example 2.1.1 provides a SAT formula which consists of eight variables $\{x_1, x_2, \cdots, x_8\}$ and four clauses $\{a_1, \cdots, a_4\}$. Moreover, a satisfying assignment to the below toy example is given by $\sigma(x_1) = 1, \sigma(x_2) = 1, \sigma(x_3) = 0, \sigma(x_4) =$

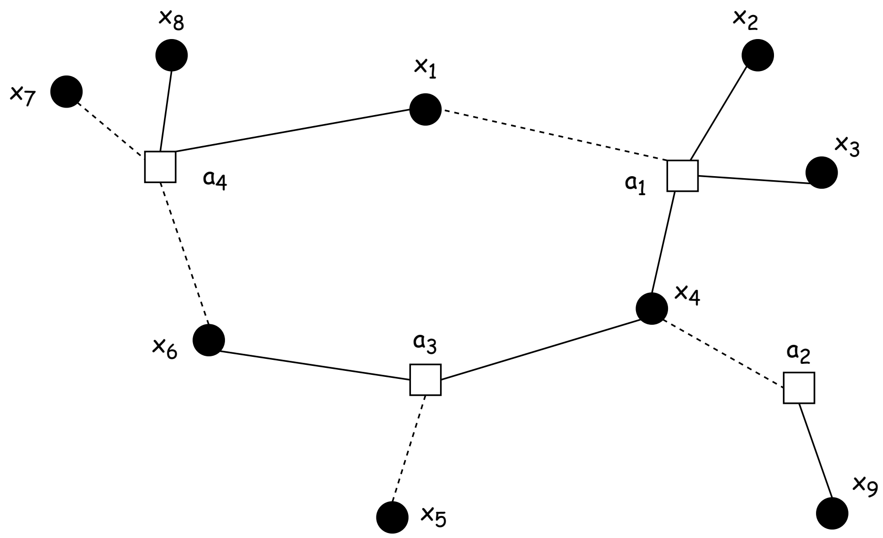


Figure 2.1: Factor Graph representation of the SAT formula in example 2.1.1

$0, \sigma(x_5) = 1, \sigma(x_6) = 1, \sigma(x_7) = 0, \sigma(x_8) = 1, \sigma(x_9) = 1.$

Example 2.1.1. $\boldsymbol{F} = \underbrace{(\neg x_1 \vee x_2 \vee x_3 \vee x_4)}_{a_1} \wedge \underbrace{(\neg x_4 \vee x_9)}_{a_2} \wedge \underbrace{(x_4 \vee \neg x_5 \vee x_6)}_{a_3} \wedge \underbrace{(x_1 \vee \neg x_6 \vee \neg x_7 \vee x_8)}_{a_4}$

In the decision version of this problem, a satisfying assignment is the configuration σ such that our formula $\boldsymbol{F}$ evaluates to 1. Thus the instance of a SAT problem is defined by the number of variables (n) and clauses (m) as well for each clause a the given choice of the subset ∂j and for each variable appears in that subset ∂j the choice of the literal l_i appears in the clause a_j by describing the formula $\boldsymbol{F}$. Using the spins $\sigma_i \in \{-1, 1\}$ with changing variable $\sigma_i = 2x_i - 1$ used in the context of Ising model in statistical physics paradigm, the clauses can be written as follows:

$$a_j(\sigma_{\partial j}) = 1 - \prod_{i \in \partial j} \mathbb{1}[\sigma_i = -\Upsilon_i^j] \tag{2.1.4}$$

where,
$$\Upsilon_i^j = \begin{cases} 1 & \text{if the literal is } x_i \text{ and,} \\ -1 & \text{if the literal is } \neg x_i \end{cases}$$

2.1.3 Why SAT ?

So far, we have discussed the basic definitions of constraint satisfaction problems and the satisfiability (SAT) problem which is a member of a larger family of CSPs. The obvious question comes into reader's mind that why we study the satisfiability problems in computer science. The SAT problems play a crucial role in theoretical computer science and sit in a prominent place among all NP-complete problems. This is because SAT is both simple by empowering combinatorial reasoning and general enough to model any kind of other problems in a quite natural style.

(i) ***SAT is simple to describe:*** In complexity theory, one of the most general theorem concerning the NP-complete problems is "*NDTM-ACCEPTANCE*" which states: given the description of a non-deterministic turing machine M, an input string x and the number of steps t, does M accept x within t steps? Although this problem is more general than SAT, but not simple. General in the sense that the theorem "NDTM-ACCEPTANCE is NP-complete" becomes a triviality whereas on the other hand the Cook-Levin theorem "3-*SAT is NP-complete*" is one of the fundamental results making SAT much more pliable by allowing combinatorial reasoning than the non-deterministic turing machine. One of the most interesting and deep research topic of SAT problems say random k-SAT, intrinsically defining a probability distribution on k-SAT formulas exhibits many interesting phenomena. By contrast, defining a probability distribution on the tuples (M, x, t), makes the instance of NDTM-ACCEPTANCE much less natural as well less interesting as compared to SAT.

(ii) ***SAT is general:*** Beside its simplicity regarding the combinatorial structure, it is still general enough to model a wide range of problems in a quite natural fashion. Thus it can be served as a 'modeling language' for many problems. In complexity theory any NP-complete problem can model another NP problem via reduction if there exists a polynomial time algorithm for that problem. But some of these reductions are more or less straightforward, whereas some are not. Let's consider the problem "*Hamiltonian Path*" where we have given a graph G and we know some of the edges say e_1, e_2, e_3 are the part of the Hamiltonian path. Then can we be able to extend this to complete a Hamiltonian Path? So, formulating this problem as a SAT problem is quite straightforward whereas the reduction from SAT to Hamiltonian Path requires the design of several clever mechanism.

2.1.4 Factor Graphs

Coming to the graphical representation of an instance of a CSP, the most prominent way to represent any CSP problem is using a graph called *factor graph* or, *tanner graph*. This is a bipartite graph $\mathbb{G} = (V, C, E)$ consists of two types of nodes where the first type $V = \{x_1, x_2, \cdots, x_n\}$ represents the variables (referred as variable nodes) and the second type $C = \{a_1, a_2, \cdots, a_m\}$ represents the clauses (referred as check/factor nodes). E is the set of edges connecting the variable and check nodes. There will be an edge between a variable node (x_i) and a check node (a_i) if the variable x_i appears in the check a_i either in original form (x_i) or as its negation $(\neg x_i)$. Furthermore, there can be a weight function associated with each check node a_i where these weights are linked to a probability distribution called *Boltzmann (Gibbs) distribution* described later in this section.

Figure 2.1 is a factor graph representation corresponds to the SAT formula $\boldsymbol{F}$ given in example 2.1.1. In the figure, the variable nodes $V = \{x_1, x_2, \cdots, x_9\}$ are represented by filled circles whereas the check nodes $C = \{a_1, \cdots, a_4\}$ are represented by the empty square. The set E is represented by both solid and dashed edge, where the solid link between a variable and check node is referred as the variable appears as its original form in that check, whereas the dashed link represents that the variable appears

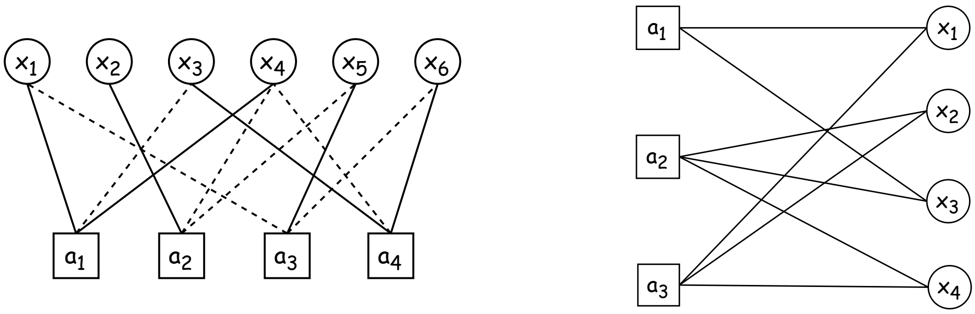


Figure 2.2: **Left:** A factor Graph representation of a 3-SAT formula $\boldsymbol{F}_{3\text{SAT}} = (x_1 \vee \neg x_3 \vee x_4) \wedge (x_2 \vee \neg x_4 \vee \neg x_5) \wedge (\neg x_1 \vee x_5 \vee \neg x_6) \wedge (x_3 \vee \neg x_4 \vee x_6)$. **Right:** A factor graph representation of a random linear system of equations (2.1.5) over $\mathbb{F}_2$.

as negation in that check node. Equivalently, the factor graph $\mathbb{G}$ also can be viewed as a *hyper-graph* where the variables are still represented as vertices but the clauses are now represent as hyper-edges which link a subset of vertices with length > 2. In this thesis we are mainly interested in the factor graph associated with the random k-SAT model and the random matrix over finite field. Let's talk about the representation of the factor graph corresponds to these two models briefly.

(i) ***random** k**-SAT:*** The most common example of SAT problem is random k-SAT where each clause can take exactly k-variables. Similar to other factor graph representations, the set V represents the boolean variables and the set C represents the clauses and there will be an edge between a variable node and a check node if and only if that particular variable is present in that check. The variables are denoted by a circle and the check/factor nodes are denoted by a square. Figure 2.2 (left) is a simple factor graph representation of a k-SAT formula with $k = 3$.

(ii) ***random linear equations over** $\mathbb{F}_2$**:*** Coming to the random linear system of equations unlike random k-SAT, it is easier to grab. Consider a linear system of equation $\boldsymbol{A}x = \boldsymbol{b}$ over $\mathbb{F}_2$ where $\boldsymbol{A}$ is an $n \times n$ matrix with each entry 1 with probability d/n where d is the average variable degree. Now, given a random vector $\boldsymbol{b} = \{0,1\}^m$, our goal is to design a factor graph corresponding to the system of linear equations. Resembling to k-SAT, here also, the set V represents the variables designed by a circle and the check nodes C represents the equations designed by a square. There will be an edge between a variable and a check if and only if that particular variable will appear in that equation. Let's take a toy example of this kind below:

$$\underbrace{\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}}_{\boldsymbol{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}}_{\boldsymbol{b}} \tag{2.1.5}$$

Finding the set of all possible solutions of the vector $\boldsymbol{X}$ of the system of linear equations $\boldsymbol{A}x = \boldsymbol{b}$ over $\mathbb{F}_2$ is referred as a well-known random CSP called random k-XORSAT where each equation contains exactly k variables. In other words the matrix $\boldsymbol{A}$ has exactly k ones in each row. So, as compared to k-SAT, the disjunction $\vee$ (OR) inside a clause is replaced by XOR denoted by $\oplus$:

$$a_j = l_{i(1)} \oplus l_{i(2)} \oplus l_{i(3)} \oplus \cdots \oplus l_{i(k)} \tag{2.1.6}$$

Equivalently, one can rewrite this as the sum $\sum_{p=1}^{k} l_{i(p)}$ of literals inside a clause equals to 1 modulo 2. One of the best known algorithms for solving this linear system of equations in polynomial time $O(n^3)$ is *Gaussian elimination* where the number of equations $m = \Theta(n)$ with n is the total number of variables present in those equations. Moreover, this k-XORSAT can be random when the clauses (equations) are drawn independently and uniformly at random from the all possible $2^k \binom{n}{k}$ XOR-clauses on the set of variables $V = \{x_1, x_2, \cdots, x_n\}$. Due to its algebraic structure, any algorithm is always easy to analyze on this model. Later in this thesis, we will look over the performance of such a physics inspired algorithm on random k-XORSAT model. Figure 2.2 (right) shows the factor graph representation of the linear system of equations over $\mathbb{F}_2$ as given in equation (2.1.5).

2.2 Statistical Physics and CSPs

One of the most striking phenomenon of science is to deal with the ever growing variety of states of matter with various properties. Here the statistical physics comes into picture which aims to explain how the complex behaviors can emerge when a large numbers of identical elementary component interact with each other. It relies on two notable steps, in one hand passing the idea from the deterministic law of physics to a probabilistic description, on the other hand it starts from a probabilistic description and tries to recover that determinism by law of large numbers at a macroscopic level. In this section we will discuss some of the basic properties of statistical physics and its connection with the constraint satisfaction problems.

2.2.1 Boltzmann (Gibbs) probability distribution

From equation (2.1.1), there exists a cost function $E : \Omega^n \to \mathbb{R}^+$ which counts the total number of clauses violated by a given assignment $\sigma_x \in \Omega^n$ of the n variables. In the mathematical optimization problem one always aims to minimize this cost function defined over the set of all possible configurations Ω^n. Once the configuration space Ω and the cost function E are fixed, the Boltzmann probability distribution for the system to be found in the set of configuration is given by,

$$\mu_\beta(\sigma_x) = \frac{1}{Z(\beta)} e^{-\beta E(\sigma_x)} \tag{2.2.1}$$

where, the normalization constant $Z(\beta)$ is known as ***Partition function*** in physics jargon and is equal to

$$Z(\beta) = \sum_{\sigma_x \in \Omega^n} e^{-\beta E(\sigma_x)} \tag{2.2.2}$$

The real parameter $T = 1/\beta$ is the temperature with β refers as the inverse temperature. In the context of CSP, to emphasize the factor graph $\mathbb{G}$ by introducing a weight function $\psi_{a_j} : \Omega^{\partial a_j} \to (0, \infty)$ to each constraint a_j, the Boltzmann distribution can be re-written as

$$\mu(\sigma_x) = \frac{1}{Z_{\mathbb{G}}} \prod_{j=1}^{m} \psi_{a_j}(\sigma_{x_{\partial j}}) \qquad \text{where, } Z_{\mathbb{G}} = \sum_{\sigma_x \in \Omega^n} \prod_{j=1}^{m} \psi_{a_j}(\sigma_{x_{\partial j}}) \tag{2.2.3}$$

However equations (2.2.1)–(2.2.2) interpolates smoothly between numerous interesting situations. In the high-temperature limit ($\beta \to 0$), one can recover uniform probability distribution whereas in the low-temperature limit ($\beta \to \infty$) it concentrates on the global maxima of the original distribution. Specifically, in the $\beta \to \infty$ limit, a configuration $\sigma_{x_0} \in \Omega$ such that $E(\sigma_x) \geq E(\sigma_{x_0})$ for any $\sigma_x \in \Omega$ is called a *ground state* with Ω_0 denotes the set of all ground states and the corresponding energy $E_0 = E(\sigma_{x_0})$ is called ***ground state energy***. Therefore,

$$\lim_{\beta \to 0} \mu_\beta(\sigma_x) = \frac{1}{|\Omega|} \quad \text{and} \quad \lim_{\beta \to \infty} \mu_\beta(\sigma_x) = \frac{1}{|\Omega_0|} \mathbb{I}(\sigma_x \in \Omega_0) \tag{2.2.4}$$

Also, in this setting the cost function E is termed as the ***Hamiltonian*** or the energy function and is defined as,

$$E_{\mathbb{G}}(\sigma_x) = -\log \prod_{j=1}^{m} \psi_{a_j}(\sigma_{x_{\partial j}}) \tag{2.2.5}$$

The most important thermodynamic potential in this regards is the ***Free energy*** which is defined as

$$F_{\mathbb{G}}(\beta) = -\frac{1}{\beta} \log Z(\beta) \tag{2.2.6}$$

whereas in calculations, it is often more convenient to use the ***Free entropy*** given by,

$$\Phi_{\mathbb{G}}(\beta) = -\beta F_{\mathbb{G}}(\beta) = \log Z(\beta) \tag{2.2.7}$$

2.2.2 Some statistical physics models

In mathematical physics a wide range of interesting phenomenon occur when we make the number of variables $n \to \infty$. From the above section it is clear to see that there is a direct map between CSPs and the statistical physics problems. Let's take an example of a spin glass models which are the generalization of Ising model with the variables treated as spins σ_j with $\sigma_j \in \{-1, +1\}$ and the coupling

J^a either takes value from $\mathbb{R}$ or from $\{-1,+1\}$. Therefore, the energy function is defined as,

$$E(\sigma_x) = -\sum_{a=1}^{m} J^a \prod_{j\in\partial a} \sigma_j \tag{2.2.8}$$

But when it comes to the interaction between general spin-glass model to p-body, p-spin model comes into picture. One of the most famous well-known model in this regard is the *Edward-Anderson* model when $p = 2$. Moreover, the spin glass model can be formatted as a constraint satisfaction problem in the $\beta \to \infty$ limit as the Boltzmann distribution in (2.2.1) minimizes our energy function $E(\sigma_x)$. For a two body interaction, the variables should have either ***ferromagnetic*** (where $J^a > 0$ and $\prod_{j\in\partial a}\sigma_j = 1$) or, ***antiferromagnetic*** (where $J^a < 0$ and $\prod_{j\in\partial a}\sigma_j = -1$). The general idea behind the ferromagnetic regime is that when β is large (low temperature case), one of the spins begins to dominate the others and the system shows a positive or negative magnetization. On the other hand when β is small (high temperature case), there is no magnetization has been observed and the regime is called ***paramagnetic***. Then the obvious observation is to pinpoint the critical inverse temperature ($\beta_{\textbf{crit}}$) where the phase transition occurs i.e., the system suddenly switches from paramagnetic to ferromagnetic.

Coming to the random k-SAT model, for each constraint a_j with $j \in [m]$ and for some $\beta > 0$ we can rewrite the weight function and Hamiltonian as follows

$$\psi_{a_j}(\sigma_{a_{\partial j}}) = \prod_{j=1}^{m} \exp\left(-\beta \cdot \mathbb{1}\{\sigma \nvDash a_j\}\right) \qquad\qquad E(\sigma) = \beta \cdot \sum_{j=1}^{m} \mathbb{1}\{\sigma \nvDash a_j\}$$

Here the Hamiltonian counts the number of unsatisfied assignments and a penalty of $-\beta$ is imposed to the satisfied clauses. As a result the partition function $Z_{\mathbb{G}}(\beta)$ approximates the number of satisfying assignments by taking the inverse temperature β to infinity. Moreover, when $\beta = \infty$, the Gibbs distribution is the uniform distribution over the solution space as the unsatisfied clauses gets a zero penalty and therefore the partition function $Z_{\mathbb{G}}(\beta)$ counts the number of solutions exactly. However, it is easier to handle the finite β case and take limit after.

Coming to random k-XORSAT which boils down to a problem of random linear system of equations over $\mathbb{F}_2$, let's take for the clause/equation a_j the weight is given by,

$$\psi_{a_j}(\sigma_{a_{\partial j}}) = \prod_{i=1}^{n} \mathbb{1}\Big\{\sum_{j=1}^{n} \boldsymbol{A}_{ij}\sigma_{x_j=0}\Big\} \qquad\qquad E(\sigma) = \sum_{i=1}^{n} \mathbb{1}\Big\{\sum_{j=1}^{n} \boldsymbol{A}_{ij}\sigma_{x_j=0}\Big\}$$

where the partition function $Z_{\mathbb{G}}$ can be computed as the cardinality of the kernel of $\boldsymbol{A}(\mathbb{G})$ and can be written as,

$$|\ker \boldsymbol{A}(\mathbb{G})| = Z_{\mathbb{G}} = \sum_{\sigma_x\in\mathbb{F}_2^n} \prod_{i=1}^{n} \mathbb{1}\Big\{\sum_{j=1}^{n} \boldsymbol{A}_{ij}\sigma_{x_j=0}\Big\}$$

The weight function $\psi_{a_j}(\sigma_{a_{\partial j}})$ can be extended by allowing the value zero when the equation is unsatisfied, but the Gibbs distribution is always well-defined as the zero vector always belongs to the

Kernel of the matrix $\boldsymbol{A}$ and therefore, $Z_{\mathbb{G}} = |\ker \boldsymbol{A}| > 0$.

In view of weight function (sometimes called *compatibility function*) defined in (2.2.3), ψ_{a_j} broadcast the temperature to the variable nodes in the factor graph $\mathbb{G}$. Then for analyzing the partition function one needs to look at the normalized limit of the "free entropy ($\Phi_{\mathbb{G}}(\beta)$)" referred as ***free entropy density*** ($\phi_{\mathbb{G}}$) and defined as,

$$\phi_{\mathbb{G}} = \phi_{\mathbb{G}}(\beta) = \lim_{n\to\infty} \frac{1}{n}\Phi_{\mathbb{G}}(\beta) = \lim_{n\to\infty} \frac{1}{n}\log Z(\beta) \tag{2.2.9}$$

A clear observation on the Free entropy density $\phi_{\mathbb{G}}$ reveals that there can happen a phase transition at the singularities of it if and only if $\phi_{\mathbb{G}}$ is non-analytic. Two common types of phase transition in this regard are the first and second order phase transition. The first order phase transition occurs when the first derivative of $\phi_{\mathbb{G}}$ w.r.t. β i.e., $\frac{\partial}{\partial\beta}\phi_{\mathbb{G}}$ is discontinuous at some $\widetilde{\beta}$ and similarly for the second order phase transition when $\frac{\partial}{\partial\beta^2}\phi_{\mathbb{G}}$ is discontinuous. Concurrently the higher order phase transitions may occur and described accordingly. Therefore, Free entropy density $\phi_{\mathbb{G}}$ is one of the most important entity for understanding the physical system and its changing behavior. A heuristic towards Free entropy density $\phi_{\mathbb{G}}$ is ***Belief Propagation***, a statistical physics inspired message passing algorithm, which we will describe in details in the next chapter.

3

Message Passing Algorithms

"Message passing algorithms have proved surprisingly successful in solving hard constraint satisfaction problems on sparse random graphs."

– Andrea Montanari et.al.

Consider a universal problem of computing marginals of a graphical model with n variables denoted as $V = \{x_1, x_2, \cdots, x_n\}$ taking values in a finite set Ω. One naive approach for computing such marginals is to take the sums over all configurations with time complexity $|\Omega|^n$. But when we are talking about tree factor graphs, computing marginals on such model takes time grows linearly with n. This can be done through a *'dynamic programming'* which recursively sums over all variables starting from the leaves and moving towards the root of the tree. Such a recursive procedure is remodeled as a distributed *'message passing'* algorithm. These algorithms operate on 'messages' associated with the edges of the factor graph and update the messages recursively based on local computations done at the vertices of the graph. In this chapter we will discuss few of such message passing methods along with few algorithms associated with these methods.

3.1 Belief Propagation

Belief Propagation (in short BP) is one of most well known iterative message passing procedure for computing marginals as well as to compute the partition function Z with respect to a measure μ (defined in (2.2.3)) of a variable x_i or any subset of variables. Moreover, it provides an efficient way to sample a configuration σ_x from μ and the best thing is all these computations can be achieved in polynomial time with respect to the sample size n. It is straightforward to prove that BP computes such marginals exactly on trees. For this purpose BP is extremely effective in the case of loopy graphs as well. The basic intuition behind this success is that BP as a local message passing procedure should be successful when the underlying model is a locally tree like structure. There are many applications of these type of factor graph models appear frequently in the field of probabilistic combinatorics as well as in statistical physics. Despite these advantages, BP becomes ineffective in the emergence of long-range correlations which in turn lead to a phase transition. In the later chapters we will see few of such application.

3.1.1 BP messages

In 1962, R.G.Gallager [49] introduced BP messages for decoding the low density parity check matrix and in 1988 by J.Pearl [92] it was again launched for the first time in the context of probabilistic inference. Lets define two type of messages $\nu^{(t)}_{x\to a}$ and $\hat{\nu}^{(t)}_{a\to x}$ associated with each edge $(x,a)\in E$ of the factor graph $\mathbb{G}=(V,C,E)$ at step t. Out of these twos the message $\nu^{(t)}_{x\to a}$ is going from a variable to a check node whereas the message $\hat{\nu}^{(t)}_{a\to x}$ is going from the check node to a variable node. More specifically, $\nu^{(t)}_{x\to a}$ is the marginal of the variable x_i when the check node a is removed whereas $\hat{\nu}^{(t)}_{a\to x}$ is the marginal of a variable x_i when all the check nodes in $\partial i\setminus a$ have been discarded. As the messages are dependent on the time parameter t, so for each $t>0$ both messages are probability distributions over Ω. Initially, both the messages are the uniform distribution over Ω i.e., $\nu^{(0)}_{x\to a}(s)=\hat{\nu}^{(0)}_{a\to x}(s)=1/\Omega$ for all $x\in V$, $a\in C$ and $s\in\Omega$. One can also initialize the messages by drawing i.i.d from a probability distribution $\mathbf{P}$ on $\mathscr{P}(\Omega)$. Furthermore, the BP equations [60] on a tree consists of the set of messages $\{\nu^{(t)}_{x\to a},\hat{\nu}^{(t)}_{a\to x}\}_{(x,a)\in E}$ with $t>0$ are given by,

$$\nu^{(t+1)}_{x\to a}(s)=\frac{1}{Z_{x,a}}\prod_{b\in\partial x\setminus a}\hat{\nu}^{(t)}_{a\to x}(s) \tag{3.1.1}$$

$$\hat{\nu}^{(t)}_{a\to x}(s)=\frac{1}{\hat{Z}_{x,a}}\sum_{\sigma\in\Omega^{\partial a}}\mathbb{1}\{\sigma_x=s\}\psi_a(\sigma)\prod_{y\in\partial a\setminus x}\nu^{(t)}_{y\to a}(\sigma_y) \tag{3.1.2}$$

where the $Z_{x,a}$,$\hat{Z}_{x,a}$ are the normalization constants of the messages. The equations (3.1.1)–(3.1.2) are referred as the Belief Propagation (BP) equations for which one can consider the fixed point equations. Moreover, all the messages are updated in parallel. It is clear from the above BP equations that if $\partial x\setminus a=\emptyset$ then equation 3.1.1 is the uniform distribution over Ω. Similarly, if $\partial a\setminus x=\emptyset$ then $\hat{\nu}^{(t+1)}_{a\to x}(s)\propto\psi_a(s)$. Figure 3.1 shows a pictorial illustration of the BP equation update rules. Moreover, the algorithm 3.1 (also see [66]) provides the iterative procedure for finding a solution of the BP equations (3.1.1–3.1.2).

The obvious question is that under which condition(s) the messages converge to a limit $(\nu^*_{x\to a},\nu^*_{a\to x})$. From [66], it is clear that on a tree of diameter $t_{\max}$ the algorithm 3.1 guarantees to find the set of messages exactly, independently of the choice of the initialization and the updating rules given in (3.1.1)–(3.1.2). Moreover in this algorithm we haven't specified any ordering of the edges (x,a) for the update of the messages. So, reshuffling the ordering, taking any random permutation of edges before

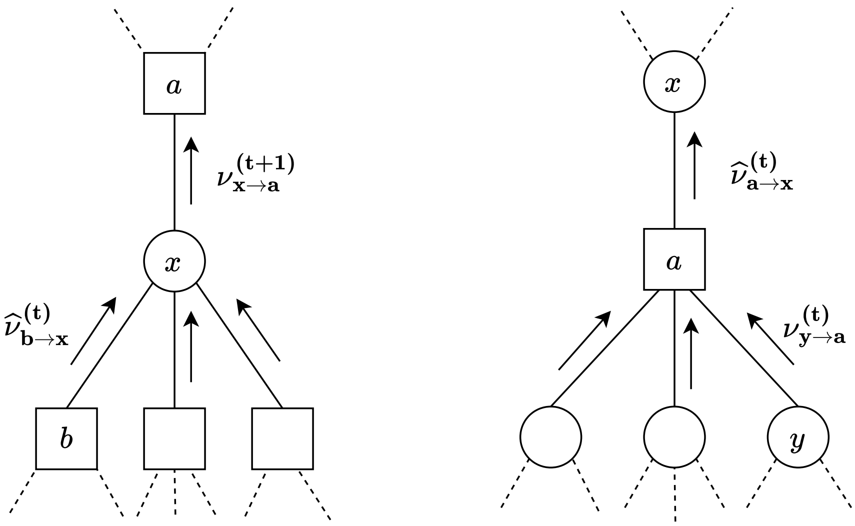


Figure 3.1: **Left:** Factor graph involved in computing $\nu^{(t+1)}_{x\to a}$ which is a function of all 'incoming messages' $\hat{\nu}^{(t)}_{b\to x}$ with $b \neq a$. **Right:** Factor graph involved in computing $\hat{\nu}^{(t)}_{a\to x}$ which is a function of all 'incoming messages' $\nu^{(t)}_{y\to a}$ with $y \neq x$.

Input: a factor graph $\mathbb{G} = (V, C, E)$, set of functional nodes $\{\psi_a\}_{a\in C}$, precision accuracy ϵ, maximum number of iterations $t_{\max}$.
Output: A set of messages $\{\nu(\cdot), \widehat{\nu}(\cdot)\}$, or state 'Not Converge' if fails.

Initialization: For each edge $\{x, a\} \in E$, initialize $\nu_{x\to a}(\cdot)$ and $\widehat{\nu}_{a\to x}(\cdot)$ as i.i.d. random variables with distribution $\mathbf{P}$.
for $t = 0, \cdots, t_{max}$ **do**
  Compute two messages: first $\{\widehat{\nu}^{(t)}_{a\to x}\}_{(x,a)\in E}$, then $\{\nu^{(t+1)}_{x\to a}\}_{(x,a)\in E}$ using (3.1.1) **if** δ *(maximum message change)* $< \epsilon$ **then**
    return set of messages $\{\nu^{(t+1)}_{x\to a}, \widehat{\nu}^{(t)}_{a\to x}\}$
return "Not Converged".

Algorithm 3.1: Belief Propagation algorithm [66]

each updating of messages are allowed as BP is used as a heuristic on factor graphs with loops without guarantee of convergence.

3.1.2 Computing marginals

Our next goal is to compute the marginals $\mu(x)$ of a variable $x \in V$. Since we have in our hand the solution of the BP equations from algorithm 3.1, using the Markov property and the finite tree factor graph model we can construct the marginal [66] of the variable x using:

$$\mu(x) \propto \prod_{a\in\partial x} \left(\sum_{\sigma_{\partial b\setminus x}, s\in\Omega} \mathbb{1}\{\sigma_x = s\} \psi_b(\sigma_{\partial b}) \prod_{y\in\partial b\setminus x} \nu_{y\to b}(\sigma_y) \right) \tag{3.1.3}$$

Then using the equation (3.1.2) we get:

$$\mu(x) = \frac{1}{Z_x} \prod_{a \in \partial x} \widehat{\nu}_{a \to x}(s) \tag{3.1.4}$$

where Z_x is a normalization constant. So far, we see how the marginal computation of a variable using the BP equations and know that the BP messages converge to a limit $(\nu^*_{x\to a}, \nu^*_{a\to x})$. The next question should come into our mind that if the limit(s) does exist then what is/are the significance of such limit(s)? Although the marginals $\mu(x)$ for all $x \in V$ are computed exactly on tree factor graph model using the BP fixed points, but generally this is not true always because of several limits depend on the initialization of factor graph which contains short cycles (cycles with bounded length). However, BP provides a good approximation on the marginal computation with the help of a correct initialization if the corresponding factor graph doesn't contain too many short cycles. Furthermore, in case of tree factor graph model one can express the free entropy density $\Phi = \frac{1}{n}\log Z$ (where Z is referred as the partition function) from the set of messages which is the solution of BP equation (3.1.1)–(3.1.2) is known as *Bethe Free Entropy* (denoted as Φ^{Bethe}) in physics jargon. In the next subsection we will see in details of this quantity.

3.1.3 Bethe-Free Entropy

In 1935 German-American physicist Hans Albrecht Eduard Bethe in [14] first introduced the Bethe free entropy density for the ferromagnetic Ising model. Using the decomposition property of Gibbs distribution, we introduce the Bethe Free Entropy Φ^{Bethe} in terms of $2|E|$ BP messages $\{\nu_{x\to a}(\cdot), \widehat{\nu}_{a\to x}(\cdot)\}$ and can be expressed as [66]

$$\Phi^{\text{Bethe}} = \frac{1}{n}\left[\sum_{x\in V} \mathscr{B}^{(t)}_x + \sum_{a\in C} \mathscr{B}^{(t)}_a - \sum_{(xa)\in E} \mathscr{B}^{(t)}_{ax}\right] \tag{3.1.5}$$

where

$$\mathscr{B}^{(t)}_x = \log\Big[\sum_{\sigma\in\Omega^{\partial a}} \psi_a(\sigma) \prod_{x\in\partial a} \nu^{(t)}_{x\to a}(\sigma_x)\Big],$$

$$\mathscr{B}^{(t)}_a = \log\Big[\sum_{s\in\Omega} \prod_{a\in\partial x} \widehat{\nu}^{(t)}_{a\to x}(s)\Big] \text{ and}$$

$$\mathscr{B}^{(t)}_{ax} = \sum_{x\in\partial a} \log\Big[\sum_{s\in\Omega} \nu^{(t)}_{x\to a}(s)\cdot\widehat{\nu}^{(t)}_{a\to x}(s)\Big]$$

Roughly speaking, $\mathscr{B}^{(t)}_x$ corresponds to the contribution of the variable nodes to the partition function, $\mathscr{B}^{(t)}_a$ to the contribution of the check nodes and $\mathscr{B}^{(t)}_{ax}$ to the contribution of the edges. The aim is that,

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\log Z] = \lim_{t\to\infty}\lim_{n\to\infty} \Phi^{\text{Bethe}} \tag{3.1.6}$$

The limit $\nu^*_{x\to a}, \nu^*_{a\to x}$ which are the fixed points of the BP equations (3.1.1)–(3.1.2) (if they exist) are expected to correspond to the stationary points of Φ^{Bethe} [101]. The quantity such as the BP messages and the Bethe-Free entropy are model specific; i.e., they depend on the factor graph or the set Ω and the convergence problem on top of this.

In this thesis, we find an explicit expression for the logarithm of the number of solutions of a random k-SAT formula $\boldsymbol{F} = \boldsymbol{F}_{d,k}(n)$ for every d within the Gibbs uniqueness threshold (discuss in Chapter 4). The result comes in terms of the Bethe-Free entropy $\mathfrak{B}_{d,k}$ which is a function defined for a probability measure $\pi \in \mathscr{P}(0,1)$. We discuss the details of this result in Chapter 7.

3.2 Algorithms

In this section we discuss a few algorithms/processes on random k-XORSAT/k-SAT instances which we used in this thesis with the technique of Belief Propagation.

3.2.1 Belief Propagation Guided Decimation

In early 2000s, physicists have proposed a message passing algorithm called *Belief Propagation Guided Decimation* (BPGD) which performs impressively on various random CSPs [72, 96] according to the computer experiments.

BPGD sets its ambitions higher than merely finding a solution to the k-XORSAT instance $\boldsymbol{F}$: the algorithm attempts to sample a solution uniformly at random. To this end BPGD assigns values to the variables $x_1, \ldots, x_n$ of $\boldsymbol{F}$ one after the other. In order to assign the next variable the algorithm attempts to compute the marginal probability that the variable is set to 'true' under a random solution to the k-XORSAT instance, given all previous assignments. More precisely, suppose BPGD has assigned values to the variables $x_1, \ldots, x_t$ already. Write $\boldsymbol{\sigma}_{\text{BP}}(x_1), \ldots, \boldsymbol{\sigma}_{\text{BP}}(x_t) \in \{0,1\}$ for their values, with 1 representing 'true' and 0 'false'. Further, let $\boldsymbol{F}_{\text{BP},t}$ be the simplified formula obtained by substituting $\boldsymbol{\sigma}_{\text{BP}}(x_1), \ldots, \boldsymbol{\sigma}_{\text{BP}}(x_t)$ for $x_1, \ldots, x_t$. We drop any clauses from $\boldsymbol{F}_{\text{BP},t}$ that contain variables from $\{x_1, \ldots, x_t\}$ only, deeming any such clauses satisfied. Thus, $\boldsymbol{F}_{\text{BP},t}$ is a XORSAT formula with variables $x_{t+1}, \ldots, x_n$. Its clauses contain at least one and at most k variables, as well as possibly a constant (the XOR of the values substituted in for $x_1, \ldots, x_t$).

Input: a random k-XORSAT formula $\boldsymbol{F}$ with variables $x_1, \ldots, x_n$ conditioned on being satisfiable
Output: an assignment $\boldsymbol{\sigma}_{\text{BP}} : \{x_1, \ldots, x_n\} \to \{0,1\}$.

for $t = 0, \ldots, n-1$ **do**
    compute the BP approximation $\mu_{\boldsymbol{F}_{\text{BP},t}}$;
    set $\boldsymbol{\sigma}_{\text{BP}}(x_{t+1}) = \begin{cases} 1 & \text{with probability } \mu_{\boldsymbol{F}_{\text{BP},t}} \\ 0 & \text{with probability } 1 - \mu_{\boldsymbol{F}_{\text{BP},t}} \end{cases}$;
return $\boldsymbol{\sigma}_{\text{BP}}$;

Algorithm 3.2: The BPGD algorithm (Section 1.2, [22]).

Let $\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}$ be a uniformly random solution of the XORSAT formula $\boldsymbol{F}_{\mathrm{BP},t}$, assuming that $\boldsymbol{F}_{\mathrm{BP},t}$ remains satisfiable. Then BPGD aims to compute the marginal probability $\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{BP},t}\right]$ that a random satisfying assignment of $\boldsymbol{F}_{\mathrm{BP},t}$ sets x_{t+1} to true. This is where Belief Propagation ('BP') comes in. An efficient message passing heuristic for computing precisely such marginals, BP returns an 'approximation' $\mu_{\boldsymbol{F}_{\mathrm{BP},t}}$ of $\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{BP},t}\right]$. Having computed the BP 'approximation', `BPGD` proceeds to assign x_{t+1} the value 'true' with probability $\mu_{\boldsymbol{F}_{\mathrm{BP},t}}$, otherwise sets x_{t+1} to 'false', then moves on to the next variable. The pseudocode is displayed as Algorithm 3.2.

Remark 3.2.1.

- *If the BP approximations are exact, i.e., if* $\boldsymbol{F}_{\mathrm{BP},t}$ *is satisfiable and* $\mu_{\boldsymbol{F}_{\mathrm{BP},t}} = \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{BP},t}\right]$ *for all* t*, then Bayes' formula shows that BPGD outputs a uniformly random solution of* $\boldsymbol{F}$*. However, there is no universal guarantee that BP returns the correct marginals.*
- *Due to the algebraic structure of the XOR operation, BPGD is easier to analyze on random* k*-XORSAT and in fact the marginal probabilities are guaranteed to be half integral as seen in below Fact 3.2.2 i.e.,*

$$\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{BP},t}\right] \in \{0, 1/2, 1\}. \tag{3.2.1}$$

Fact 3.2.2. *The BP messages and marginals are half-integral for all* t*, i.e., for all* $t \geq 0$ *and* $s \in \{0,1\}$ *we have*

$$\mu_{F,x\to a,\ell}(s), \mu_{F,a\to x,\ell}(s), \mu_{F,x,\ell}(s) \in \{0, 1/2, 1\}. \tag{3.2.2}$$

Furthermore, for all $\ell > 2\sum_{a\in C(F)} |\partial a|$ *we have* $\mu_{F,x,\ell}(s) = \mu_{F,x,\ell+1}(s)$*. (Since the total number of messages is bounded by* $2\sum_{a\in C(F)} |\partial a|$*, the BP messages will have converged point wise after this number of iterations.)*

3.2.2 Decimation Process

In addition to the BPGD algorithm itself, the heuristic work [96] considers an idealized version of the algorithm, the *decimation process*. This is a thought experiment that highlights the conceptual reasons behind the success/failure of BPGD algorithm. Just like BPGD, it also assigns values to variables one after the other but instead of the BP 'approximations', the decimation process uses the *actual* marginals given its previous decisions. To be precise, suppose that the input formula $\boldsymbol{F}$ is satisfiable and that variables $x_1, \ldots, x_t$ have already been assigned values $\boldsymbol{\sigma}_{\mathrm{DC}}(x_1), \ldots, \boldsymbol{\sigma}_{\mathrm{DC}}(x_t)$ in the previous iterations. Obtain $\boldsymbol{F}_{\mathrm{DC},t}$ by substituting the values $\boldsymbol{\sigma}_{\mathrm{DC}}(x_1), \ldots, \boldsymbol{\sigma}_{\mathrm{DC}}(x_t)$ for $x_1, \ldots, x_t$ and dropping any clauses that do not contain any of $x_{t+1}, \ldots, x_n$. Thus, $\boldsymbol{F}_{\mathrm{DC},t}$ is a XORSAT formula with variables $x_{t+1}, \ldots, x_n$.

Let $\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}$ be a random satisfying assignment of $\boldsymbol{F}_{\mathrm{DC},t}$. Then the decimation process sets x_{t+1} according to the true marginal $\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]$, thus ultimately returning a uniformly random satisfying assignment of $\boldsymbol{F}$. The pseudocode is displayed as Algorithm 3.3.

Input: a random k-XORSAT formula $\boldsymbol{F}$, conditioned on being satisfiable
Output: an assignment $\boldsymbol{\sigma}_{\mathrm{DC}} : \{x_1, \ldots, x_n\} \to \{0,1\}$.

1 **for** $t = 0, \ldots, n-1$ **do**
2 compute $\pi_{\boldsymbol{F}_{\mathrm{DC},t}} = \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]$;
3 set $\boldsymbol{\sigma}_{\mathrm{DC}}(x_t) = \begin{cases} 1 & \text{with probability } \pi_{\boldsymbol{F}_{\mathrm{DC},t}} \\ 0 & \text{with probability } 1 - \pi_{\boldsymbol{F}_{\mathrm{DC},t}} \end{cases}$;
4 **return** $\boldsymbol{\sigma}_{\mathrm{DC}}$;

Algorithm 3.3: The decimation process (Section 1.4, [22]).

Remark 3.2.3. *If the 'BP approximations' are correct, the decimation process and BPGD are identical. The key question should come to our mind that for what parameter regimes these two processes coincide or diverge.*

Later in Chapter 6 we will see in details the connection of BPGD and decimation process and their phase transitions and the performance of BPGD by providing the exact success/failure probability regimes which verifies mathematically the heuristic work by Ricci-Tersenghi and Semerjian [96].

3.2.3 Unit Clause Propagation

In this thesis, we analyze two variants of *Unit Clause Propagation*, one for random 2-SAT [23] and another one for random k-XORSAT model [22].

Employed by all modern SAT solvers as a sub-routine, Unit Clause Propagation is a linear time algorithm that tracks the implication of the partial assignments. As we know that random 2-SAT problem is in P, the polynomial algorithm, that solves it, is a sequential assignment procedure that a variable is assigned to a given value at each time, ends either when all the variables are assigned and the resultant assignment is SAT (or when it has proven the formula UNSAT).

At each step, once a variable is assigned, the initial CNF-formula can be simplified according to a reduction process. Suppose we set a variable i to $x_i = 1$ (the case $x_i = 0$ is symmetric). Each clause containing the literal x_i is satisfied by this assignment and can be removed from the formula. On the other hand, clauses containing the opposite literal $\neg x_i$ cannot be satisfied by this assignment; thus, the literal $\neg x_i$ is removed from those clauses, reducing their length by one. As a consequence, this reduction may produce a 0-clause: for instance, if in the original formula $\boldsymbol{F}$ there was already a unit clause $c = \neg x_i$, then setting $x_i = 1$ immediately creates a contradiction, and $\boldsymbol{F}|x_i = 1$ is UNSAT. In that case, backtracking is required: one must undo the assignment $x_i = 1$ and instead try $x_i = 0$ in order to proceed.

During this process, whenever the simplified formula contains a unit clause, that variable is forced to take the unique value satisfying it. This assignment may generate new unit clauses, which in turn must also be satisfied, and so on. This cascading sequence of forced steps is known as Unit Clause Propagation (UCP). However, for worst-case 2-SAT instances, UCP alone does not guarantee success: it must be combined with backtracking to systematically explore assignments when contradictions arise.

Let's consider a 2-CNF formula $\boldsymbol{F}$ along with a set $\mathscr{L}$ of literals. These literals are deemed to be 'true'. The algorithm then pursues direct logical implications, thereby identifying additional 'implied' literals that need to be true so that no clause gets violated. This procedure is outlined in Steps 1–2 of Algorithm 3.4; the outcome of Steps 1–2 is independent of the order in which literals/clauses are processed.

Input: A 2-CNF Φ along with a set $\mathscr{L}$ of literals deemed true.

1 **while** there exists a clause $a \equiv l \vee \neg l'$ with $l' \in \mathscr{L}$ and $l \notin \mathscr{L}$ **do**
2 add literal l to $\mathscr{L}$;
3 For variables $x \in V(\Phi)$ such that $x \in \mathscr{L}$ or $\neg x \in \mathscr{L}$ let

$$\sigma_x = \begin{cases} 1 & \text{if } x \in \mathscr{L} \text{ and } \neg x \notin \mathscr{L}, \\ -1 & \text{if } \neg x \in \mathscr{L} \text{ and } x \notin \mathscr{L}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathscr{C}$ be the set of all clauses a such that $\sigma_x = 0$ for all $x \in \partial a$ and return $\mathscr{L}, \mathscr{C}, \sigma$;

Algorithm 3.4: Pessimistic Unit Clause Propagation ('PUC') (Section 2.4, [23]).

Clearly, trouble occurs if PUC ends up placing both a literal l and its negation $\neg l$ into the set $\mathscr{L}$. Our 'pessimistic' Unit Clause variant makes no attempt at mitigating such contradictions. Instead, Step 3 just constructs a partial assignment where all conflicting literals are set to a dummy value zero. In addition to this, PUC identifies the set $\mathscr{C}$ of *conflict clauses* that contain conflicted variables only.

Now consider a 2-CNF $\boldsymbol{F}$ on a set of variables $V(\boldsymbol{F})$. For each possible literal $l \in \{x, \neg x : x \in V(\boldsymbol{F})\}$ we run PUC $(\boldsymbol{F}, \mathscr{L} = \{l\})$. Let $\mathscr{C}(\boldsymbol{F}, \{l\})$ be the set of conflict clauses returned by PUC. Obtain the *pruned formula* $\widehat{\boldsymbol{F}}$ from $\boldsymbol{F}$ by removing all clauses in $\mathscr{C}(\boldsymbol{F}) = \bigcup_l \mathscr{C}(\boldsymbol{F}, \{l\})$. Then it is easy to verify the following fact:

Fact 3.2.4. *For any 2-CNF $\boldsymbol{F}$ the pruned 2-CNF $\widehat{\boldsymbol{F}}$ is satisfiable.*

Remark 3.2.5. *The pruned formula $\widehat{\boldsymbol{F}}$ could have far fewer clauses than the original formula $\boldsymbol{F}$. Accordingly, even if $\boldsymbol{F}$ is satisfiable the number $Z(\widehat{\boldsymbol{F}})$ of satisfying assignments of $\widehat{\boldsymbol{F}}$ could dramatically exceed $Z(\boldsymbol{F})$.*

As UCP returns all the assignments that were forced due to the presence of unit clauses, there are three possible output for this process:

- the output is an assignment of all the variables in V, then $\boldsymbol{F}$ is SAT which in turn applies that the assignment produced is a solution to $\boldsymbol{F}$.

- If the obtained simplified formula doesn't contain unit clauses, then the output is a partial assignment which can be extended to a complete satisfying assignment if and only if the input formula is satisfiable i.e., if and only if $\boldsymbol{F}$ is SAT.

- the output is UNSAT which in turn the input formula $\boldsymbol{F}$ is UNSAT.

Coming to the analysis of UCP on random k-XORSAT, due to the Fact 3.2.2, the BPGD algorithm effectively reduces to UCP, a purely combinatorial algorithm [66, 96]. It works exactly same as we discussed before by attempting to assign random values to as yet unassigned variables one after the other. After each such random assignment the algorithm pursues the 'obvious' implications of its decisions. Specifically, the algorithm substitutes its chosen truth values for all occurrences of the already assigned variables. If this leaves a 'unit clause', the algorithm assigns that variable so as to satisfy the unit clause. If a conflict occurs because two unit clauses impose opposing values on a variable, the algorithm declares that a conflict has occurred, sets the variable to false and continues; of course, in the event of a conflict the algorithm will ultimately fail to produce a satisfying assignment. The pseudocode for the algorithm is displayed in Algorithm 3.5.

1 Let $U = \emptyset$ and let $\boldsymbol{\sigma}_{\mathrm{UC}} : U \to \{0,1\}$ be the empty assignment;
2 **for** $t = 0, \dots, n-1$ **do**
3   **if** $x_{t+1} \notin U$ **then**
4   add x_{t+1} to U;
5   choose $\boldsymbol{\sigma}_{\mathrm{UC}}(x_{t+1}) \in \{0,1\}$ uniformly at random;
6   **while** $\boldsymbol{F}[\boldsymbol{\sigma}_{\mathrm{UC}}]$ *contains a unit clause* a **do**
7   let x be the variable in a;
8   let $s \in \{0,1\}$ be the truth value that x needs to take to satisfy a;
9   **if** *another unit clause* a' *exists that requires* x *be set to* $1-s$ **then**
10   output 'conflict' and let $\boldsymbol{\sigma}_{\mathrm{UC}}(x) = 0$;
11   **else**
12   add x to U and let $\boldsymbol{\sigma}_{\mathrm{UC}}(x) = s$;
13 **return** $\boldsymbol{\sigma}_{\mathrm{UC}}$;

Algorithm 3.5: The UCP algorithm for random k-XORSAT instance $\boldsymbol{F}$ (Section 6.1, [22]).

Let $\boldsymbol{F}_{\mathrm{UC},t}$ denote the simplified random k-XORSAT formula obtained after the first t iterations (in which the truth values chosen for $x_1, \dots, x_t$ and any values implied by unit clauses have been substituted). We notice that the values assigned during Steps 6–12 are deterministic consequences of the choices in Step 5. In particular, the order in which unit clauses are processed Steps 6–12 does not affect the output of the algorithm. Later in this thesis, we will see in details the connection/performance of UCP algorithm with the previously discussed BPGD algorithm on random k-XORSAT in Chapter 6.

3.2.4 Pure Literal Pursuit

In our third result of this thesis [21] on the lower bound of Gibbs Uniqueness threshold on random k-SAT, the main result relies on the algorithm called *Pure Literal Pursuit*('PULP'). Its purpose is to trace the repercussions of setting a relatively small number of variables to specific truth values which will allow us to compare the number of satisfying assignments that set a few chosen variables to specific values to the total number of satisfying assignments. Given a k-CNF $\boldsymbol{F}$ and a set $\mathscr{L}$ of literals of $\boldsymbol{F}$ that we deem to be set to 'true'. We would like to identify a superset $\bar{\mathscr{L}} \supseteq \mathscr{L}$ of literals ($\bar{\mathscr{L}}$ is a 'closure' of $\mathscr{L}$) with the following properties:

PULP1 every clause a that contains a literal from $\neg\bar{\mathscr{L}} = \{\neg l : l \in \bar{\mathscr{L}}\}$ also contains a literal from $\bar{\mathscr{L}}$.

PULP2 there is no literal l such that $l, \neg l \in \bar{\mathscr{L}}$.

It may be impossible to satisfy **PULP1** and **PULP2** simultaneously. In this case we ask PULP to report a 'contradiction'. But if **PULP1–PULP2** can be satisfied, we aim to find a closure $\bar{\mathscr{L}}$ of as small size $|\bar{\mathscr{L}}|$ as possible.

The combinatorial idea behind **PULP1–PULP2** is as follows. Deeming the literals from the initial set $\mathscr{L}$ 'true', our goal is to reconcile this assumption with the formula $\boldsymbol{F}$. To this end we enhance the set $\mathscr{L}$. Clearly, any clause that contains the negation $\neg l$ of a literal l that we deem true also needs to contain another literal l' that is set to true. This is what **PULP1** asks. Furthermore, it would be contradictory to deem both l and its negation $\neg l$ true; this is **PULP2**.

In order to identify a 'small' closure $\bar{\mathscr{L}}$ the PULP algorithm resorts to *pure literal elimination*, let's consider a variable x is *pure* in a CNF formula $\boldsymbol{F}$ if $\operatorname{sign}(x,a) = \operatorname{sign}(x,b)$ for any two clauses $a, b \in \partial x$. Clearly, if our objective is to construct a satisfying assignment, we might as well set all pure variables x to the value that satisfies all clauses $a \in \partial x$ and disregard these clauses henceforth. In light of this observation, pure literal elimination repeatedly removes all clauses that contain a pure variable. Naturally, every round of clause removals may create new pure variables, and thus more clauses may be ripe for removal in the next round. For a clause a of the original formula $\boldsymbol{F}$ let $\mathfrak{h}_a(\boldsymbol{F}) \geq 1$ be the number of the round at which pure literal elimination removes a. If a is never removed then we set $\mathfrak{h}_a(\boldsymbol{F}) = \infty$.

The PULP algorithm invokes a slightly modified version of pure literal elimination to accommodate the initial set $\mathscr{L}$ of literals. Specifically, for a variable x of a CNF $\boldsymbol{F}$ and $s \in \{\pm 1\}$ let $\boldsymbol{F}[x \mapsto s]$ be the CNF obtained by removing all clauses $a \in \partial x$ with $\operatorname{sign}(x,a) = s$ and removing the literal $-s \cdot x$ from all $a \in \partial x$ with $\operatorname{sign}(x,a) = -s$. The definition reflects that if we set x to value s, all $a \in \partial^s x$ will be satisfied, while all $a \in \partial^{-s} x$ will have to be satisfied by one of their other constituent literals. Further, let

$$\mathfrak{h}_x(s,\boldsymbol{F}) = \begin{cases} 0 & \text{if } \partial_{\boldsymbol{F}}^{-s} x = \emptyset, \\ \max\left\{\mathfrak{h}_a(\boldsymbol{F}[x \mapsto s]) : a \in \partial_{\boldsymbol{F}}^{-s} x\right\} & \text{otherwise.} \end{cases} \in [0,\infty]. \tag{3.2.3}$$

We refer to $\mathfrak{h}_x(s,\boldsymbol{F})$ as the *height of literal $s \cdot x$ in $\boldsymbol{F}$*.

The PULP algorithm, displayed as Algorithm 3.6, harnesses the heights as follows. In its attempt to precipitate **PULP1** and **PULP2** the algorithm iteratively enhances the set $\mathscr{L}$ of literals deemed to be 'true'. For any clause a that violates **PULP1** and that contains a literal $l \notin \neg\mathscr{L}$ the algorithm adds one such literal l *of minimum height* to $\mathscr{L}$. This choice is intended to keep the ultimate size of the closure small; one could say that PULP uses height as a proxy of 'size'. If at any point the algorithm encounters a clause a that consists of literals from $\neg\mathscr{L}$ only, the algorithm reports a contradiction and aborts.

Remark 3.2.6. *To break ties that may occur in the execution of Steps 3 and 7 of PULP we assume that the variables and clauses of $\boldsymbol{F}$ are numbered so that Steps 3 and 7 can choose the clause/variable with the smallest number that satisfies the respective requirements. In due course we will run PULP on (finite*

Input: A k-CNF $\boldsymbol{F}$ and a set $\mathscr{L}$ of literals.

Let $\bar{\mathscr{L}} = \mathscr{L}$;
while *there is a clause a that contains a literal from $\neg\bar{\mathscr{L}}$ but no literal from $\bar{\mathscr{L}}$* **do**
  Pick such a clause a that minimizes the distance from the initial set $\mathscr{L} = \{|l| : l \in \mathscr{L}\}$;
  if *a consists of literals $l \in \neg\bar{\mathscr{L}}$ only* **then**
    return 'contradiction' and halt;
  else
    choose $x \in \partial a$ with $x, \neg x \notin \bar{\mathscr{L}}$ that minimizes $\mathfrak{h}_x(\mathrm{sign}(x,a), \boldsymbol{F})$ and add $\mathrm{sign}(x,a)\cdot x$ to $\bar{\mathscr{L}}$
return $\bar{\mathscr{L}}$

Algorithm 3.6: The PULP algorithm (Section 2.3, [21])

subtrees of) the Galton-Watson tree $\mathbb{T}$. To number the variables and clauses of $\mathbb{T}$ we equip each of them with an independent Gaussian label. Since $\mathbb{T}$ comprises a countable number of clauses/variables, these labels will almost surely be pairwise distinct.

The analysis of this algorithm can be found in Chapter 7 briefly.

3.3 Warning Propagation

Warning Propagation is a purely combinatorial message passing algorithm in the same family like Belief Propagation. For some graph based matrix and constraint satisfaction models the 'discrete' version (where the messages are from the finite alphabet Ω instead of being probability distributions) of Belief Propagation is treated as Warning Propagation which helps to find direct implications of a recursive processes associated with graphs. So, for a graph $\mathbb{G}$ let $\mathscr{M}(G)$ be the set of all vectors $(\omega_{u\to v})_{(u,v)\in V(\mathbb{G})^2:\{u,v\}\in E(\mathbb{G})} \in \Omega^{2|E(\mathbb{G})|}$. Here also the parallelism holds for updating messages based on some fixed rules like in BP. The update rule φ is defined for $d \in \mathbb{N}$ and $\binom{\Omega}{d}$, the set of all d-ary multisets with elements from Ω:

$$\varphi : \bigcup_{d\geq 0} \binom{\Omega}{d} \to \Omega \tag{3.3.1}$$

which takes any multiset of input messages and produces an output messages. Now the corresponding Warning Propagation operator is defined as,

$$\begin{aligned}
&\mathrm{WP}_{\mathbb{G}} : \mathscr{M}(\mathbb{G}) \to \mathscr{M}(\mathbb{G})\\
&\omega = (\omega_{u\to v})_{uv} \to \left(\varphi\left(\{\omega_{u\to v} : uv \in E(\mathbb{G}), u \neq v\}\right)\right)_{uv}
\end{aligned}$$

The message from node u to node v is updated according to the WP updated rule applied to the multiset of messages that u receives from all of its neighbors except v. In factor graph setting, similar as BP, WP also associates two message sequences $(\omega_{F,x\to a}, \omega_{F,a\to x})$ with every adjacent clause/variable pair. In this thesis we provide a detailed analysis of Warning Propagation when we analyze the performance of *Belief Propagation Guided Decimation*('BPGD') on random k-XORSAT model (more details can be found in Chapter 6).

Due to the half-integrality of the BP messages and marginals in random k-XORSAT, BP is equivalent to Warning Propagation. The messages take one of three possible discrete values $\{\mathtt{f},\mathtt{u},\mathtt{n}\}$ ('frozen', 'uniform', 'null'). To trace the BP messages for which the two values $\{\mathtt{n},\mathtt{u}\}$ would be necessary. However, the third value $\mathtt{f}$ will prove useful in order to compare the BP approximations (computed using 'BPGD' algorithm discussed in Section 3.2.1) with the actual marginals(computed using 'Decimation' process discussed in Section 3.2.2). Although the messages initially in BP are given as uniform i.e.,

$$\omega_{F,x\to a,0}(s) = \omega_{F,a\to x,0}(s) = 1/2 \qquad\qquad (s \in \{0,1\}).$$

we launch WP from all frozen start values.

$$\omega_{F,x\to a,0} = \omega_{F,a\to x,0} = \mathtt{f} \qquad\qquad \text{for all } a, x. \tag{3.3.2}$$

Subsequently the messages get updated according to the rules:

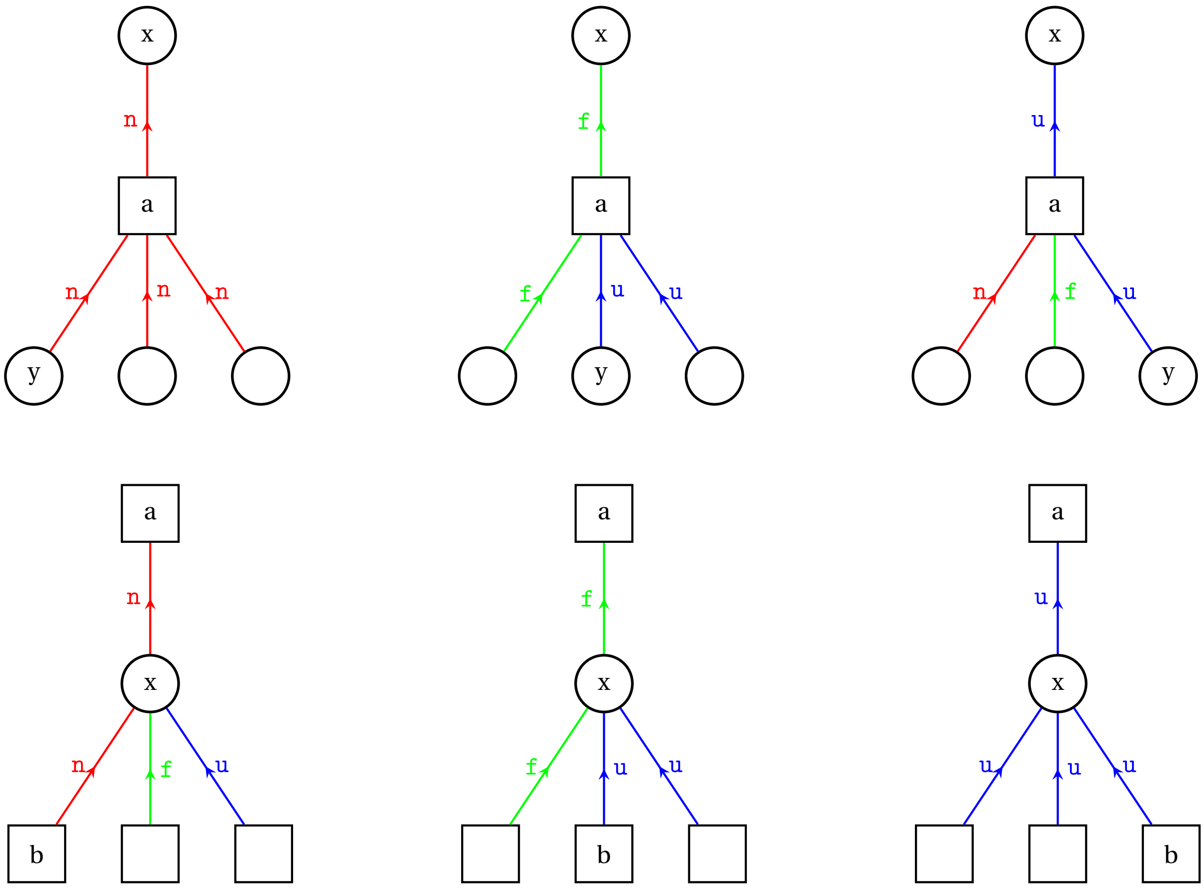


Figure 3.2: **Up**: A local snapshot of Warning Propagation update rules for message $\nu_{F,a\to x,\ell}$ defined in (3.3.3). **Down**: Similarly, a local snapshot of Warning Propagation update rules for message $\nu_{F,x\to a,\ell}$ defined in (3.3.4)

$$\omega_{F,a\to x,\ell+1} = \begin{cases} \mathtt{n} & \text{if } \omega_{F,y\to a,\ell} = \mathtt{n} \text{ for all } y \in \partial a \setminus \{x\}, \\ \mathtt{f} & \text{if } \omega_{F,y\to a,\ell} \neq \mathtt{u} \text{ for all } y \in \partial a \setminus \{x\} \text{ and } \omega_{F,y\to a,\ell} \neq \mathtt{n} \text{ for at least one } y \in \partial a \setminus \{x\}, \\ \mathtt{u} & \text{otherwise,} \end{cases} \tag{3.3.3}$$

$$\omega_{F,x\to a,\ell+1} = \begin{cases} \mathtt{n} & \text{if } \omega_{F,b\to x,\ell} = \mathtt{n} \text{ for at least one } b \in \partial x \setminus \{a\}, \\ \mathtt{f} & \text{if } \omega_{F,b\to x,\ell} \neq \mathtt{n} \text{ for all } b \in \partial x \setminus \{a\} \text{ and } \omega_{F,b\to x,\ell} = \mathtt{f} \text{ for at least one } b \in \partial x \setminus \{a\}, \\ \mathtt{u} & \text{otherwise.} \end{cases} \tag{3.3.4}$$

In addition to the messages we also define the *mark* of variable node x by letting

$$\omega_{F,x,\ell} = \begin{cases} \mathtt{n} & \text{if } \omega_{F,b\to x,\ell} = \mathtt{n} \text{ for at least one } b \in \partial x, \\ \mathtt{f} & \text{if } \omega_{F,b\to x,\ell} \neq \mathtt{n} \text{ for all } b \in \partial x \text{ and } \omega_{F,b\to x,\ell} = \mathtt{f} \text{ for at least one } b \in \partial x, \\ \mathtt{u} & \text{otherwise.} \end{cases} \tag{3.3.5}$$

We conclude the chapter by establishing a relationship between BP and WP.

Fact 3.3.1. *For all $t \geq 0$ and all x, a we have*

$$\nu_{x\to a,\ell}(1) = 1/2 \quad \Leftrightarrow \quad \omega_{F,x\to a,\ell} \neq \mathtt{n}, \tag{3.3.6}$$

$$\nu_{a\to x,\ell}(1) = 1/2 \quad \Leftrightarrow \quad \omega_{F,a\to x,\ell} \neq \mathtt{n}, \tag{3.3.7}$$

$$\nu_{x,\ell}(1) = 1/2 \quad \Leftrightarrow \quad \omega_{F,x,\ell} \neq \mathtt{n}. \tag{3.3.8}$$

In the next chapter we will have a detailed overview of the phase transitions of different random satisfiability problems.

4

Phase Transitions in random CSPs

"Recent research indicates that many convex optimization problems with random constraints exhibit a phase transition as the number of constraints increases.."

–Dennis Amelunxen et al.

In this chapter we provide a detailed analysis on the phase transitions occurring in random constraint satisfaction problems when the clause density α is varied. Particularly, we will take as an example of mainly two random CSP models, one is random k-SAT and another is random k-XORSAT which are the main theme of this thesis, but many other random CSP ensembles share the same qualitative nature.

4.1 The Satisfiabilty Transition

Recall that the satisfiability threshold $\alpha_{sat}(k)$ separates a phase $\alpha < \alpha_{sat}(k)$ where the random instances are SAT w.h.p. to a phase where random instances are UNSAT w.h.p. One of the most well-known technique for an estimation of the satisfiability threshold discussed in [15, 59, 65, 72, 79] is the *cavity method*. However, the existence of the satisfiability transition is not yet proven for all k values. The below conjecture summarizes the fact.

Conjecture 4.1.1. *Let $\boldsymbol{F} = \boldsymbol{F}(k,\alpha)$ be a random CNF formula with n variables and $m = \alpha n$ clauses with α is the clause density, drawn from the random k-SAT ensemble. Then for any $k \geq 2$ there exists a* *constant $\alpha_{sat}(k)$ such that for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}[\boldsymbol{F}(k, \alpha_{sat}(k) - \varepsilon) \text{ is } SAT] = 1 \qquad \lim_{n\to\infty} \mathbb{P}[\boldsymbol{F}(k, \alpha_{sat}(k) + \varepsilon) \text{ is } SAT] = 0$$

This conjecture has been proved for $k = 2$ with $\alpha_{sat} = 1$ by Chvatal and Reed in 1992 [26] and Goerdt in 1996 [53]. Later in 2015, Ding, Sly and Sun prove the satisfiability conjecture for large but finite k value and show that the value of $\alpha_{sat}(k)$ is given by the one step symmetry breaking cavity method prediction. However, Friedgut in 1999 [47] provides a partial result in this regards:

Theorem 4.1.2. *For every $k \geq 2$, there exists a sequence $\alpha_k(n)$ such that for all $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}[\boldsymbol{F}(k, \alpha_k(n) - \varepsilon) \text{is } SAT] = 1 \qquad \lim_{n\to\infty} \mathbb{P}[\boldsymbol{F}(k, \alpha_k(n) + \varepsilon) \text{is } SAT] = 0$$

The above theorem provides the transition from SAT to UNSAT which takes place in a window smaller than any fixed ε for large enough n. However, it remains to prove the convergence of the sequence $\alpha_k(n)$ to some value $\alpha_{sat}(k)$ as $n \to \infty$ to prevent from possible oscillations. There are several methods to derive the upper and the lower bounds rigorously on this sequence $\alpha_k(n)$. The most well-known choice for deriving the upper bound on the number of satisfying assignments is to apply the Markov's inequality (or, the first moment bound) whereas for obtaining the lower bound on the number of solutions, Chebyshev's inequality (or, second moment method) which is more delicate to implement than the first moment bound.

More precisely, define a function on the set of instances $\mathscr{U}(\boldsymbol{F})$ such that,

$$\mathscr{U}(\boldsymbol{F}) = \begin{cases} 0, & \text{if } \boldsymbol{F} \text{ is UNSAT} \\ \geq 1, & \text{otherwise} \end{cases}$$

Then after applying Markov's inequality one can get,

$$\mathbb{P}[\boldsymbol{F} \text{ is SAT}] \leq \mathbb{E}[\mathscr{U}(\boldsymbol{F})]$$

As we don't know how to compute the quantity $\mathscr{U}(\boldsymbol{F}) = \mathbb{1}[\boldsymbol{F} \text{ is SAT}]$, the first choice will be to use $\mathscr{U}(\boldsymbol{F}) = Z(\boldsymbol{F})$, the number of solutions of $\boldsymbol{F}$. Then given an assignment σ, by linearity of expectation and by uniformity in the clause generation $\mathbb{E}[Z(\boldsymbol{F})]$ is given by,

$$\mathbb{E}[Z(\boldsymbol{F})] = 2^n(1-2^{-k})^m = \exp[n(\log 2 + \alpha \log(1-2^{-k}))]$$

When $n \to \infty$ we get,

$$\mathbb{E}[Z(\boldsymbol{F})] \to \begin{cases} 0, & \text{if } \alpha > \alpha_u(k), \\ +\infty, & \text{if } \alpha < \alpha_u(k) \end{cases}$$

with $\alpha_u(k) = -\frac{\log 2}{\log(1-2^{-k})}$. Because of the number of satisfying assignments $Z(\boldsymbol{F})$ can take exponentially large values as $n \to \infty$ and its fluctuations can be exponentially large as well, one can expect that the above upper bound α_u is not tight. Later in 1998, Kirousis et.al. [57] define the function $\mathscr{U}(\boldsymbol{F})$ as the number of locally maximal SAT assignment which counts the number of solutions in a small subclass of solutions. Thus they obtain a new more tighter upper bound $\widehat{\alpha_u}(k)$ that is the solution of the equation

$$\alpha \log(1-2^k) + \log(2 - \exp(-k\alpha/(2^k-1))) = 0$$

Coming to the lower bound on the value $\alpha_k(n)$, in 2006 Achlioptas and Moore [4] first introduced the method for obtaining the lower bound using second moment method. Applying the second moment to the function $\mathscr{U}(\boldsymbol{F})$ we get,

$$\mathbb{P}[\boldsymbol{F} \text{ is SAT}] \geq \frac{\mathbb{E}[\mathscr{U}(\boldsymbol{F})]^2}{\mathbb{E}[\mathscr{U}(\boldsymbol{F})^2]}$$

It can be shown that applying this method to the number of solutions $Z(\boldsymbol{F})$ doesn't provide a useful bound because the fraction $\frac{\mathbb{E}[Z(\boldsymbol{F})]^2}{\mathbb{E}[Z(\boldsymbol{F})^2]}$ disappears for any non-zero value of α. Instead one can choose another function $\mathscr{U}$ to be the size of the subset of the set of solutions. Using this, Achlioptas and Moore in [4] showed that for any $k \geq 3$,

$$\lim_{n\to\infty} \mathbb{P}[\boldsymbol{F}(k,\alpha) \text{ is SAT}] = 1, \quad \text{if } \alpha \leq 2^{k-1}\log 2 - 2$$

For $n \to \infty$ this lower bound along with the upper bound provides the scaling $\alpha_{sat}(k) = O(2^k)$. In the next section we will discuss briefly 'quenched' and 'annealed' techniques which is very useful in the analysis of satisfiability threshold on random CSP problems.

4.2 Quenched and Annealed Techniques

Let us consider a graphical model $\mathscr{G}$ and its corresponding measure $\mu_{\mathscr{G}(\boldsymbol{F})}$. The support of the model is the set of solutions of a given instance $\boldsymbol{F}$:

$$\mu_{\mathscr{G}(\boldsymbol{F})}(\sigma) = \begin{cases} 0, & \text{if } \sigma \text{ is not a solution} \\ >0, & \text{otherwise} \end{cases}$$

When σ is not a solution i.e., $\mu_{\mathscr{G}(\boldsymbol{F})}(\sigma) = 0$ we can introduce the parameter inverse temperature (β) and bringing the normalized free entropy (defined in (2.2.7)) as follows:

$$\Phi_{\mathscr{G}} = \Phi\left(\mathscr{G}(\boldsymbol{F}), \beta\right) = \frac{1}{n}\log Z(\mathscr{G}(\boldsymbol{F}), \beta) \tag{4.2.1}$$

Now when the instance is randomly drawn (for instance from the random k-SAT ensemble), the measure $\mu_{\mathscr{G}(\boldsymbol{F})}$ becomes random. Determining the typical properties of measure $\mu_{\mathscr{G}(\boldsymbol{F})}$ and the normalized free entropy density $\Phi_{\mathscr{G}}$ are two most important quantities for estimating the partition function which counts the number of solutions in any random satisfiability problems. For this the quenched free entropy density is defined over the random ensemble of instances:

$$\Phi^{\text{que}}(\mathscr{G}) = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\log Z(\mathscr{G}(\boldsymbol{F}))] \tag{4.2.2}$$

Usually, this quantity cannot be evaluated exactly on random satisfiability problems, but can be estimated using non-rigorous *cavity method* which we will discuss later in this chapter. There is a natural upper bound on this quantity provided by normalized annealed free entropy density,

$$\Phi^{\text{ann}}(\mathscr{G}) = \lim_{n\to\infty} \frac{1}{n}\log\mathbb{E}[Z(\mathscr{G}(\boldsymbol{F}))] \tag{4.2.3}$$

It is straightforward from Jensen's inequality applied on the number of solutions $Z(\mathscr{G}(\boldsymbol{F}))$ yields:

$$\Phi^{\text{ann}}(\mathscr{G}) \geq \Phi^{\text{que}}(\mathscr{G}) \tag{4.2.4}$$

Later in this chapter, when we will talk about the replica symmetry trick, we will again introduce the 'quenched free energy' quantity. In the second result of this thesis, we harness 'quenched' arguments which was partly developed in some prior work [12,29] on the rank of random matrices over finite fields to establish a precise connection between the decimation process and the performance of BPGD [22].

4.3 Gibbs measure and Long range correlation

The effectiveness of belief propagation relies on a basic assumption that the adjacent variable nodes becomes weakly correlated with respect to the resulting distribution when a check node is pruned from the factor graph. But when the factor graph contains small loops or variables are correlated at a long distance then the above hypothesis may break down [66]. So, in factor graphs with locally tree like structure, the long range correlation is responsible for the failure of BP. Thus a phase transition will occur with the emergence of such long range correlation separating 'weakly correlated' and 'highly correlated' phase. The central tool behind the study of any random CSPs such as random k-SAT and random k-XORSAT through the lens of statistical mechanics is '*Gibbs measure*' which encodes the uniform distributions over all set of solutions. Thus it provides insight into the correlation decay, algorithmic tractability and various phase transitions [38]. This section develops the framework of Gibbs measure in sparse random structures [43] by introducing correlation decay and Gibbs uniqueness property followed by the landscape of phase transition – *replica symmetry, one-step replica symmetry breaking* (1-RSB cavity method) and the *(non)-reconstruction* / *reconstruction* properties on trees.

4.3.1 Gibbs measure on random CSPs

Recall the Boltzmann(Gibbs) distribution from Chapter 2, with $\boldsymbol{F}$ be an instance of a random CSP consists of variable set $V = \{x_1, x_2, \cdots, x_n\}$ and constraint set $C = \{c_1, c_2, \cdots, c_m\}$. A satisfying assignment

$\sigma \in \{0,1\}^n$ is a solution if all the constraints are satisfied. The associated Gibbs measure at inverse temperature $\beta \geq 0$ is defined as,

$$\mu_\beta(\sigma) = \frac{1}{Z_\beta(\boldsymbol{F})} \exp\big(-\beta H(\sigma)\big)$$

where $H(\sigma)$ is number of violated constraints under the assignment σ and the partition function $Z_\beta(\boldsymbol{F})$ is given by,

$$Z_\beta(\boldsymbol{F}) = \sum_{\tau \in \{0,1\}^n} \exp\big(-\beta H(\tau)\big)$$

Remark 4.3.1.

- *At $\beta = \infty$, the Gibbs measure $\mu_\beta(\boldsymbol{F})$ is the uniform distribution over all satisfying assignments.*
- *At finite β, $\mu_\beta(\boldsymbol{F})$ interpolates between uniform randomness and strong bias toward solutions.*

Thus the Gibbs measure encodes the solution space geometry and serves as a basis for analyzing the correlations between variables.

4.3.2 Correlation decay and Gibbs Uniqueness

In statistical mechanics, physical systems that have only short range correlation should relax rapidly to their equilibrium distribution. The reason behind this depends on the different degrees of freedom. If the degrees of freedom are independent, then the system relaxes on microscopic scales (namely the relaxation time of a single particle, spin etc.), whereas if they are not independent but their correlations are short-ranged, they can be harsh in such a way that they become nearly independent. For two variables x_i and $x_j \in V$, define their correlation under the Gibbs measure [38, 40] μ_β by

$$\mathrm{Corr}_{\mu_\beta}(x_i, x_j) = \mathbb{E}_{\mu_\beta}[x_i x_j] - \mathbb{E}_{\mu_\beta}[x_i]\mathbb{E}_{\mu_\beta}[x_j]$$

So we need to use a measure of how much the joint distribution $\mu_{ij}(\cdot,\cdot)$ of x_i and x_j is different from their product marginals $\mu_i(\cdot)$ times $\mu_j(\cdot)$. Thus defining the two-point correlation [66] by averaging their variation distance $||\mu_{ij}(\cdot,\cdot) - \mu_i(\cdot)\mu_j(\cdot)||$ over the vertices i, j:

$$\mathrm{Corr}^{(2)} \equiv \frac{1}{n} \sum_{i,j \in V} ||\mu_{ij}(\cdot,\cdot) - \mu_i(\cdot)\mu_j(\cdot)||$$

Remark 4.3.2.

- *Correlation decay occurs if correlations $Corr_{\mu_\beta}(x_i, x_j)$ vanish with the distance between two nodes $i, j \in V$ in the graph, i.e.,*

$$Corr_{\mu_\beta}(x_i, x_j) \to 0 \quad as \quad dist(i,j) \to \infty$$

- *Absence of correlation decay indicates long-range dependencies which often indicates the clustering or broken symmetry.*

These phenomena are linked to the uniqueness of Gibbs measure on the infinite tree limit (local weak convergence) [21, 40].

Gibbs Uniqueness

Begin with the Galton-Watson tree $\mathbb{T} = \mathbb{T}_{d,k}$, which is generated by a two-type branching process. The two types are *variable nodes* and *clause nodes*. The process starts with a single root variable node r. The offspring of any variable node is a $\mathrm{Po}(d)$ number of clause nodes, while every clause node begets precisely $k-1$ variable nodes. Additionally, independently for each clause node a and every variable node x that is either a child or the parent of a a *sign*, denoted $\mathrm{sign}(x,a) \in \{\pm 1\}$, is chosen uniformly at random. The resulting random tree $\mathbb{T}$ models the local structure of the random k-CNF formula $\boldsymbol{F} = \boldsymbol{F}(n,d,k)$ in the sense of local weak convergence [9, 62].

For an integer $\ell \geq 0$ let $\mathbb{T}^{(\ell)}$ be the finite tree obtained by removing all variable and clause nodes at a distance greater than 2ℓ from the root r. We identify the finite tree $\mathbb{T}^{(\ell)}$ with a Boolean formula whose variables/clauses are precisely the variable/clause nodes of $\mathbb{T}^{(\ell)}$. Let $S(\mathbb{T}^{(\ell)}) \neq \emptyset$ be the set of satisfying assignments of this formula and let $\boldsymbol{\tau}^{(\ell)} \in S(\mathbb{T}^{(\ell)})$ be a uniformly random satisfying assignment. Moreover, let $\partial^{2\ell} r$ be the set of variable nodes of $\mathbb{T}^{(\ell)}$ at distance precisely 2ℓ from the root r. Then for given d, k the tree $\mathbb{T} = \mathbb{T}_{d,k}$ has the *Gibbs uniqueness property* (see [59] for more details) if

$$\lim_{\ell\to\infty} \mathbb{E}\left[\max_{\tau\in S(\mathbb{T}^{(\ell)})} \left|\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(r)=1 \mid \mathbb{T}\right] - \mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(r)=1 \mid \mathbb{T}, \forall x\in\partial^{2\ell} r : \boldsymbol{\tau}^{(\ell)}(x) = \tau(x)\right]\right|\right] = 0. \tag{4.3.1}$$

In words, in the limit of large ℓ the truth value $\boldsymbol{\tau}^{(\ell)}(r)$ of the root r is asymptotically independent of the truth values $\{\boldsymbol{\tau}^{(\ell)}(x)\}_{x\in\partial^{2\ell} r}$ of the variables at distance 2ℓ from r. In this thesis, we explicitly derive the lower bound on the $d_{\mathrm{uniq}}(k)$ threshold for any $k \geq 3$. The details of this is discussed in Chapter 7.

Remark 4.3.3.

- ***Uniqueness regime***: *correlation decay along the tree $\mathbb{T}$ and the belief propagation converges to a unique fixed point.*

- ***Non-Uniqueness regime***: *In the infinite tree $\mathbb{T}$, suitably chosen boundary conditions can give rise to distinct extremal Gibbs measures, commonly referred to as* pure states. *On the other hand, in large finite random graphs, this phenomenon appears in the thermodynamic limit as a decomposition of the solution space into multiple well-separated clusters, which correspond to these pure states.*

- *In random k-SAT, the breakdown of Gibbs uniqueness (a state where the solution is not unique and has many distinct clauses) occurs well before the clustering threshold (see Figure 4.2) whereas in random k-XORSAT the phase transition for the appearance of a linear number of frozen variables*

(variables whose assignments are determined by the problem structure) occurs simultaneously with the emergence of the 2-core.

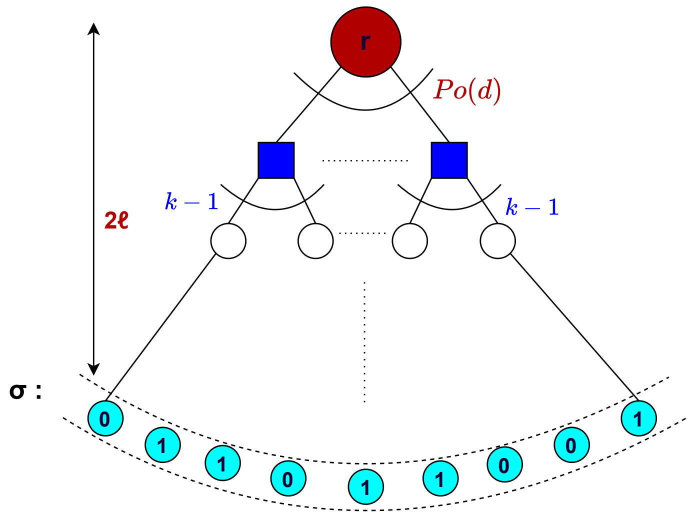


Figure 4.1: Galton-Watson tree $\mathbb{T}$ with Gibbs Uniqueness property

4.3.3 Replica Symmetry

The replica symmetry condition originally come from the method to study the partition function Z. As estimating the partition function is hard because of the sum runs over the exponential number of indices (say for k-SAT we have 2^n possible summands). Therefore to get the leading exponential order of partition function, a reasonable approximation is given by the term $\lim_{n\to\infty} \frac{1}{n} \log Z$. Then the average log-partition function (in other words 'quenched free energy') defined in Section 4.2 for computing the moments $\mathbb{E}[Z^\ell]$ for any fixed $\ell \in \mathbb{N}$:

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}[\log Z] = \lim_{n\to\infty} \lim_{\ell\to 0} \frac{1}{n\ell} \log \mathbb{E}[Z^\ell]$$

This technique is known as *replica symmetry trick* [36, 70, 71]. To see the working mechanism the most suitable model to consider is *random energy model*(REM). Although this model does not describe any realistic physical system but is a good example for studying the concept of replica theory. Consider the model, we have 2^n possible assignments σ with each has an energy $E(\sigma)$. Therefore ℓ-th moment of Z is given by,

$$Z^\ell = \sum_{i_1, i_2, \cdots, i_\ell}^{2^n} \exp\left[\beta\left(-E_{i_1} - \cdots - E_{i_\ell}\right)\right]$$

where, β denotes the inverse temperature. This quantity can be considered as a partition function of a new system given by ℓ-tuples $\{i_1, \cdots, i_\ell\}$ with energies $E_{i_1,\cdots,i_\ell} = E_{i_1} + \cdots + E_{i_\ell}$. which implies that the new system is obtained by taking ℓ independent copies of the original model with each copy refers as '*replica*'. Then the above equation can be written as,

$$Z^{\ell} = \sum_{i_1,\cdots,i_\ell=1}^{2^n} \prod_{j=1}^{2^n} \exp\left[-\beta E_j \left(\sum_{a=1}^{\ell} \mathbb{1}(i_a = j)\right)\right]$$

Therefore for obtaining the average by taking the linearity of expectation and the i.i.d Gaussian energies E_j, we get,

$$\mathbb{E}\left[Z^{\ell}\right] = \sum_{i_1,\cdots,i_\ell=1}^{2^n} \exp\left[\frac{\beta^2 n}{4} \sum_{a,b=1}^{\ell} \mathbb{1}(i_a = i_b)\right] \tag{4.3.2}$$

Now we can write the indicator in the equation (4.3.2) using $\ell \times \ell$ *overlap matrix* Q where each entry is given by $Q_{ab} = \mathbb{1}\{i_a = i_b\} \in \{0,1\}$. Then we can rewrite the above equation as below:

$$\mathbb{E}\left[Z^{\ell}\right] = \sum_{Q} \mathcal{N}_n(Q) \exp\left[\frac{\beta^2 n}{4} \sum_{a,b=1}^{\ell} Q_{ab}\right]$$

where $\mathcal{N}_n(Q)$ is the number of configurations $(i_1,\cdots,i_\ell)$ whose overlap matrix is Q_{ab} and the sum over $\sum_Q$ runs over the symmetric matrix $\{0,1\}^{\ell\times\ell}$ matrices. Using the large deviation principle for an entropy function $s(Q)$ which only depends on matrix Q of the REM model we obtain $\mathcal{N}_n(Q) = \exp\left(n(s(Q)+o(1))\right)$. Then after taking the log to the above equation we get,

$$\log\mathbb{E}\left[Z^{\ell}\right] = n(\max_Q \gamma(Q) + o(1)) \qquad \gamma(Q) = \frac{\beta^2}{4} \sum_{a,b=1}^{n} Q_{ab} + s(Q) \tag{4.3.3}$$

Here, γ is symmetric under permutation of replicas for $\ell > 1$ For any permutation π on set $|\ell|$ we have $\gamma(Q) = \gamma(Q^{\pi})$ with elements $Q^{\pi}_{ab} = Q_{\pi(a)\pi(b)}$. This γ is symmetric due to the fact that the replicas at the beginning are identical and that implies $Q_{ab} = q_0 \in \{0,1\}$ for all $a \neq b$ — this is called in physics ***replica symmetry***(RS). Now the immediate consequence comes for the maximization over γ with the matrix Q yields two maxima: one is Q_1 where all entries are one and second one is the identity matrix Q_0. From [66] there exist a precise threshold $\beta_c = 2\sqrt{(\log 2)/\ell}$, with $\beta \le \beta_c$, the global maximum obtained at Q_0 (identity matrix) whereas with $\beta > \beta_c$, the maximum attained at Q_1 (all 1-matrix). Thus heuristically, by putting the threshold value β_c for the identity matrix Q_0 we obtain the following prediction for $\beta \le \beta_c$:

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[Z] &= \lim_{n\to\infty}\lim_{\ell\to 0} \frac{1}{\ell n} \log\mathbb{E}\left[Z^{\ell}\right] \\ &= \lim_{\ell\to 0} \frac{1}{\ell}\gamma(Q_0) = \frac{\beta^2}{4} - \log 2\end{aligned}$$

Moreover, two problems come into picture for $\ell < 1$: first one we need to maximize over the negative number of variables. For the remedy of this problem Giorgio Parisi transformed the problem into a minimization problem referred as *'Parisi axioms'*. Mathematically,

$$\log \mathbb{E}\left[Z^{\ell}\right] = n\left(\min_{Q} \gamma(Q) + o(1)\right)$$

For $\beta > \beta_c$ the sum in (4.3.3) is dominated by Q which are not replica symmetric. In order to improve on the RS result, one can the subspace of matrices to be optimized over. Proposed by Parisi in the more complicated case of spin glass mean-field theory the *replica symmetry breaking*, prescribes a recursive procedure for defining larger and larger spaces of matrices Q where one searches for saddle points. The first step of this procedure called ***one step replica symmetry breaking***(1RSB) [66]. For better understanding let us suppose ℓ is a multiple of x and we group ℓ replicas into ℓ/x groups of x in such as way with $q_0 \neq q_1$,

$$\begin{cases} Q_{aa} & = 1, \\ Q_{ab} & = q_1, \quad \text{if } a \text{ and } b \text{ are in same group} \\ Q_{ab} & = q_0, \quad \text{if } a \text{ and } b \text{ are in different group} \end{cases}$$

In random k-SAT model, for any variable x_1 and x_2 and a sample assignment σ from the Boltzmann distribution we have,

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\mu(\sigma_{x_1} = \sigma_{x_2} = 1) - \mu(\sigma_{x_1} = 1)\cdot\mu(\sigma_{x_2} = 1)\right|\right] = 0$$

If the replica symmetry condition approximately holds for random factor graphs, then the limiting normalized log-partition function is predicted by,

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\log Z\right] = \sup_{\pi\in\mathscr{P}^2(\Omega)} \mathscr{B}(\pi) \tag{4.3.4}$$

where, π is the probability measure defined in the interval $[0,1]$ and $\mathscr{P}(\Omega)$ is the probability distribution on Ω and $\mathscr{B}:\mathscr{P}^2(\Omega)\to\mathbb{R}$. In physics paradigm the quantity in the r.h.s. of (4.3.4) is referred as *'Bethe free entropy'*. Our third result of this thesis on random k-SAT provide a detailed evaluation of this quantity. For more details refer to Chapter 7 and [21].

In the next subsection we will see the clustering transition upto which this replica symmetry holds for random k-SAT and beyond which this 1RSB holds that the cluster breaks into exponentially many solution clusters.

4.3.4 Clustering transition: Reconstruction Property

when we are in the satisfiable regime $\alpha < \alpha_{\text{SAT}}(k)$, the clustering transition (also known as reconstruction or, dynamic transition) occurs. The set of typical solutions is rather well connected, that means any solution can be reached from the other by intermediate solutions. But above the clustering threshold α_{clus} the typical solutions break into an exponentially many clusters or pure states which are internally well connected, but separated one from each other by free-energy barriers. So, the definition concerning typical solutions with respect to the measure $\mu = \mu_{\mathscr{G}(\boldsymbol{F})}$ chosen to describe the set of solutions. Moreover introduced by Montanari and Semerjian in [82] the clustering transition can be interpreted as the birth of the point-to-set correlation function under that probability measure $\mu_{\mathscr{G}(\boldsymbol{F})}$ chosen to describe the set of solutions. Given a variable node v and a set of variables $\mathscr{V}$, the point-to-set correlation function is defined for spin variables over the measure μ:

$$\text{Cor}(v,\mathscr{V}) = \sum_{\sigma_{\mathscr{V}}} \mathbb{P}_{\mu}(\sigma_{\mathscr{V}}) \left(\sum_{\sigma_v} \mathbb{P}_{\mu}(\sigma_v|\sigma_{\mathscr{V}})\,\sigma_v \right)^2 - \left(\sum_{\sigma_i} \mathbb{P}_{\mu}(\sigma_v)\,\sigma_v \right)$$

In the unclustered regime when the distance from variable v to $\mathscr{V}$ on the graph of interaction grows, the point-to-set interaction function $\text{Cor}(v,\mathscr{V})$ vanishes, whereas in the clustered regime it doesn't decay to zero. Further in [82], the author justifies the terminology dynamic transition by showing that this correlation implies the divergence of the relaxation time of local stochastic processes that obey the detailed balanced condition such as '*Markov Chain Monte Carlo*'. In the cavity method pf random CSPs, this clustering threshold refers to as the appearance of a non-trivial solution of one-step of Replica Symmetry Breaking (1RSB) with parisi parameter $\mathscr{X} = 1$.

Let us briefly summarize the above thing. In the unclustered phase, the typical solution belongs to a single cluster. Thus the thermodynamic properties of the measure μ are well defined by the rigorous technique called Replica Symmetric cavity method which in particular estimate Φ^{RS} that is the quenched free entropy density defined in (4.2.2). On the other hand, in the case of 1-RSB cavity method, the solution set splits into an exponential number of disjoint clusters (also called '*pure states*'). In the next subsection we will see different phases of these transitions in random k-SAT and k-XORSAT model.

In the second result of this thesis, we employ the concept of (non-) reconstruction property in the context of computing the marginal probability of the root of a bipartite graph $G(\boldsymbol{F}_{\text{DC},t})$ associated with a random k-XORSAT formula $\boldsymbol{F}_{\text{DC},t}$ generated by the decimation process (defined in Chapter 3). Roughly speaking, non-reconstruction means that the marginal $\pi_{\boldsymbol{F}_{\text{DC},t}} = \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\text{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\text{DC},t}\right]$ is determined by short-range rather than long-range effects. As we know that Belief Propagation is a local algorithm, one might expect that the (non-)reconstruction phase transition coincides with the threshold up to which `BPGD` succeeds (more details can be found in [22]).

For a (variable or clause) vertex v of $G(\boldsymbol{F}_{\text{DC},t})$ let ∂v be the set of vs neighbors. More generally, for an integer $\ell \geq 1$ let $\partial^{\ell} v$ be the set of vertices of $G(\boldsymbol{F}_{\text{DC},t})$ at shortest path distance precisely ℓ from v. From Figure 4.1 also one can refer that computing the marginal of the root r, whatever assignments at a distance greater than 2ℓ will not affect the marginal computation of root in the case of (non)-reconstruction phase. Following [59], we say that $\boldsymbol{F}_{\text{DC},t}$ has the *non-reconstruction property* if

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{E}\left[\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\text{DC},t}}(x_{t+1}) = 1 \,\middle|\, \boldsymbol{F}_{\text{DC},t}, \left\{\boldsymbol{\sigma}_{\boldsymbol{F}_{\text{DC},t}}(y)\right\}_{y\in\partial^{2\ell} x_{t+1}}\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\text{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\text{DC},t}\right]\right| \mid \boldsymbol{F}_{\text{XOR}} \text{ sat.}\right] = 0. \tag{4.3.5}$$

Conversely, $\boldsymbol{F}_{\mathrm{DC},t}$ has the *reconstruction property* if

$$\liminf_{\ell\to\infty}\liminf_{n\to\infty}\mathbb{E}\left[\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1})=1\,\middle|\,\boldsymbol{F}_{\mathrm{DC},t},\left\{\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(y)\right\}_{y\in\partial^{2\ell}x_{t+1}}\right]-\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1})=1\mid\boldsymbol{F}_{\mathrm{DC},t}\right]\right|\,\middle|\,\boldsymbol{F}_{\mathrm{XOR}}\text{ sat.}\right]>0. \tag{4.3.6}$$

4.4 Different phases in random k-SAT and random k-XORSAT

There are several phase transitions affecting the structure of the set of solutions of typical instances predicted by *cavity method* in the satisfiable phase. Last few subsections describe the concept of correlation decay and pure state decomposition in the clustering transition. When we deal with models on locally tree like structure, we have encountered three main phases (also in random k-SAT) which can be studied using appropriate cavity method.

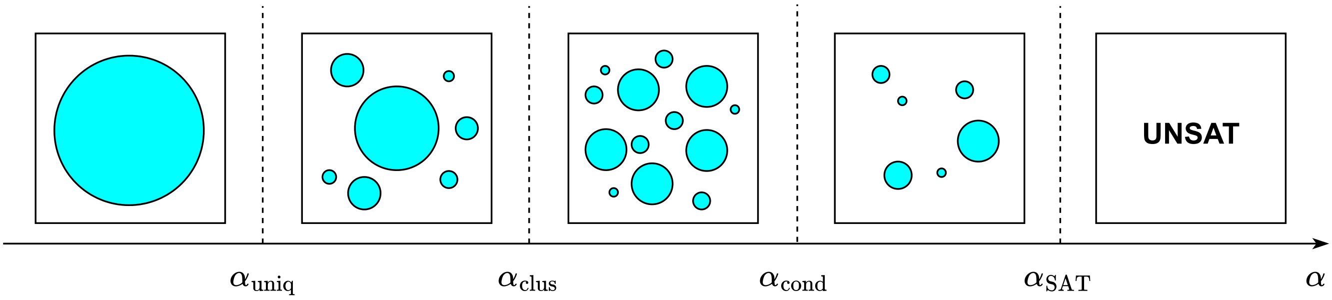


Figure 4.2: Phase diagram of k-SAT adapted and modified from [66]. **Left** to **Right**: Uniqueness, Clustering (*Replica Symmetry*), Clustering → Condensation (*dynamic* 1*RSB*), Condensation → Satisfiability (*static* 1*RSB*), UNSAT

- **Replica Symmetry:** We further divide this phase into two different regimes.
  - (i). $\alpha<\alpha_{\mathrm{uniq}}$: In this phase there exists no trivial decomposition into pure states. The system is in the replica symmetry phase with one big cluster of solution.
  - (ii). $\alpha_{\mathrm{uniq}}<\alpha<\alpha_{\mathrm{clus}}$: Although replica symmetry holds in this regime but the clusters form one big cluster along with exponentially tiny and scarce cluster.
- **Dynamic 1RSB** ($\alpha_{\mathrm{clus}}<\alpha<\alpha_{\mathrm{cond}}$): In this phase the Gibbs measure $\mu(\cdot)$ admits a non-trivial decomposition into an exponentially number of pure states. From the correlation point of view this phase is stable to small perturbations, but it is reconstructible. The solution space undergoes clustering that is exponentially many solution clusters of small size.
- **Static 1RSB** ($\alpha_{\mathrm{clus}}<\alpha<\alpha_{\mathrm{cond}}$): This is the 'original' 1RSB phase analogous to the low-temperature phase of the REM. This phase is not stable to small perturbations, and it is reconstructible implies it has long-range correlations. The number of solutions in this regime is bounded.

- **UNSAT**: Beyond the satisfiability threshold α_{SAT}, there exists no solution and the regime is referred as *Unsatisfiable phase.*

Remark 4.4.1. *In terms of statistical physics, one replica symmetry breaking (1RSB) cavity method gives a full overview of the solution space geometry of random k-SAT. According to physics method, the satisfiability threshold for random k-SAT turns out to be*

$$\alpha_{sat}(k) = 2^k \log 2 - \frac{1+\log 2}{2} + o_k(1)$$

Further for larger k values the satisfiability threshold have been studied rigorously [34, 41].

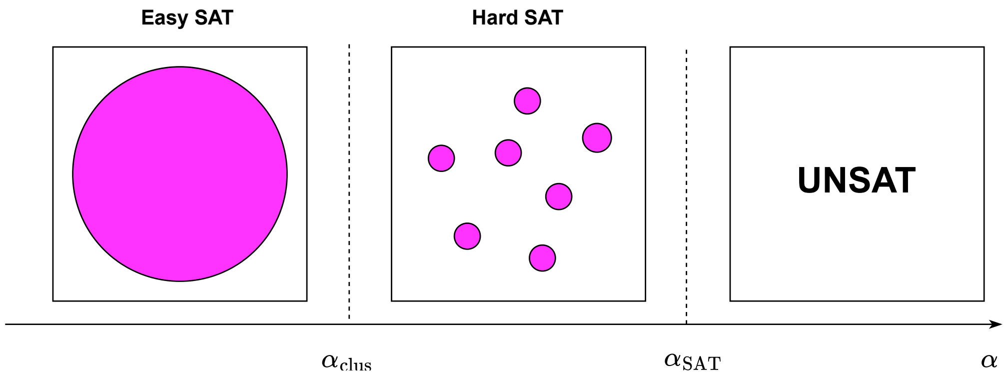


Figure 4.3: Phase diagram of k-XORSAT adapted and modified from [66]. **Left** to **Right**: Clustering (*Easy SAT*), Satisfiability(*Hard SAT*), UNSAT

k	3	4	5
α_{clus}	0.81847	0.77228	0.70178
α_{SAT}	0.91794	0.97677	0.99244

Table 4.1: The thresholds α_{clus} and α_{SAT} for various k [66]

On the other hand, in case of random k-XORSAT, the whole regime $\alpha < \alpha_{\text{SAT}}$ is the satisfiable phase. This means there exist solutions to the random linear system with high probability and more specifically, the number of solutions is given by, $Z = e^{n(1-\alpha)}$ where n is the number of variables of random k-XORSAT formula. From the figure it is clear that, the threshold α_{clus} separates two phases:

- $\alpha < \alpha_{\text{clus}}$: This phase is referred as '***Easy SAT***' where there is a single cluster of solutions.

- $\alpha_{\text{clus}} < \alpha < \alpha_{\text{SAT}}$: This phase is referred as '***Hard SAT*** [1]' where the solutions of linear system are grouped into well-separated clusters.

[1]The k-XORSAT problem can, of course, always be solved by '*Gaussian elimination* in $O(n^3)$ time with n is the number of involved variables and $m = \Theta(n)$ is the number of equations.

- $\alpha > \alpha_{\text{SAT}}$: In this regime there exists no solution and is referred as '***UNSAT***' phase.

Table 4.1 shows the thresholds for small k values. For the large k, $\alpha_{\text{clus}} \approx \frac{\log k}{k}$ and $\alpha_{\text{SAT}} \approx 1 - e^{-k} + O(e^{-2k})$ [66]. In the second result of this thesis, we analyze the performance of BPGD algorithm on thresholds of α values, more precisely on the threshold values of average variable degree d (where $\alpha = d/k$) with constant $k \geq 3$.

Based on:
The number of random 2-SAT solutions is asymptotically log-normal [23]
Arnab Chatterjee, Amin Coja-Oghlan, Noela Müller, Connor Riddlesden, Maurice Rolvien, Pavel Zakharov, Haodong Zhu
Proc. 28th RANDOM (2024)

5

A Central Limit Theorem for random 2-SAT solutions

"The central limit theorem is the most fundamental theory in modern statistics.$\cdots$ With the central limit theorem, parametric tests have higher statistical power than non-parametric tests, which do not require probability distribution assumptions."

–Sang Gyu Kwak et.al.

Till now we have explored the random constraint satisfaction models, the different message passing algorithms and the phase transitions depend on the values of variable to clause density (α) and the solution space geometry based on the values of α. The goal of this chapter is to provide a deeper insight on the estimation of partition function i.e., the number of solutions (more precisely, the logarithm of the number of satisfying assignments) in random 2-SAT, the simplest random CSP model.

5.1 Motivation and History

The hunt for satisfiability thresholds has been a guiding theme of research into random constraint satisfaction problems [6, 24, 41]. Once the satisfiability threshold has been pinpointed, the next obvious question should come into one's mind is to determine the distribution of satisfying assignments within the satisfiable phase [59]. Indeed, the number of such solutions is intimately tied to phase transitions that affect the solution space geometry, which in turn impacts the computational nature of finding or sampling solutions [1, 18, 44]. Despite its importance, the problem of counting solutions in random CSPs remains difficult, with few general-purpose tools currently available. In those instances where precise, rigorous results have been obtained, such as for random NAE-SAT or XORSAT, the proofs commonly rely on the method of moments (e.g., [4, 42, 93, 95]). A necessary condition for the success of this approach is that the problem exhibits certain symmetries which are absent in many interesting cases [6, 31]. Random 2-SAT, the simplest random constraint satisfaction problem lacking the aforementioned symmetry properties, is therefore an intriguing topic of study.

The number of satisfying assignments in random 2-SAT

Let $\boldsymbol{F}_{\text{2SAT}} = \boldsymbol{F}_{n,m}$ be a random 2-CNF on n Boolean variables $x_1,\dots,x_n$ with m clauses, drawn independently and uniformly from all $4\binom{n}{2}$ possible 2-clauses. Further assume that $m \sim dn/2$ for a fixed real $d > 0$. Since 1990s it has been known that $\boldsymbol{F}_{\text{2SAT}}$ is satisfiable w.h.p. if $d < 2$, and unsatisfiable w.h.p. if d>2 [26, 53]. Whereas the first order approximation to the number of satisfying assignments has been studied recently [2]. Alongside this, calculating the number of satisfying assignments $Z(\boldsymbol{F}_{\text{2SAT}})$ is a #P-hard task [1] [100]. Nonetheless, **Monasson** and **Zecchina** in [77] put forward a conjecture on the exponential order of the number of satisfying assignments of random 2-CNFs using physics inspired technique. In 2021 **Achlioptas et al.** [2] provides a first order on the logarithm of the number of satisfying assignments using law of large number approximation by introducing a function $\phi(d) > 0$ such that for all $d < 2$, i.e.,throughout the entire satisfiable phase,

$$\log Z(\boldsymbol{F}_{\text{2SAT}}) = n\phi(d) + o(n) \qquad \text{w.h.p.,} \tag{5.1.1}$$

In this thesis we determine not only the leading order of $\log Z(\boldsymbol{F}_{\text{2SAT}})$ but also its fluctuations. We also provide a precise result by showing that the logarithm of the number of satisfying assignments converges to a Gaussian throughout the satisfiable regime – the first central limit theorem ('CLT') of this type for any random CSPs.

5.2 Main Result.

In this section we state our first result of this thesis [23].

Let $\mathscr{P}(\mathbb{R}^2)$ be the set of all (Borel) probability measures on $\mathbb{R}^2$. For $0 < d < 2$ and $0 \le t \le 1$ we define an operator

$$\texttt{logBP}^{\otimes}_{d,t} : \mathscr{P}\left(\mathbb{R}^2\right) \to \mathscr{P}\left(\mathbb{R}^2\right), \qquad \rho \mapsto \hat{\rho} = \texttt{logBP}^{\otimes}_{d,t}(\rho), \tag{5.2.1}$$

as follows. Let

$$(\boldsymbol{\xi}_{\rho,i})_{i\ge1},\ (\boldsymbol{\xi}'_{\rho,i})_{i\ge1},\ (\boldsymbol{\xi}''_{\rho,i})_{i\ge1}, \qquad \boldsymbol{\xi}_{\rho,i} = \begin{pmatrix}\boldsymbol{\xi}_{\rho,i,1}\\ \boldsymbol{\xi}_{\rho,i,2}\end{pmatrix}, \boldsymbol{\xi}'_{\rho,i} = \begin{pmatrix}\boldsymbol{\xi}'_{\rho,i,1}\\ \boldsymbol{\xi}'_{\rho,i,2}\end{pmatrix}, \boldsymbol{\xi}''_{\rho,i} = \begin{pmatrix}\boldsymbol{\xi}''_{\rho,i,1}\\ \boldsymbol{\xi}''_{\rho,i,2}\end{pmatrix}$$

be random vectors with distribution ρ, let $\boldsymbol{d} \overset{\text{d}}{=} \text{Po}(td)$, $\boldsymbol{d}', \boldsymbol{d}'' \overset{\text{d}}{=} \text{Po}((1-t)d)$ and let $\boldsymbol{s}_i, \boldsymbol{s}'_i, \boldsymbol{s}''_i, \boldsymbol{r}_i, \boldsymbol{r}'_i, \boldsymbol{r}''_i$ for

[1] #P-hard comprises counting problems that ask for the existence of the number of solutions for a given NP decision problem. A problem is said to be #P-hard if it is as hard as any problem in the class #P. For this reason, any problem in #P can be reduced to it in polynomial time.

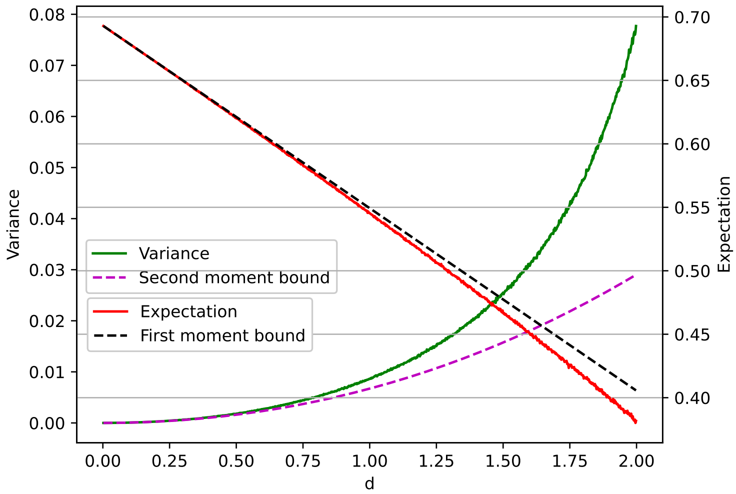


Figure 5.1: Numerical approximations to the function $\phi(d)$ from (5.1.1) (red) and the variance $\eta(d)^2$ from (5.2.5) (green). The black dashed line is the first moment bound $d \mapsto \log(2) + \frac{d}{2}\log(3/4)$ whereas the purple dashed line is the second moment bound. (Figure 1, [23])

$i \geq 1$ be uniformly random on $\{\pm 1\}$, all mutually independent. Then $\hat{\rho}$ is the distribution of the vector

$$\begin{pmatrix} \sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \log\left(\frac{1}{2}\left(1+\boldsymbol{r}_i \tanh(\boldsymbol{\xi}_{\rho,i,1}/2)\right)\right) + \sum_{i=1}^{\boldsymbol{d}'} \boldsymbol{s}'_i \log\left(\frac{1}{2}\left(1+\boldsymbol{r}'_i \tanh(\boldsymbol{\xi}'_{\rho,i,1}/2)\right)\right) \\ \sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \log\left(\frac{1}{2}\left(1+\boldsymbol{r}_i \tanh(\boldsymbol{\xi}_{\rho,i,2}/2)\right)\right) + \sum_{i=1}^{\boldsymbol{d}''} \boldsymbol{s}''_i \log\left(\frac{1}{2}\left(1+\boldsymbol{r}''_i \tanh(\boldsymbol{\xi}''_{\rho,i,2}/2)\right)\right) \end{pmatrix} \in \mathbb{R}^2 .$$

In addition, define a function $\mathscr{B}^{\otimes}_{d,t} : \mathscr{P}(\mathbb{R}^2) \to (0,\infty]$ by letting

$$\mathscr{B}^{\otimes}_{d,t}(\rho) = \mathbb{E}\left[\prod_{h=1}^{2} \log\left(1 - \frac{1}{4}(1+\boldsymbol{r}_1 \tanh(\boldsymbol{\xi}_{\rho,1,h}/2))(1+\boldsymbol{r}_2 \tanh(\boldsymbol{\xi}_{\rho,2,h}/2))\right)\right]. \tag{5.2.2}$$

The main theorem establishes a CLT for the logarithm of the number of solutions of $\boldsymbol{F}_{2\mathrm{SAT}}$, with a standard deviation that is closely connected to the aforementioned function $\mathscr{B}^{\otimes}_{d,t}$:

Theorem 5.2.1 (Theorem 1.1, [23])**.** *For any* $0 < d < 2$, $t \in [0,1]$ *there exists a unique probability measure* $\rho_{d,t} \in \mathscr{P}(\mathbb{R}^2)$ *such that*

$$\rho_{d,t} = \mathtt{logBP}^{\otimes}_{d,t}(\rho_{d,t}) \qquad \textit{and} \qquad \int_{\mathbb{R}^2} |\xi|_2^2 \,\mathrm{d}\rho_{d,t}(\xi) < \infty. \tag{5.2.3}$$

Furthermore,

$$\lim_{n\to\infty} \frac{\log Z(\boldsymbol{F}_{2SAT}) - \mathbb{E}[\log Z(\boldsymbol{F}_{2SAT}) \mid Z(\boldsymbol{F}_{2SAT}) > 0]}{\sqrt{m}} = \boldsymbol{\Gamma}_{\eta(d)} \qquad \textit{in distribution, where} \tag{5.2.4}$$

$$\eta(d)^2 = \int_0^1 \mathscr{B}^{\otimes}_{d,t}(\rho_{d,t})\mathrm{d}t - \mathscr{B}^{\otimes}_{d,0}(\rho_{d,0}) \in (0,\infty). \tag{5.2.5}$$

Remark 5.2.2. *Note that, the conditioning on* $\log Z(\boldsymbol{F}_{2SAT}) > 0$ *is necessary in* (5.2.4), *because even for* $d < 2$ *the formula* $\boldsymbol{F}_{2SAT}$ *is unsatisfiable with probability* $\Omega(n^{-1})$, *in which case* $\log Z(\boldsymbol{F}_{2SAT}) = -\infty$. *Moreover, the* L^2*-bound ensures that the integral* (5.2.5) *is well-defined. Finally,* (5.2.4) *implies that,*

$$\mathbb{P}\left[\log Z(\boldsymbol{F}_{2SAT}) - \mathbb{E}[\log Z(\boldsymbol{F}_{2SAT}) \mid Z(\boldsymbol{F}_{2SAT}) > 0] < z\sqrt{m}\right] \sim \mathbb{P}\left[\boldsymbol{\Gamma}_{\eta(d)} < z\right] \qquad (z \in \mathbb{R}). \tag{5.2.6}$$

So, the first result of this thesis [23] addresses two questions.

Question 1**.** How to show the asymptotic normality in (5.2.4) ?

Question 2**.** How to calculate the variance of the formula effectively?

Evaluating Standard Deviation.

Towards answering the second question, we know that the proof of the uniqueness of the stochastic fixed point $\rho_{d,t}$ from (5.2.3) is based on the contraction method, a fixed point iteration will converge rapidly. In effect, for any d, t a discrete distribution that approximates $\rho_{d,t}$ arbitrarily well (in Wasserstein distance) can be computed via a randomized heuristic algorithm called *population dynamics* [66, Chapter 14]. Since $\mathscr{B}^{\otimes}_{d,t}(\rho_{d,t})$ varies continuously in d and t, $\eta(d)^2$ can thus be approximated within any desired accuracy, see Figure 5.1.

Study Fixed Point $\rho_{d,t}$.

The distribution of each coordinate of the fixed point has been studied by Múller, Neininger and Zhu in [86]. It is shown that when $d \le 1$, the corresponding measure is purely discrete. After applying the transformation $\frac{1+\tanh(\cdot/2)}{2}$ the support of the measure consists of rational numbers in $(0,1)$ for all $0 < d < 2$. Moreover, when $d \in (1,2)$, the measure acquires a continuous part.

5.3 Proof Strategy

The main hurdle towards the proof of Theorem 5.2.1 is to compute the variance of $\log Z(\boldsymbol{F}_{2SAT})$ given satisfiability. The key idea, inspired by spin glass theory [25] but new to any random CSPs, is to count the *joint* number of satisfying assignments of two correlated random formulas. Once this is accomplished Theorem 5.2.1 will follow from a general martingale central limit theorem. To get habituated we first revisit the method of moments, the reasons it fails on random 2-SAT and the combinatorial interpretation of the law of large numbers (5.1.1).

5.3.1 Method of Moments fails.

The default approach to estimating the number of solutions to a random CSP is the venerable second moment method [6]. Its core idea is to show that the second moment of the number of solutions is of the same order as the square of the expected number of solutions. If so then the moment computation together with small subgraph conditioning yields the precise limiting distribution of the number of solutions [35, 97]. However, this approach works only if the log of the number of solutions superconcentrates around the log of the expected number of solutions which does not hold in random 2-SAT. In fact, a straightforward calculation yields that,

$$\frac{1}{n}\log\mathbb{E}[Z(\boldsymbol{F}_{2\mathrm{SAT}})] \sim \log 2 + \frac{d}{2}\log(3/4). \tag{5.3.1}$$

The formula on the r.h.s. is displayed as the black dashed line in Figure 5.1. As can be verified analytically, this line strictly exceeds the function $\phi(d)$ from (5.1.1) for any $0 < d < 2$. Consequently, (5.1.1) implies that $\log Z(\boldsymbol{F}_{2\mathrm{SAT}}) \le \log\mathbb{E}[Z(\boldsymbol{F}_{2\mathrm{SAT}})] - \Omega(n)$ w.h.p. In other words, the expected number of solutions $\mathbb{E}[Z(\boldsymbol{F}_{2\mathrm{SAT}})]$ overshoots the typical number of solutions by an exponential factor w.h.p. (for details, see the discussion in [4, 7]).

Rather than relying on the method of moments, Monasson and Zecchina in [77] proposed a physics-inspired approach for estimating $\log Z(\boldsymbol{F}_{2\mathrm{SAT}})$ using the Belief Propagation algorithm which is discussed in Chapter 3 of this thesis. This approach was established rigorously by Achlioptas et.al. [2].

5.3.2 BP Approximation.

The Belief propagation messages introduced in Chapter 3 get updated iteratively by an operator

$$\mathrm{BP} : (\nu_{x\to a}, \nu_{a\to x})_{a,x\in\partial a} \mapsto (\hat\nu_{x\to a}, \hat\nu_{a\to x})_{a,x\in\partial a} = \mathrm{BP}((\nu_{x\to a}, \nu_{a\to x})_{a,x\in\partial a}). \tag{5.3.2}$$

where, for $\partial a = \{x, y\}$ the updated messages $\hat\nu_{a\to x}(\pm 1)$ are defined by

$$\hat\nu_{a\to x}(\mathrm{sign}(x,a)) = \frac{1}{1+\nu_{y\to a}(\mathrm{sign}(y,a))}, \qquad \hat\nu_{a\to x}(-\mathrm{sign}(x,a)) = \frac{\nu_{y\to a}(\mathrm{sign}(y,a))}{1+\nu_{y\to a}(\mathrm{sign}(y,a))}. \tag{5.3.3}$$

Moreover, for a variable x and a clause $a \in \partial x$ we define [2]

$$\hat\nu_{x\to a}(s) = \frac{\prod_{b\in\partial x\setminus\{a\}} \nu_{b\to x}(s)}{\prod_{b\in\partial x\setminus\{a\}} \nu_{b\to x}(1) + \prod_{b\in\partial x\setminus\{a\}} \nu_{b\to x}(-1)} \qquad (s \in \{\pm 1\}) \ . \tag{5.3.4}$$

The purpose of BP is to heuristically ‘approximate’ the marginal probabilities that a random satisfying assignment $\boldsymbol{\sigma} = \boldsymbol{\sigma}_{\boldsymbol{F}_{2\mathrm{SAT}}}$ of $\boldsymbol{F}_{2\mathrm{SAT}}$ will set a certain variable to a specific truth value. The ‘approximation’ given by the set $(\nu_{x\to a}, \nu_{a\to x})_{a,x\in\partial a}$ of messages reads

$$\nu_x(s) = \frac{\prod_{b\in\partial x} \nu_{b\to x}(s)}{\prod_{b\in\partial x} \nu_{b\to x}(1) + \prod_{b\in\partial x} \nu_{b\to x}(-1)} \qquad (s \in \{\pm 1\}). \tag{5.3.5}$$

Equation (5.3.5) suggests that the BP operator should be iterated until a fixed point is reached; that is, until $\hat\nu_{x\to a} = \nu_{x\to a}$ and $\hat\nu_{a\to x} = \nu_{a\to x}$ hold for all x, a. We then evaluate (5.3.5) and substitute them into a general expression known as *Bethe free entropy*, which yields the BP *approximation* of $\log Z(\boldsymbol{F}_{2\mathrm{SAT}})$.

[2] For the sake of tidyness, if the above denominator vanishes we simply let $\hat\mu_{x\to a}(\pm 1) = \frac{1}{2}$.

This BP approximation is accurate when the bipartite graph induced by the clause-variable incidences of the 2-CNF $\boldsymbol{F}_{\text{2SAT}}$ is acyclic, but it may produce inaccurate results in the presence of cycles.

5.3.3 Towards calculating variance.

The proof of the formula (5.1.1) combines the Gibbs uniqueness property discussed in Chapter and the local convergence to the Galton-Watson tree with a coupling argument called the '*Aizenman-Sims-Starr* scheme' [2]. Unfortunately it is not clear how the order of the standard deviation of $\log Z(\boldsymbol{F}_{\text{2SAT}})$ could be derived because the main problem relies on the convergence of the Gibbs uniqueness property diminishes as d approaches the satisfiability threshold. To overcome this challenge, we develop a combinatorial interpretation of $\log^2(Z(\boldsymbol{F}_{\text{2SAT}}))$ by constructing a correlated pair $(\boldsymbol{F}_1(M,M'),\boldsymbol{F}_2(M,M'))$ for any given integers $M,M'\geq 0$ of formulas on the variable set $V_n=\{x_1,\dots,x_n\}$ as follows. Let $(\boldsymbol{a}_i)_{i\geq 1}$, $(\boldsymbol{a}'_i)_{i\geq 1}$, $(\boldsymbol{a}''_i)_{i\geq 1}$ be sequences of mutually independent uniformly random clauses on V_n. Then

$$\boldsymbol{F}_1(M,M')=\boldsymbol{a}_1\wedge\cdots\wedge\boldsymbol{a}_M\wedge\boldsymbol{a}'_1\wedge\cdots\wedge\boldsymbol{a}'_{M'}\quad\text{and,}\tag{5.3.6}$$

$$\boldsymbol{F}_2(M,M')=\boldsymbol{a}_1\wedge\cdots\wedge\boldsymbol{a}_M\wedge\boldsymbol{a}''_1\wedge\cdots\wedge\boldsymbol{a}''_{M'}.\tag{5.3.7}$$

Thus, the two formulas share clauses $\boldsymbol{a}_1,\dots,\boldsymbol{a}_M$. Additionally, each contains another M' independent clauses. In particular, $\boldsymbol{F}_1(m,0)$, $\boldsymbol{F}_2(m,0)$ are identical, while $\boldsymbol{F}_1(0,m)$, $\boldsymbol{F}_2(0,m)$ are independent. For computing the variance given that $\boldsymbol{F}_1(M,m-M)$ and $\boldsymbol{F}_2(M,m-M)$ are satisfiable for all M, we can write a telescoping sum

$$\begin{aligned}&\log Z(\boldsymbol{F}_1(m,0))\cdot\log Z(\boldsymbol{F}_2(m,0))-\log Z(\boldsymbol{F}_1(0,m))\cdot\log Z(\boldsymbol{F}_2(0,m))\\&\quad=\sum_{M=1}^{m}\log Z(\boldsymbol{F}_1(M,m-M))\cdot\log Z(\boldsymbol{F}_2(M,m-M))\\&\qquad\qquad-\log Z(\boldsymbol{F}_1(M-1,m-M+1))\cdot\log Z(\boldsymbol{F}_2(M-1,m-M+1)).\end{aligned}\tag{5.3.8}$$

Clearly, if we could take the expectation on the l.h.s. of (5.3.8), we would precisely obtain the variance of $\log Z(\boldsymbol{F}_{\text{2SAT}})$. However, we cannot just take the expectation of (5.3.8), because some $\boldsymbol{F}_h(M,m-M)$ may be unsatisfiable for $(h=1,2)$. potentially leading to occurrences of $\log 0$. To address this issue we replace $\log Z(\boldsymbol{F}_{\text{2SAT}})$ with more tractable random variable sharing same limiting distribution whose construction is based on *'Unit Clause Propagation'* discussed in Chapter 3, Section 3.2.3. Thus we obtain a pruned formula $\hat{\boldsymbol{F}}_{\text{2SAT}}$ from the original 2-CNF $\boldsymbol{F}_{\text{2SAT}}$ and the following Fact can be verified (for more details of the proof see [23]).

Fact 5.3.1 (Fact 2.2, [23])**.** *For any 2-CNF $\boldsymbol{F}_{\text{2SAT}}$, the pruned 2-CNF $\hat{\boldsymbol{F}}_{\text{2SAT}}$ is satisfiable.*

Note that, even if $\boldsymbol{F}_{\text{2SAT}}$ is satisfiable the number $Z(\hat{\boldsymbol{F}}_{\text{2SAT}})$ of satisfying assignments of $\hat{\boldsymbol{F}}_{\text{2SAT}}$ could dramatically exceed $Z(\boldsymbol{F}_{\text{2SAT}})$ as the pruned formula $\hat{\boldsymbol{F}}_{\text{2SAT}}$ generally have far fewer clauses than the original formula $\boldsymbol{F}_{\text{2SAT}}$. However, the following proposition shows that on a random formula, the impact of pruning is modest.

Proposition 5.3.2 (Proposition 2.3, [23])**.** $|\log Z(\hat{\boldsymbol{F}}_{2SAT}) - \log Z(\boldsymbol{F}_{2SAT})| \leq n^{1/3}$. *with probability* $1 - o(n^{-1/2})$.

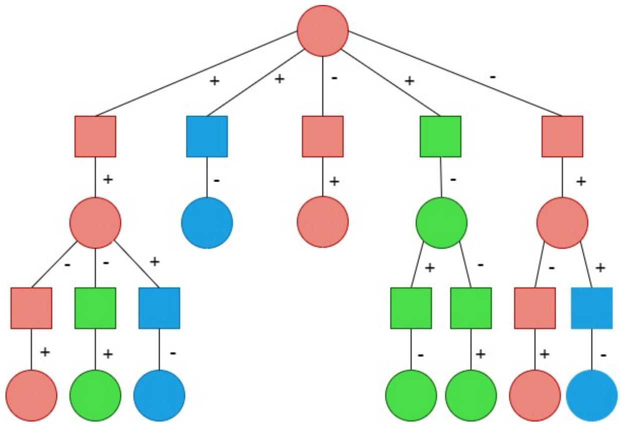


Figure 5.2: An illustration of the correlated GW-tree $\boldsymbol{T}^{\otimes}$ (Figure 1, [23])

As the error bound from Proposition 5.3.2 is tight, it suffices to establish a CLT for the log of the number of satisfying assignments of the pruned formula $\log Z(\hat{\boldsymbol{F}}_{\mathrm{2SAT}})$. Revisiting the telescoping sum (5.3.8) we obtain the following lemma Lemma 5.3.3 which expresses the variance as a sum of local changes. For example, $\boldsymbol{\Phi}_1(M, m-M)$ is obtained from $\boldsymbol{\Phi}_1(M-1, m-M)$ by adding a single random clause, namely $\boldsymbol{a}_M$. On the other hand, only a few clauses are pruned from random formulas w.h.p. Expanding the variance $\mathrm{Var}\left(\log Z(\hat{\boldsymbol{F}}_{\mathrm{2SAT}})\right)$ as follows:

Lemma 5.3.3 (Lemma 2.4, [23])**.** *Let*

$$\Delta(M) = \mathbb{E}\left[\log\left(\frac{Z(\hat{\boldsymbol{F}}_1(M, m-M))}{Z(\hat{\boldsymbol{F}}_1(M-1, m-M))}\right) \cdot \log\left(\frac{Z(\hat{\boldsymbol{F}}_2(M, m-M))}{Z(\hat{\boldsymbol{F}}_2(M-1, m-M))}\right)\right], \tag{5.3.9}$$

$$\Delta'(M) = \mathbb{E}\left[\log\left(\frac{Z(\hat{\boldsymbol{F}}_1(M-1, m-M+1))}{Z(\hat{\boldsymbol{F}}_1(M-1, m-M))}\right) \cdot \log\left(\frac{Z(\hat{\boldsymbol{F}}_2(M-1, m-M+1))}{Z(\hat{\boldsymbol{F}}_2(M-1, m-M))}\right)\right]. \tag{5.3.10}$$

Then $\mathrm{Var}\left[\log Z(\hat{\boldsymbol{F}}_{2SAT})\right] = \sum_{M=1}^{m} \Delta(M) - \Delta'(M)$.

Now for the analysis of the correlated formulas we need the following expressions to evaluate.

Proposition 5.3.4 (Proposition 2.5, [23])**.** *Let* $1 \leq M \leq m$. *Then,*

$$\frac{Z(\hat{\boldsymbol{F}}_h(M, m-M))}{Z(\hat{\boldsymbol{F}}_h(M-1, m-M))} = 1 - \prod_{y \in \partial \boldsymbol{a}_M} \mathbb{P}\left[\boldsymbol{\sigma}_y \neq \mathrm{sign}(y, \boldsymbol{a}_M) \mid \hat{\boldsymbol{F}}_h(M-1, m-M), \boldsymbol{a}_M\right] + o(1) \quad (h = 1, 2),$$

$$\frac{Z(\hat{\boldsymbol{F}}_1(M-1, m-M+1))}{Z(\hat{\boldsymbol{F}}_1(M-1, m-M))} = 1 - \prod_{y \in \partial \boldsymbol{a}'_{m-M+1}} \mathbb{P}\left[\boldsymbol{\sigma}_y \neq \mathrm{sign}(y, \boldsymbol{a}'_{m-M+1}) \mid \hat{\boldsymbol{F}}_1(M-1, m-M), \boldsymbol{a}'_{m-M+1}\right] + o(1),$$

$$\frac{Z(\hat{\boldsymbol{F}}_2(M-1, m-M+1))}{Z(\hat{\boldsymbol{F}}_2(M-1, m-M))} = 1 - \prod_{y \in \partial \boldsymbol{a}''_{m-M+1}} \mathbb{P}\left[\boldsymbol{\sigma}_y \neq \mathrm{sign}(y, \boldsymbol{a}''_{m-M+1}) \mid \hat{\boldsymbol{F}}_2(M-1, m-M), \boldsymbol{a}'_{m-M+1}\right] + o(1).$$

We construct a Galton-Watson tree $\boldsymbol{T}^{\otimes}$ that approximates the *joint* distribution of the local structure of the pair $(\hat{\boldsymbol{F}}_1(M-1, m-M), \hat{\boldsymbol{F}}_2(M-1, m-M))$.

Shared variables/clauses are indicated in red, 1-distinct variables/clauses in green and 2-distinct ones in blue in Figure 5.2. (for more details refer to [23]) From $\boldsymbol{T}^{\otimes}$ we extract a pair $(\boldsymbol{T}_1, \boldsymbol{T}_2)$ of correlated random trees. Specifically, $\boldsymbol{T}_h$ is obtained from $\boldsymbol{T}^{\otimes}$ by deleting all $(3-h)$-distinct variables and clauses. Hence, the parameter t determines how 'similar' $\boldsymbol{T}_1, \boldsymbol{T}_2$ are. As we have generated a pair of random formulas and take a uniformly random pair of satisfying assignments, the joint distribution of any of n coordinates can be viewed on the heatmaps (shown in Figure 5.3): almost independent formulas on the left and highly correlated formulas on the right.

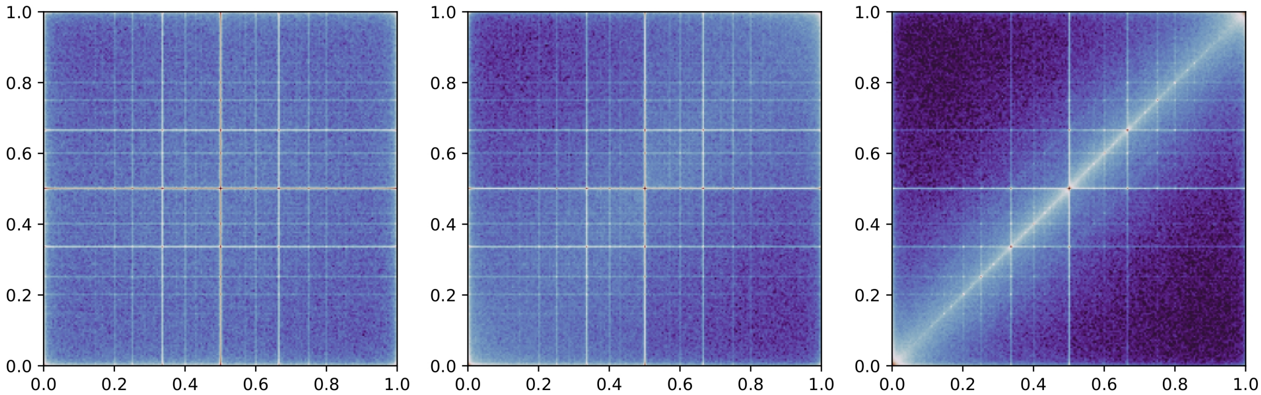


Figure 5.3: Marginal distribution on two correlated formulas for $d = 0.9$ and $M = 0.1m, 0.5m, 0.9m$ (Figure 2, [23])

Now in our hand we have a pair of correlated formulas the next step is to run BP on the random trees $(\boldsymbol{T}_1, \boldsymbol{T}_2)$ to find the joint distribution of the truth values $\boldsymbol{\sigma}_{\boldsymbol{T}_1^{(2\ell)},o}, \boldsymbol{\sigma}_{\boldsymbol{T}_2^{(2\ell)},o}$ assigned to the root o. Fortunately, due to the Markovian nature of the Galton-Watson tree $\boldsymbol{T}^{\otimes}$, the bottom-up BP computation on a random tree can be expressed by a fixed point iteration on the space of probability distributions on $\mathbb{R}^2$. The most appropriate operator $\mathtt{logBP}^{\otimes}_{d,t}$ expresses the updates of the log-likelihood ratios of the BP messages from (5.3.3)–(5.3.4). Thus the followings hold:

Proposition 5.3.5 (Proposition 2.8, [23])**.** *There exists a unique* $\rho_{d,t} \in \mathscr{P}(\mathbb{R}^2)$ *that satisfies* (5.2.3) *and* $\lim_{\ell\to\infty} \rho^{(\ell)}_{d,t} = \rho_{d,t}$ *weakly.*

Corollary 5.3.6 (Corollary 2.11, [23])**.** *With* $\eta(d)^2$ *from* (5.2.5) *we have* $\eta(d) > 0$ *and* $\mathrm{Var}\log Z(\hat{\boldsymbol{F}}_{2SAT}) \sim m\eta_d^2$.

The proof of Proposition 5.3.5 is based on a contraction argument, for any d, t the distribution $\rho_{d,t}$ can be approximated effectively within any given accuracy via a fixed point iteration. From the contraction argument on the evaluation of the functional $\mathscr{B}^{\otimes}_{d,t}$ on $\rho_{d,t}$ yield finite values for any $d \in (0,2)$ and $t \in [0,1]$ which implies the finiteness of the variance.

Lemma 5.3.7 (Lemma 10.1, [23])**.** *For any* $d \in (0,2)$ *and* $t \in [0,1]$, $\mathscr{B}^{\otimes}_{d,t}(\rho_{d,t}) < \infty$. *Moreover, for any* $d \in (0,2)$, $\eta(d)^2 < \infty$.

Finally, along with Proposition 5.3.4, the BP arguments on correlated formulas give us the variance of $\log Z(\hat{\boldsymbol{F}}_{2\text{SAT}})$.

5.4 Establishing the Central Limit Theorem

Once finished calculating variance we set up a filtration $(\mathfrak{F}_{n,M})_{0\le M\le m_n}$ by letting $\mathfrak{F}_{n,M}$ be the σ-algebra generated by $\boldsymbol{a}_1,\dots,\boldsymbol{a}_M$. The conditional expectations is given by,

$$\boldsymbol{Z}_{n,M} = m^{-1/2}\mathbb{E}\left[\log Z(\hat{\boldsymbol{F}}_{2\text{SAT}}) \mid \mathfrak{F}_{n,M}\right] \tag{5.4.1}$$

then form a Doob martingale. Let $\boldsymbol{X}_{n,M} = \boldsymbol{Z}_{n,M} - \boldsymbol{Z}_{n,M-1}$ be the martingale differences.

Proposition 5.4.1 (Proposition 2.12, [23]). *For all $0<d<2$ the martingale (5.4.1) satisfies*

$$\lim_{n\to\infty}\mathbb{E}\left[\max_{1\le M\le m}|\boldsymbol{X}_{n,M}|\right]=0 \qquad \textit{and,} \tag{5.4.2}$$

$$\lim_{n\to\infty}\mathbb{E}\left|\eta(d)^2-\sum_{M=1}^{m}\boldsymbol{X}_{n,M}^2\right|=0. \tag{5.4.3}$$

Clearly, the above conditions can be checked easily with the help of pruning argument.

Thus we conclude this chapter by deriving the main result of this chapter from the following general martingale central limit theorem, which is a special case of [55, Theorem 3.2].

Theorem 5.4.2 ([55, Theorem 3.2]). *Let $(\boldsymbol{Z}_{n,i},\mathfrak{F}_{n,i})_{0\le i\le m_n,n\ge1}$ be a zero-mean, square-integrable martingale array with differences $\boldsymbol{X}_{n,i}=\boldsymbol{Z}_{n,i}-\boldsymbol{Z}_{n,i-1}$ for $1\le i\le m_n$. Assume that there exists a constant η^2 such that*

$$\lim_{n\to\infty}\max_{1\le i\le m_n}|\boldsymbol{X}_{n,i}|=0 \qquad \textit{in probability,} \tag{5.4.4}$$

$$\lim_{n\to\infty}\sum_{i=1}^{m_n}\boldsymbol{X}_{n,i}^2=\eta^2 \qquad \textit{in probability,} \tag{5.4.5}$$

$$\mathbb{E}\left[\max_{1\le i\le m_n}\boldsymbol{X}_{n,i}^2\right] \qquad \textit{is bounded in } n. \tag{5.4.6}$$

Then $\boldsymbol{Z}_{n,m_n}$ converges in distribution to a Gaussian distribution with mean zero and variance η^2.

Observe that,Proposition 5.4.1 directly implies the conditions of Theorem 5.4.2. Lastly, for the finiteness and positiveness of the variance, Lemma 5.3.7 guarantees that $\eta(d)<\infty$, while Corollary 5.3.6 shows that $\eta(d)>0$.

Based on:
***Belief Propagation Guided Decimation on Random k-XORSAT* [22]**
Arnab Chatterjee, Amin Coja-Oghlan, Mihyung Kang, Lena Krieg,
Maurice Rolvien, Gregory Sorkin
Proc. 52nd ICALP (2025)

6
Performance of BPGD on random k-XORSAT

" BPGD enhances conventional BP by sequentially fixing (decimating) variable nodes based on their belief values and reducing the solution space."

–Masoumeh Alinia et.al.

As we established a central limit theorem holds for the number of solutions of random 2-SAT formulae in the previous chapter, we now shift our attention from random 2-SAT to another random satisfiability problem namely *random k-XORSAT*, one of the simplest examples of random contraint satisfaction problems exhibiting sharp phase transition. More precisely, we will analyze the performance of Belief Propagation Guided Decimation ('BPGD') introduced in Chapter 3 by mathematically verified the heuristic work by Ricci-Tersenghi and Semerjian [96]. In addition to this, we study a thought experiment called *'decimation process'* (also initiated in Chapter 3) for which we identify a (non)-reconstruction and condensation phase transition. Begin the chapter with some motivation and background behind this work.

6.1 Motivation and History

The random k-XORSAT exhibits many features common to other intensely studied random CSPs, such as random k-SAT. At the same time, the random k-XORSAT is mathematically more compliant than say random k-SAT because a XORSAT instance translates into a linear system over $\mathbb{F}_2$ as XOR operation is equivalent to addition modulo two. In addition, the algebraic nature of the problem induces strong symmetry properties that simplify its study [12]. Since early 2000, in combinatorics as well as in statistical physics there has been contributing intriguing 'prediction' on different random CSPs. Furthermore, in 2008 **Ricci-Tersenghi** and **Semerjian** in [96] put forward a heuristic analysis of BPGD on both random k-SAT and k-XORSAT. Later **Coja-Oghlan** and **Pachon-Pinzon**, demonstrated both 'decimation process' and 'BPGD' rigorously on random k-SAT [28,33] by assuming clause length k is sufficiently large due to the lack of inherent symmetry in random k-SAT. In a recent paper [102], **Yung** in 2024 a first step towards the rigorous analysis of BPGD on random k-XORSAT has been undertaken. However, Yung's analysis turns out to be not tight. Specifically, apart from requiring spurious lower bounds on the clause length k, Yung's results do not quite establish the precise connection between the decimation process and the performance of BPGD. One reason for this is that [102] relies on 'annealed' techniques, i.e., essentially moment computations. Here we instead harness 'quenched' arguments to proof the success probability of BPGD and to make a precise connection between BPGD and the decimation process. The next section provides a very brief overview of the main problem addressed in this work along with their results.

6.2 Problem Statement and Results.

let $\boldsymbol{F}_{\text{XOR}} = \boldsymbol{F}_(n,d,k)$ be a random k-XORSAT formula with variables $x_1,\dots,x_n$ and $\boldsymbol{m}$ random clauses of length k where $\boldsymbol{m} \stackrel{\text{d}}{=} \text{Po}(dn/k)$. The $\boldsymbol{m}$ clauses are drawn uniformly and independently out of the set of all $2^k\binom{n}{k}$ possibilities. Thus, $d>0$ equals the average number of clauses that a given variable x_i appears in. Moreover, every clause of $\boldsymbol{F}_{\text{XOR}}$ is an XOR of precisely k distinct variables with $k \geq 3$, each of which may or may not come with a negation sign. Mathematically, if we are handed a number of independent random constraints (clauses) c_i of the type

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

where each y_{ij} is either one of n available Boolean variables $x_1,\dots,x_n$ or a negation $\neg x_1,\dots,\neg x_n$. As we know that boolean XOR boils down to addition over $\mathbb{F}_2$, this problem can be rephrased as the full rank problem for the random matrix $\boldsymbol{A}$ with $q=2$, $\boldsymbol{k}=k$ fixed to a deterministic value. Furthermore, the random negation patterns of the constraints amount to choosing a random right-hand side vector $\boldsymbol{y}$ for which we are to solve $\boldsymbol{A}x=\boldsymbol{y}$.

Let $\boldsymbol{A}$ be a matrix representation of a random k-XORSAT formula $\boldsymbol{F}_{\text{XOR}}$. In addition to the matrix $\boldsymbol{A}$, define $\boldsymbol{A}' = \boldsymbol{A}'(\theta)$ by adding θn $(0 \leq \theta \leq 1)$ new rows, each with exactly a single one, at the bottom of $\boldsymbol{A}$. Equivalently, the Tanner graph $\boldsymbol{G}'$ of $\boldsymbol{A}'$ is obtained by adding $\text{Po}(\lambda n)$ unary check nodes to the Tanner graph $\boldsymbol{G}$ of $\boldsymbol{A}$ where $\lambda = -\log(1-\theta)$. This process is called *'Pinning'* which helps to remove mostly *'short linear relations'*. Below Figure 6.1 gives a rough sketch of the matrix $\boldsymbol{A}$ and $\boldsymbol{A}'$ with respect to the tanner graph $\boldsymbol{G}$ and $\boldsymbol{G}'$. So the second result of this thesis [22] addresses two questions.

Question 1. How does the solution space geometry will change when 'pinning' occurs?

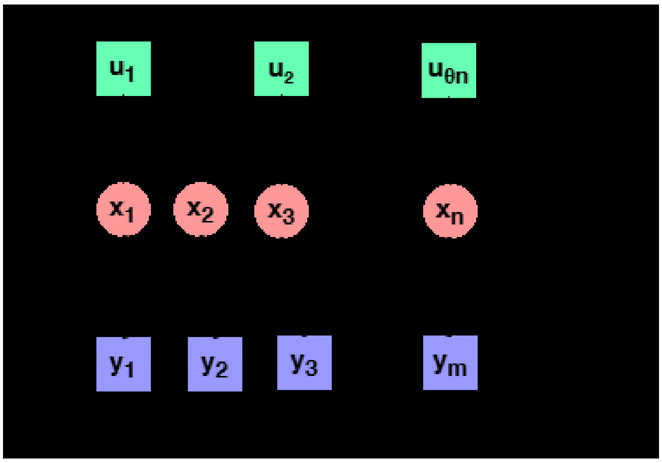


Figure 6.1: Matrix $\boldsymbol{A}$ and $\boldsymbol{A}'$ corresponds to the Tanner graph $\boldsymbol{G}$ and $\boldsymbol{G}'$

Question 2. Establish a link between the performance of the BPGD algorithm and phase transition in decimation process.

6.2.1 Analysis of BPGD

In order to state the main results we need to introduce a few threshold values. To this end, given d, k and a real parameter $\lambda \geq 0$, consider the probability generating functions of $\boldsymbol{D}$ corresponds to the variable node and treat as a Poisson random variable and $\boldsymbol{K}$ corresponds to the check nodes and treat as a two-point distribution, either k or 1 (as the check nodes contains two types of nodes, one with degree k and another is the unit clause.) and is given by,

$$D(z) = \exp((\lambda + d)(z-1))$$
$$K(z) = \frac{k\lambda}{k\lambda + d} z + \frac{d}{k\lambda + d} z^k. \tag{6.2.1}$$

Definition 6.2.1. *The Bethe free entropy* $\boldsymbol{\Phi}$ *of the matrix* $\boldsymbol{A}'$ *is defined by*

$$\boldsymbol{\Phi}(z) = D\Big(1 - \frac{K'(z)}{K'(1)}\Big) - \frac{D'(1)}{K'(1)}\Big(1 - K(z) - (1-z)K'(z)\Big).$$

Also, consider a function ϕ:

$$\phi(z) = 1 - D'\left(1 - K'(z)/K'(1)\right)/D'(1).$$

Remark 6.2.2.

- $\boldsymbol{\Phi}'(z) = D'(1)K''(z)(\phi(z) - z)/K'(1)$.
- *So, the stationary points of* $\boldsymbol{\Phi}$ *coincide with the fixed points of* ϕ *which is verified in [22].*

Substituting for the specific distributions from (6.2.1) we get the following expressions for $\phi(z)$ and $\mathbf{\Phi}(z)$:

$$\phi_{d,k,\lambda}:[0,1]\to[0,1], \qquad z\mapsto 1-\exp\left(-\lambda-dz^{k-1}\right), \tag{6.2.2}$$

$$\mathbf{\Phi}_{d,k,\lambda}:[0,1]\to\mathbb{R}, \qquad z\mapsto \exp\left(-\lambda-dz^{k-1}\right)-\frac{d(k-1)}{k}z^k+dz^{k-1}-\frac{d}{k}. \tag{6.2.3}$$

Let $\alpha_*(\lambda)=\alpha_*(d,k,\lambda)\in[0,1]$ be the smallest and $\alpha^*(\lambda)=\alpha^*(d,k,\lambda)\geq\alpha_*(d,k,\lambda)\in[0,1]$ the largest fixed point of $\phi_{d,k,\lambda}$. Figure 6.2 visualizes $\mathbf{\Phi}_{d,k,\lambda}(z)$ for different values of λ.

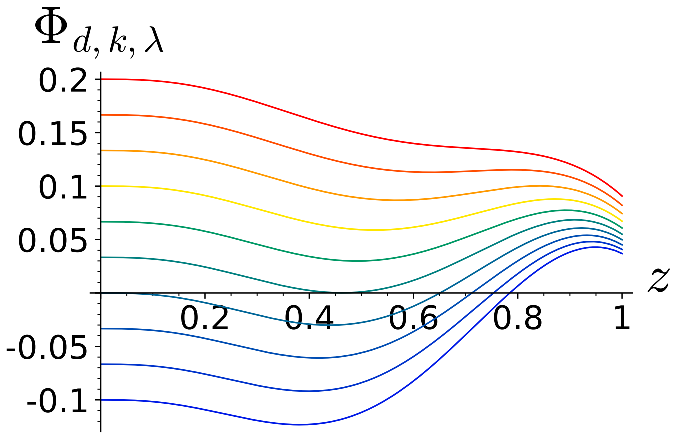


Figure 6.2: $\mathbf{\Phi}_{d,k,\lambda}$ for $k=3$ and $d=2.4$, for λ from 0 to 0.3 (maximum at $z=0$) and from 0.4 to 0.9 (Figure 1, [22])

In addition to this, define few threshold values of d.

$$d_{\min}(k)=\left(\frac{k-1}{k-2}\right)^{k-2}, \tag{6.2.4}$$

$$d_{\text{core}}(k)=\sup\left\{d>0:\alpha^*(0)=0\right\}, \tag{6.2.5}$$

$$d_{\text{SAT}}(k)=\sup\left\{d>0:\Phi_{d,k,0}(\alpha^*(0))\leq\Phi_{d,k,0}(0)\right\}. \tag{6.2.6}$$

where, the value $d_{\text{SAT}}(k)$ is the random k-XORSAT satisfiability threshold [12, 42, 93] and $d_{\text{core}}(k)$ equals the threshold for the emergence of a giant 2-core within the k-uniform hypergraph induced by $\mathbf{\Phi}$ [12, 75]. A bit of calculus reveals that

$$0<d_{\min}(k)<d_{\text{core}}(k)<d_{\text{SAT}}(k)<k.$$

Now we state our second result of this thesis [22].

Theorem 6.2.3 (Theorem 1.1, [22])**.** *Let $k\geq3$.*

(i). If $d < d_{\min}(k)$, then

$$\lim_{n\to\infty} \mathbb{P}\left[\texttt{BPGD}(\boldsymbol{F}_{XOR}) \textit{ finds a satisfying assignment}\right] = \exp\left(-\frac{d^2(k-1)^2}{4}\int_0^1 \frac{z^{2k-4}(1-z)}{1-d(k-1)z^{k-2}(1-z)}\,\mathrm{d}z\right).$$

(ii). If $d_{\min}(k) < d < d_{\mathrm{SAT}}(k)$, then

$$\mathbb{P}\left[\texttt{BPGD}(\boldsymbol{F}_{XOR}) \textit{ finds a satisfying assignment}\right] = o(1).$$

The above theorem determines the precise clause-to-variable densities where BPGD succeeds/fails and mathematically verified the heuristic work by Ricci-Tersenghi and Semerjian [96]. To be precise, in the 'successful' regime BPGD does not actually succeed with high probability, but with an explicit prob- ability strictly between zero and one. The most significant ingredient towards turning the heuristic arguments from [96] into a rigorous proof is a formula for the nullity of the check matrix of the XORSAT instance $\boldsymbol{F}_{\mathrm{DC},t}$ from the decimation process introduced in Chapter 3. The following proposition establishes a relationship between the matrix $\boldsymbol{A}_t = A_{\boldsymbol{F}_{\mathrm{DC},t}}$ and the function $\boldsymbol{\Phi}_{d,k,\lambda}$ for the pre-defined d and λ.

Proposition 6.2.4 (Proposition 2.6, [22])**.**

$$\lim_{n\to\infty} \mathrm{nul}\boldsymbol{A}_t = \Phi_{d,k,\lambda}(\alpha_{\max}) \qquad \textit{in probability.}$$

6.2.2 Phase Transition of Decimation process

In addition to the success probability of BPGD algorithm, we also mathematically confirm the predictions of phase transition heuristically introduced by Ricci-Tersenghi and Semerjian [96] and investigate how they relate to the performance of BPGD. The next two theorems identify precise regime of d,t where different phase transitions of the decimation process hold. Before going to the statement of the theorems let us introduce few values of λ. [accordingly the θ values follow from $\lambda = -\log(1-\theta)$.

$$\lambda_* = \lambda_*(d,k) = -\log(1-z_*) - \frac{z_*}{(k-1)(1-z_*)} > \lambda^*$$
$$\text{where, } \lambda^* = \lambda^*(d,k) = \max\left\{0, -\log(1-z^*) - \frac{z^*}{(k-1)(1-z^*)}\right\} \geq 0$$

Additionally, let $\lambda_{\mathrm{cond}}(d,k)$ be the solution to the ODE

$$\frac{\partial\lambda_{\mathrm{cond}}(d,k)}{\partial d} = -\frac{\alpha^*(\lambda_{\mathrm{cond}}(d,k))^k - \alpha_*(\lambda_{\mathrm{cond}}(d,k))^k}{k(\alpha^*(\lambda_{\mathrm{cond}}(d,k)) - \alpha_*(\lambda_{\mathrm{cond}}(d,k)))}, \qquad \lambda_{\mathrm{cond}}(d_{\mathrm{SAT}}(k),k) = 0 \tag{6.2.7}$$

To be precise, while θ_{cond} matches the predictions of [96], the ODE formula (6.2.7) for the threshold, which is easy to evaluate numerically, does not appear in [96]. Instead of the ODE formulation,

[96] define λ_{cond} as the (unique) $\lambda \geq 0$ such that $\mathbf{\Phi}(\alpha_*) = \mathbf{\Phi}(\alpha^*)$; (we showed in [22] that both are equivalent.)

Theorem 6.2.5 (Theorem 1.2, [22])**.** *Let $k \geq 3$ and let $0 \leq t = t(n) \leq n$ be a sequence such that $\lim_{n\to\infty} t/n = \theta \in (0,1)$.*

(i). If $d < d_{\min}(k)$, then $\boldsymbol{F}_{\mathrm{DC},t}$ has the non-reconstruction property w.h.p.

(ii). If $d_{\min}(k) < d < d_{\mathrm{SAT}}(k)$ and $\theta < \theta^*$ or $\theta > \theta_{\mathrm{cond}}$, then $\boldsymbol{F}_{\mathrm{DC},t}$ has the non-reconstruction property w.h.p.

(iii). If $d_{\min}(k) < d < d_{\mathrm{SAT}}(k)$ and $\theta^* < \theta < \theta_{\mathrm{cond}}$, then $\boldsymbol{F}_{\mathrm{DC},t}$ has the reconstruction property w.h.p.

Recall $\mu_{\boldsymbol{F}_{\mathrm{DC},t}}$ denote the BP 'approximation' of the correct marginal $\pi_{\boldsymbol{F}_{\mathrm{DC},t}}$ of variable x_{t+1} in the formula $\boldsymbol{F}_{\mathrm{DC},t}$ created by the decimation process.

Theorem 6.2.6 (Theorem 1.3, [22])**.** *Let $k \geq 3$ and let $0 \leq t = t(n) \leq n$ be a sequence such that $\lim_{n\to\infty} t/n = \theta \in (0,1)$.*

(i). If $0 < d < d_{\min}(k)$ then $\mu_{\boldsymbol{F}_{\mathrm{DC},t}} = \pi_{\boldsymbol{F}_{\mathrm{DC},t}}$ w.h.p.

(ii). If $d_{\min}(k) < d < d_{\mathrm{SAT}}(k)$ and $\theta < \theta_{\mathrm{cond}}$ or $\theta > \theta_*$, then $\mu_{\boldsymbol{F}_{\mathrm{DC},t}} = \pi_{\boldsymbol{F}_{\mathrm{DC},t}}$ w.h.p.

(iii). If $d_{\min}(k) < d < d_{\mathrm{SAT}}(k)$ and $\theta_{\mathrm{cond}} < \theta < \theta_*$, then $\mathbb{E}\left|\mu_{\boldsymbol{F}_{\mathrm{DC},t}} - \pi_{\boldsymbol{F}_{\mathrm{DC},t}}\right| = \Omega(1)$.

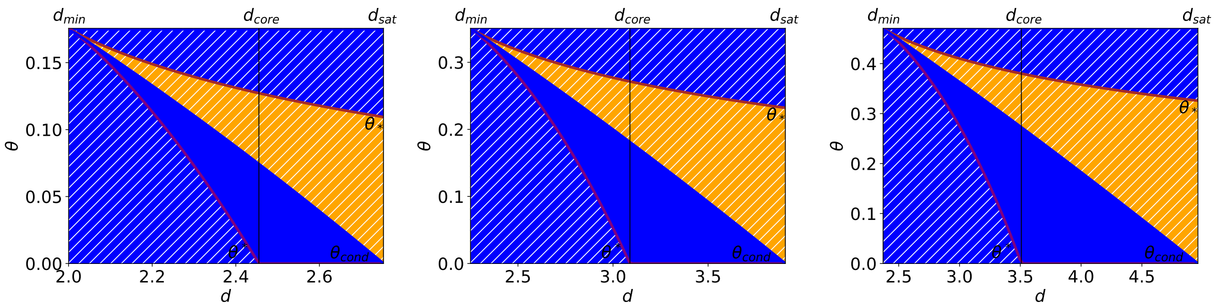
The upshot of Theorems 6.2.5–6.2.6 is that the relation between the accuracy of BP and reconstruction is subtle.

Remark 6.2.7. *As long as $d < d_{\min}$ non-reconstruction holds throughout and the BP approximations are correct. But if $d_{\min} < d < d_{\mathrm{SAT}}$ and $\theta^* < \theta < \theta_{\mathrm{cond}}$, then Theorem 6.2.5 (iii) shows that reconstruction occurs. Nonetheless, Theorem 6.2.6 (ii) demonstrates that the BP approximations remain valid in this regime. By contrast, for $\theta_{\mathrm{cond}} < \theta < \theta_*$ we have non-reconstruction by Theorem 6.2.5 (iii), but Theorem 6.2.6 (iii) shows that BP misses its mark with a non-vanishing probability. Finally, for $\theta > \theta_*$ everything is in order once again as BP regains its footing and non-reconstruction holds. Unfortunately* `BPGD` *is unlikely to reach this happy state because the algorithm is bound to make numerous mistakes at times $t/n \in (\theta_{\mathrm{cond}}, \theta_*)$.*

Figure 6.3 illustrates Theorems 6.2.5–6.2.6, displays the phase diagram in terms of d and $\theta \sim t/n$ for $k = 3,4,5$.

Figure 6.3 description:

- **Hatched area:** displays the regime $\theta < \theta^*$ and $\theta_{\mathrm{cond}} < \theta$ where non reconstruction holds.
- **Non-Hatched area:** displays the regime $\theta^* < \theta < \theta_{\mathrm{cond}}$ where we have reconstruction.
- **Blue area:** displays $\theta < \theta_{\mathrm{cond}}$ and $\theta > \theta_*$ where BP is correct.
- **Orange area:** BP is inaccurate.



(a) $k = 3$ (b) $k = 4$ (c) $k = 5$

Figure 6.3: The phase diagrams for $k = 3, 4, 5$ with $d \in (d_{\min}, d_{\text{SAT}})$ on the horizontal and θ on the vertical axis (Figure 3, [22]).

6.3 Proof Strategy.

Thanks to the half integrality of the messages introduced in [Chapter 3, Fact 3.2.2], BP is equivalent to Warning Propagation in random k-XORSAT. Theorem 6.2.5–6.2.6 rely on the count of null variables in the WP algorithm. Recall $\omega_{F,x\to a}, \omega_{F,a\to x}, \omega_{F,x} \in \{\mathtt{f}, \mathtt{u}, \mathtt{n}\}$ be the WP limits from Chapter 3. Furthermore, let $V_{\mathtt{f},\ell}(F)$, $V_{\mathtt{u},\ell}(F)$, $V_{\mathtt{n},\ell}(F)$ be the sets of variables with the respective mark after $\ell \geq 0$ iterations and $V_{\mathtt{f}}(F), V_{\mathtt{u}}(F), V_{\mathtt{n}}(F)$ be the sets of variables where the limit $\omega_{F,x}$ takes the respective value.

The following statement traces WP on the random formula $\boldsymbol{F}_{\text{DC},t}$ produced by the decimation process.

Proposition 6.3.1 (Proposition 2.5, [22])**.** *Let $\varepsilon > 0$ and assume that $d > 0$, $t = t(n) \sim \theta n$ satisfy one of the following conditions:*

(i). $d < d_{\min}$, or

(ii). $d > d_{\min}$ and $\theta \notin \{\theta_*, \theta^*\}$.

Then there exists $\ell_0 = \ell_0(d, \theta, \varepsilon) > 0$ such that for any fixed $\ell \geq \ell_0$ with $\lambda = -\log(1-\theta)$ w.h.p. we have

$$\left|t + |V_{\mathtt{n},\ell}(\boldsymbol{F}_{\text{DC},t})| - \alpha_* n\right| < \varepsilon n, \quad \left|t + |V_{\mathtt{f},\ell}(\boldsymbol{F}_{\text{DC},t})| - (\alpha^* - \alpha_*) n\right| < \varepsilon n, \quad \left|V_{\mathtt{n}}(\boldsymbol{F}_{\text{DC},t}) \triangle V_{\mathtt{n},\ell}(\boldsymbol{F}_{\text{DC},t})\right| < \varepsilon n.$$

Along with the above proposition, in order to investigate the accuracy of BP it suffices to compare the *numbers* of variables marked $\mathtt{n}$ by WP with the true marginals. The following corollary summarizes the result.

Corollary 6.3.2 (Corollary 2.9, [22])**.** *For any d, θ the following statements are true.*

(i). If $d < d_{\min}$, or *$d > d_{\min}$ and $\theta < \theta_{\text{cond}}$,* or *$d > d_{\min}$ and $\theta > \theta_*$, then $|V_0(\boldsymbol{F}_{\text{DC},t}) \triangle V_{\mathtt{n}}(\boldsymbol{F}_{\text{DC},t})| = o(n)$ w.h.p.*

(ii). If $d > d_{\min}$ and $\theta_{\text{cond}} < \theta < \theta_*$, then $|V_0(\boldsymbol{F}_{\text{DC},t}) \triangle V_{\mathtt{n}}(\boldsymbol{F}_{\text{DC},t})| = \Omega(n)$ w.h.p.

The Corollary 6.3.2 directly implies Theorem 6.2.6 which in turn implies Theorem 6.2.3 (ii). For the (non-)reconstruction thresholds in Theorem 6.2.5 we need to investigate the *conditional* marginals given the values of variables at a certain distances from x_{t+1} as in the (non)-reconstruction property defined in Chapter 4. This is where the extra value f from the construction of WP enters.

Corollary 6.3.3 (Corollary 2.10, [22]). *Assume that $d > d_{\min}$ and let $\varepsilon > 0$.*

(i). If $\theta < \theta_{\text{cond}}$, then for any fixed ℓ we have $|V_{\mathtt{f},\ell}(\boldsymbol{F}_{\text{DC},t}) \cap V_{0,\ell}$[1] $(\boldsymbol{F}_{\text{DC},t})| < \varepsilon n$ w.h.p.

(ii). If $\theta > \theta_{\text{cond}}$, then there exists $\ell_0 = \ell_0(d,\theta,\varepsilon)$ such that for any fixed $\ell > \ell_0$ we have

$$|(V_{\mathtt{n},\ell}(\boldsymbol{F}_{\text{DC},t}) \cup V_{\mathtt{f},\ell}(\boldsymbol{F}_{\text{DC},t})) \triangle V_{0,\ell}(\boldsymbol{F}_{\text{DC},t})| < \varepsilon n \qquad w.h.p.$$

Comparing the number of actually frozen variables with the ones marked f by WP, we obtain Theorem 6.2.5.

Coming to the proof of the success probability of BPGD, the Corollary 6.3.3 directly implies that the BP approximations of the marginals are mostly correct for $d < d_{\min}$ on the formula $\boldsymbol{F}_{\text{DC},t}$ obtained by the decimation process. The difficulty in analyzing BPGD lies in proving that the estimates of the algorithm are not just mostly correct, but correct up to only a bounded expected number of discrepancies over the entire execution of the algorithm. To prove this fact we combine the method of differential equations with a precise analysis of the sources of the remaining bounded number of discrepancies which comes from the presence of short (i.e., bounded-length) cycles (we call this as 'toxic cycles') in the graph $G(\boldsymbol{F})$. (more details of the proof can be found in [22]).

Again due to the half-integrality fact 3.2.2 on random k-XORSAT, we know that BPGD boils down to the pure combinatorial algorithm called *'Unit Clause Propagation'* (UCP) (pseudocode of the UCP algorithm can be found in Chapter 3). We conclude this chapter by stating the following proposition which can be verified easily (details can be found in [22]).

Proposition 6.3.4 (Proposition 6.1, [22]). *We have,*

$$\mathbb{P}\left[\texttt{BPGD}\ \textit{outputs a satisfying assignment of}\ \boldsymbol{F}_{XOR}\right] = \mathbb{P}\left[\texttt{UCP}\ \textit{outputs a satisfying assignment of}\ \boldsymbol{F}_{XOR}\right].$$

This proposition implies that the success probability of BPGD established in Theorem 6.2.3 is equivalent to that of the UCP algorithm. So, the second result of the thesis establishes a sharper bound on the clause length for $k \geq 3$. Depending on the regime of d, both the BPGD and UCP algorithms may succeed or fail, thereby substantially improved over Yung's result [102] which shows the lower bounds on the clause length ($k \geq 9$ for UCP and $k \geq 13$ for BPGD).

[1] $V_{0,\ell}(F)$ be the set of variables x_i such that $\sigma_i = 0$ for all $\sigma \in \ker A_F$ for which $\sigma_h = 0$ for all variables $x_h \in \partial^\ell x_i$

Based on: ***The random k-SAT Gibbs Uniqueness Threshold revisited*** *[21]*
Arnab Chatterjee, Amin Coja-Oghlan, Catherine Greenhill,
Vincent Pfenninger, Maurice Rolvien, Pavel Zakharov, Konstantinos Zampetakis
arXiv:2506.01359 (2025)

7
On the Gibbs Uniqueness in random k-SAT

"··· the theory of Gibbs measures, which provides a very effective and flexible way to define collections of "locally dependent" random variables."

–Amir Dembo et.al.

Unlike in previous chapter where we have investigated the random k-XORSAT problem, analyzing the performance of BPGD and its connection to the decimation process associated with different phase transitions, we now turn to random k-SAT, one of the most extensively studied random constraint satisfaction problems. Beyond its central role in computational complexity, it provides a natural framework for exploring the fundamental phenomena such as the number of satisfying assignments, clustering, reconstruction and Gibbs uniqueness phase transition. In this chapter our focus is to establish rigorous lower bounds on the number of satisfying assignments of random k-SAT upto the Gibbs uniqueness threshold inspired from the statistical physics inspired mechanism *'replica symmetric solution'* [77, 78].

7.1 Motivation and History

Since the time of 1990s, pinpointing the satisfiability threshold on random k-SAT, defined as the largest clause to variable density upto which satisfying assignments exist [6], has been a guiding theme of research in the area of random CSPs. For every $k \geq 3$, indeed the physics inspired '*cavity method*' predicts the satisfiability threshold exactly [72] but for 'small' $k \geq 3$ in random k-SAT it is hard nut to crack. In statistical physics, one of the most important quantity is to determine the exact number of satisfying assignments (also known as '*partition function*' in physics jargon) and then the satisfiability threshold. More recently three prior contributions stand out to prove the physics prediction correctly. Firstly, in [90] **Panchenko** and **Talagrand** proved a rigorous upper bound on the physics formula using a proof technique called '*interpolation method*'. Secondly, **Achlioptas et al.** [2] proved the physics formula in the case $k = 2$ which is conceptually easier than $k \geq 3$. **Montanari** and **Shah** in [83] provided a correct approximation on the number of 'good' satisfying assignment all but $o(n)$ clauses for all $k \geq 3$. Our paper verifies the correct number of satisfying assignments given by physics method '*replica symmetry solution*' for any $k \geq 3$ upto Gibbs uniqueness threshold. Moreover we derive a lower bound on the Gibbs uniqueness threshold which improves significantly over the work of **Montanari** and **Shah** in [83] for small $k \geq 3$.

7.2 Main Results.

In this section we state our third result of this thesis [21].

Let $\boldsymbol{F}_{k\text{SAT}} = \boldsymbol{F}_{d,k}(n)$ be the random k-CNF on n Boolean variables $x_1, \ldots, x_n$ with $\boldsymbol{m} \overset{\text{d}}{=} \text{Po}(dn/k)$ clauses $a_1, \ldots, a_{\boldsymbol{m}}$. Similarly like in random 2-SAT(Chapter 5) and in random k-XORSAT (Chapter 6) the clauses a_i are drawn independently and uniformly from the set of all $2^k\binom{n}{k}$ possible clauses with k distinct variables. The parameter d prescribes the expected number of clauses associated with a given variable appears to it. Also, let $S(\boldsymbol{F}_{k\text{SAT}})$ be the set of satisfying assignments of $\boldsymbol{F}_{k\text{SAT}}$ and let $Z(\boldsymbol{F}_{k\text{SAT}}) = |S(\boldsymbol{F}_{k\text{SAT}})|$ i.e., the number of satisfying assignments of the random k-SAT formula $\boldsymbol{F}_{k\text{SAT}}$. The aim of this chapter are twofold:

(i). To study the logarithm of the number of satisfying assignments of random k-SAT (in statistical physics it is called 'partition function' as defined in Chapter 2) i.e., the quantity $\frac{1}{n} \log Z(\boldsymbol{F}_{k\text{SAT}})$ as $n \to \infty$, which is given by the prediction of *'replica symmetry solution'* in terms of *'Bethe Free entropy'* which is a function defined for a probability measure $\pi \in \mathscr{P}(0,1)$.

(ii). When the above quantity is well-defined we further rigorously analyzed the lower bound obtained for the Gibbs uniqueness threshold for any $k \geq 3$ which is significantly improved for small k values over the work in [83].

7.2.1 Limit in probability of log-partition function in random k-SAT

Along with the Gibbs uniqueness property defined in Chapter 4 as a final preparation we need to illuminate the 'replica symmetric solution' from [77, 78]. This prediction comes in terms of a fixed point problem on the space $\mathscr{P}(0,1)$ of probability measures on the open unit interval. Consider the *Belief Propagation operator*

$$\text{BP}_{d,k} : \mathscr{P}(0,1) \to \mathscr{P}(0,1), \qquad \pi \mapsto \hat{\pi} = \text{BP}_{d,k}(\pi) \tag{7.2.1}$$

defined as follows. Let $\boldsymbol{d}^+, \boldsymbol{d}^- \stackrel{\mathrm{d}}{=} \mathrm{Po}(d/2)$ be Poisson variables with expectation $d/2$. Moreover, let $(\boldsymbol{\mu}_{\pi,i,j})_{i,j\geq 1}$ be a sequence of i.i.d. random variables, each following distribution π. All these random variables are mutually independent. Further, let

$$\boldsymbol{\mu}_{\pi,i} = 1 - \prod_{j=1}^{k-1} \boldsymbol{\mu}_{\pi,i,j} \quad \text{for } i \geq 1, \qquad \text{and} \qquad \hat{\boldsymbol{\mu}}_\pi = \frac{\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\pi,2i-1}}{\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\pi,2i-1} + \prod_{i=1}^{\boldsymbol{d}^+} \boldsymbol{\mu}_{\pi,2i}}. \tag{7.2.2}$$

Then $\hat{\pi}$ is the distribution of $\hat{\boldsymbol{\mu}}_\pi$. Furthermore, define the *Bethe free entropy*

$$\mathfrak{B}_{d,k}(\pi) = \mathbb{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\pi,2i} + \prod_{i=1}^{\boldsymbol{d}^+} \boldsymbol{\mu}_{\pi,2i-1}\right) - \frac{d(k-1)}{k} \log\left(1 - \prod_{j=1}^{k} \boldsymbol{\mu}_{\pi,1,j}\right)\right], \tag{7.2.3}$$

provided that the expectation on the r.h.s. exists.

Theorem 7.2.1 (Theorem 1.1, [21])**.** *Let $k \geq 3$ and assume that $0 < d < d_{\mathrm{uniq}}(k)$. Then the weak limit*

$$\pi_{d,k} = \lim_{\ell\to\infty} \mathrm{BP}_{d,k}^{\ell}(\delta_{1/2}) \in \mathscr{P}(0,1) \tag{7.2.4}$$

exists and

$$\lim_{n\to\infty} \frac{1}{n} \log Z(\boldsymbol{\Phi}) = \mathfrak{B}_{d,k}(\pi_{d,k}) \qquad \textit{in probability.} \tag{7.2.5}$$

where, $\mathrm{BP}_{d,k}^{\ell}$ is the ℓ-fold application of the operator $\mathrm{BP}_{d,k}$ and $\delta_{1/2} \in \mathscr{P}(0,1)$ be the atom at $1/2$. Although the formula (7.2.5) is not explicit, but the proof the Theorem 7.2.1 reveals that the convergence of the weak limit $\pi_{d,k}$ occurs rapidly. In the next subsection we will introduce few threshold values of d and provide a improved lower bound on the number of solutions of random k-SAT for any $k \geq 3$.

7.2.2 Lower bound on Gibbs uniqueness

Before we dive into the second result of this chapter we first introduce few known and our threshold values of d corresponds to the number of satisfying assignments of random k-SAT. From the title of this chapter one natural question arises: *How can we determine the Gibbs uniqueness threshold d_{uniq} in random k-SAT?*

The best known current result for d_{uniq} is in the case of $k = 2$ which coincides with the value d_{SAT} for random 2-SAT. As the precise value of d_{uniq} is not known for $k \geq 3$, Montanari and Shah in [83] proved that this value is upper bounded by the pure literal threshold d_{pure} defined in [19, 74]:

$$d_{\mathrm{uniq}}(k) \leq d_{\mathrm{pure}}(k) = \min_{z>0} \frac{z}{(1-\exp(-z/2))^{k-1}}. \tag{7.2.6}$$

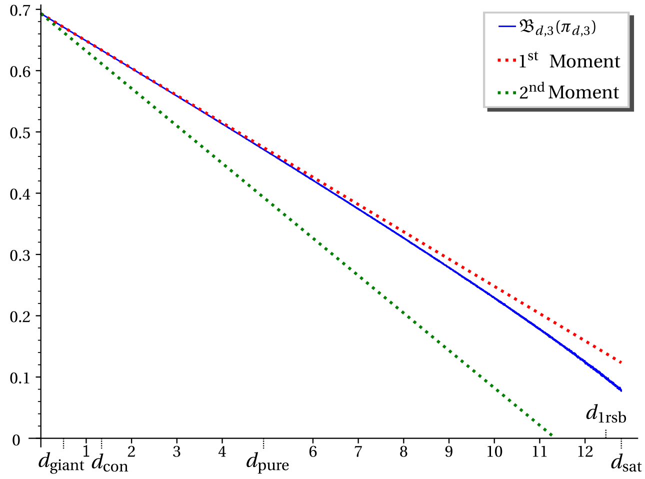


Figure 7.1: Comparison of $\mathfrak{B}_{d,k}(\pi_{d,k})$ with known bounds for $\lim_{n\to\infty} \frac{1}{n} \log Z(\mathbf{\Phi})$ for $k = 3$. [21]

Complementing the upper bound (7.2.6), Montanari and Shah derived a lower bound $d_{\text{MS}}(k)$:

$$d_{\text{MS}}(k) = \sup\left\{d > 0 : d(k-1)\left(1 - \exp(-d/2)/4\right)\left(1 - \exp(-d/2)/2\right)^{k-2} < 1\right\} \leq d_{\text{uniq}}(k). \tag{7.2.7}$$

But unfortunately, this bound is not tight even for $d = 2$. Along this their bound only yields the number of 'good' assignments satisfying all but $o(n)$ clauses, rather than of actual satisfying assignments. In the following theorem we derived a new lower bound d_{our} on the Gibbs uniqueness threshold d_{uniq} on the number of actual satisfying assignments of random k-SAT.

Theorem 7.2.2 (Theorem 1.2, [21])**.** *For all $k \geq 3$ we have*

$$d_{\text{uniq}}(k) \geq d_{\text{our}}(k) := \sup\left\{d > 0 : \frac{d(k-1)}{2}\left(1 - \exp(-d/2)/2\right)^{k-2} < 1\right\}. \tag{7.2.8}$$

An easy calculation exposes that for every $k \geq 2$,

$$d_{\text{MS}}(k) < d_{\text{our}}(k)$$

Beside that, the best prior rigorous bounds on the number of satisfying assignments beyond the giant component threshold $d_{\text{giant}} = 1/(k-1)$ from the first and second moment methods. The first moment bound reads

$$\frac{1}{n}\log Z(\boldsymbol{F}_{k\text{SAT}}) \leq \log 2 + \frac{d}{k}\log(1 - 2^{-k}) + o(1) \quad \text{w.h.p.} \tag{7.2.9}$$

Moreover, Achlioptas and Peres [7] perform a second moment argument on the number of satisfying assignments that enjoy a peculiar additional condition required to keep the second moment under control. They show that w.h.p.

$$\frac{1}{n}\log Z(\boldsymbol{F}_{k\mathrm{SAT}}) \geq (1-d)\log 2 + \frac{d}{k}\log\left[\left(\lambda^{1/2}+\lambda^{-1/2}\right)^k - \lambda^{-k/2}\right] + o(1), \tag{7.2.10}$$

$$\text{where}(1-\lambda)(1+\lambda)^{k-1} = 1, \lambda > 0. \tag{7.2.11}$$

Figure 7.1 illustrates the bounds (7.2.9)–(7.2.10) along with (7.2.5) for $k = 3$.

Figure 7.1 description:

- The red dotted line depicts the first moment upper bound (7.2.9).
- The green dotted line represents the lower bound provided by (7.2.10).
- The blue line displays a numerical approximation of $\mathfrak{B}_{d,3}(\pi_{d,3})$. To obtain our values, we generated 10^6 samples from $\pi \approx \mathrm{BP}^{25}_{d,3}(\delta_{1/2})$ and then evaluated the corresponding empirical average of the expression in (7.2.3).

Finally combining Theorem 7.2.1–Theorem 7.2.2 we obtain the following corollary:

Corollary 7.2.3 (Corollary 1.3, [21])**.** *For $k \geq 3$ and $d < d_{\mathrm{our}}$ the following holds from* (7.2.5)

$$\lim_{n\to\infty}\frac{1}{n}\log Z(\boldsymbol{\Phi}) = \mathfrak{B}_{d,k}(\pi_{d,k})$$

7.3 Proof Strategy

The proof of the two results of this chapter comprises several steps discussing below:

7.3.1 Existence of fixed point.

The existence of the limit $\pi_{d,k}$ is an easy consequence of the Gibbs uniqueness property discussed in Chapter 4 for every $d < d_{\mathrm{uniq}}$ and the limit $\pi_{d,k} = \lim_{\ell\to\infty}\mathrm{BP}^{\ell}_{d,k}(\delta_{1/2})$ is a fixed point of the Belief Propagation operator from 7.2.1. The following proposition implies that the limit defined in (7.2.4) exists with respect to W_1-metric.

Proposition 7.3.1 (Proposition 2.1, [21])**.** *The W_1-limit $\pi_{d,k} = \lim_{\ell\to\infty}\mathrm{BP}^{\ell}_{d,k}(\delta_{1/2})$ exists and*

$$\mathbb{E}\left[\log^2 \boldsymbol{\mu}_{\pi_{d,k},1,1}\right] + \mathbb{E}\left|\log\left(\prod_{i=1}^{\boldsymbol{d}^-}\boldsymbol{\mu}_{\pi_{d,k},2i} + \prod_{i=1}^{\boldsymbol{d}^+}\boldsymbol{\mu}_{\pi_{d,k},2i-1}\right)\right| + \mathbb{E}\left|\log\left(1-\prod_{j=1}^{k}\boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right| < \infty. \tag{7.3.1}$$

In addition, $\boldsymbol{\mu}_{\pi_{d,k},1,1}$ and $1-\boldsymbol{\mu}_{\pi_{d,k},1,1}$ are identically distributed.

Along with the fixed point existence below we will discuss the upper and lower bound of the value of $\log Z(\boldsymbol{F}_{k\mathrm{SAT}})$ using '*interpolation method*' and '*Aizenmann-Sims-Starr scheme*' respectively towards the proof of our result in the regime $d < d_{\mathrm{uniq}}$.

7.3.2 Interpolation method: matching upper bound

In mathematical physics as well as in random constraint satisfaction problems, there are several literature [31, 32, 90] deals with the interpolation method to provide a matching upper bound on the normalized log-partition function when $n \to \infty$. The basic idea is to construct a family of random CSPs, parameterized by $t \in [0,1]$ which coincides with the final limiting random graph $\hat{\boldsymbol{G}}$ of interest while for t =0 the CSP is so simple that we can calculate the partition function easily.

Setup

Define a family of interpolating CSPs $\{G_t\}_{t\in[0,1]}$:

- At $t = 0$: G_0 is a trivial/decoupled CSP, often consisting of independent clauses on single variables.
- At $t = 1$: $G_1 = \hat{\boldsymbol{G}}$, the full random CSP model of interest. The clauses consists of exactly k variables.

Indeed at $t = 0$ the logarithm of the partition function asymptotically equal to $n\mathfrak{B}_{d,k}$. To obtain the matching upper bound on $\mathbb{E}[\log Z(\hat{\boldsymbol{G}})]$ one can show that the mean of the logarithm of the partition function is a monotonically increasing function of t.

In our model, the interpolation method along with Proposition 7.3.1 easily implies that,

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\log(Z(\boldsymbol{\Phi}) \vee 1)\right] \le \mathfrak{B}_{d,k}(\pi_{d,k})$$

Ultimately Theorem 7.2.1 is the direct consequence of the following corollary and the proposition.

Corollary 7.3.2 (Corollary 2.2, [21])**.** *If $d < d_{\text{uniq}}(k)$ then w.h.p. we have*

$$\frac{1}{n}\log Z(\boldsymbol{\Phi}) \le \mathfrak{B}_{d,k}(\pi_{d,k}) + o(1)$$

Proposition 7.3.3 (Proposition 2.3, [21])**.** *If $d < d_{\text{uniq}}(k)$ then*

$$\mathbb{E}\left[\log(Z(\boldsymbol{\Phi}_{d,k}(n+1)) \vee 1)\right] - \mathbb{E}\left[\log(Z(\boldsymbol{\Phi}_{d,k}(n)) \vee 1)\right] = \mathfrak{B}_{d,k}(\pi_{d,k}) + o(1).$$

In order to evaluate the expectation from Proposition 7.3.3 we harness a 'soft' version of the k-SAT problem where violated clauses are discouraged but not strictly forbidden. Define for a real $\beta > 0$:

$$Z_\beta(\boldsymbol{F}_{k\text{SAT}}) = \sum_{\sigma\in\{\pm1\}^{V(\boldsymbol{F}_{k\text{SAT}})}} \prod_{a\in C(\boldsymbol{F}_{k\text{SAT}})} \exp(-\beta \mathbb{1}\{\sigma \not\models a\}). \tag{7.3.2}$$

The above definition of the partition function ensures that $Z_\beta(\boldsymbol{F}_{k\text{SAT}}) \ge Z(\boldsymbol{F}_{k\text{SAT}})$ for all $\beta > 0$. Then by means of interpolating argument we can say the following theorem and lemma.

Theorem 7.3.4 ([90, Theorem 1]). *For any $k \geq 3$, any $\beta > 0$ and any probability measure π on $[0,1]$ we have*

$$\frac{1}{n}\mathbb{E}\left[\log Z_\beta(\boldsymbol{F}_{kSAT})\right] \leq \mathbb{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}^-}\boldsymbol{\mu}_{\beta,\pi,2i}+\prod_{i=1}^{\boldsymbol{d}^+}\boldsymbol{\mu}_{\beta,\pi,2i-1}\right)-\frac{d(k-1)}{k}\log\left(1-\left(1-e^{-\beta}\right)\prod_{j=1}^{k}\boldsymbol{\mu}_{\pi,1,j}\right)\right], \quad (7.3.3)$$

$$\text{where,} \quad \boldsymbol{\mu}_{\beta,\pi,i} = 1-(1-\exp(-\beta))\prod_{j=1}^{k-1}\boldsymbol{\mu}_{\pi,i,j} \qquad (\text{for } i \geq 1).$$

The monotone convergence theorem for the measure $\pi = \pi_{d,k}$ we get the explicit expression of the r.h.s. in (7.3.3). Now the routine application of Azuma-Hoeffding implies to this soft model with $\beta < \infty$ gives the below concentration bound:

Lemma 7.3.5 (Lemma 5.2, [21]). *For any fixed $\beta > 0$ we have $\mathbb{P}\left[\left|\log Z_\beta(\boldsymbol{F}_{kSAT}) - \mathbb{E}\log Z_\beta(\boldsymbol{F}_{kSAT})\right| > \sqrt{n}\log n\right] = o(1/n)$.*

This lemma implies that the clauses of the random formula $\boldsymbol{F}_{k\mathrm{SAT}}$ are drawn independently, and adding or removing a single clause can alter the value of $\log Z_\beta(\cdot)$ by no more than $\pm\beta$. Finally, the Corollary 7.3.2 directly implies from Theorem 7.3.4 and Lemma 7.3.5.

7.3.3 Aizenmann-Sims-Starr: matching lower bound

The key step of this chapter is to establish a lower bound on $\log Z(\boldsymbol{F}_{k\mathrm{SAT}})$ that matches the upper bound from Corollary 7.3.2.

Here, we couple the random k-CNF $\boldsymbol{F}_{k\mathrm{SAT}}(n) = \boldsymbol{F}_{d,k}(n)$ with n variables with the random k-CNF $\boldsymbol{F}_{k\mathrm{SAT}}(n+1) = \boldsymbol{F}_{d,k}(n+1)$ with $n+1$ variables. Most important part of the lower bound proof of Proposition 7.3.3 consists of the coupling argument discussed below along with the necessary tail bound.

CPL1 Let $\boldsymbol{F}'_{k\mathrm{SAT}}$ be a random k-CNF with variables $x_1,\ldots,x_n$ and $\boldsymbol{m}' \stackrel{\mathrm{d}}{=} \mathrm{Po}(d(n-k+1)/k)$ clauses.

CPL2 Obtain $\boldsymbol{F}''_{k\mathrm{SAT}}$ from $\boldsymbol{F}'_{k\mathrm{SAT}}$ by adding another $\boldsymbol{\Delta}'' \stackrel{\mathrm{d}}{=} \mathrm{Po}(d(k-1)/k)$ independent random clauses.

CPL3 Obtain $\boldsymbol{F}'''_{k\mathrm{SAT}}$ from $\boldsymbol{F}'_{k\mathrm{SAT}}$ by adding one new variable x_{n+1} and $\boldsymbol{\Delta}''' \stackrel{\mathrm{d}}{=} \mathrm{Po}(d)$ independent random clauses that each contain x_{n+1} and $k-1$ other variables from $\{x_1,\ldots,x_n\}$.

Figure 7.2 shows a graphical representation of the Aizenmann-Sims-Starr scheme (coupling technique used above). Based on the coupling we have the following fact and the tail bound:

Fact 7.3.6 (Fact 2.4, [21]). *For any $d > 0$ we have $Z(\boldsymbol{F}_{kSAT}(n)) \stackrel{d}{=} Z(\boldsymbol{F}''_{kSAT})$ and $Z(\boldsymbol{F}_{kSAT}(n+1)) \stackrel{d}{=} Z(\boldsymbol{F}'''_{kSAT})$.*

Proposition 7.3.7 (Proposition 2.7, [21]). *For $d < d_{\mathrm{uniq}}(k)$ we have*

$$\mathbb{E}\left[\left|\log\frac{Z(\boldsymbol{F}''_{kSAT})\vee 1}{Z(\boldsymbol{F}'_{kSAT})\vee 1}\right|^{3/2}+\left|\log\frac{Z(\boldsymbol{F}'''_{kSAT})\vee 1}{Z(\boldsymbol{F}'_{kSAT})\vee 1}\right|^{3/2}\right] = O(1). \quad (7.3.4)$$

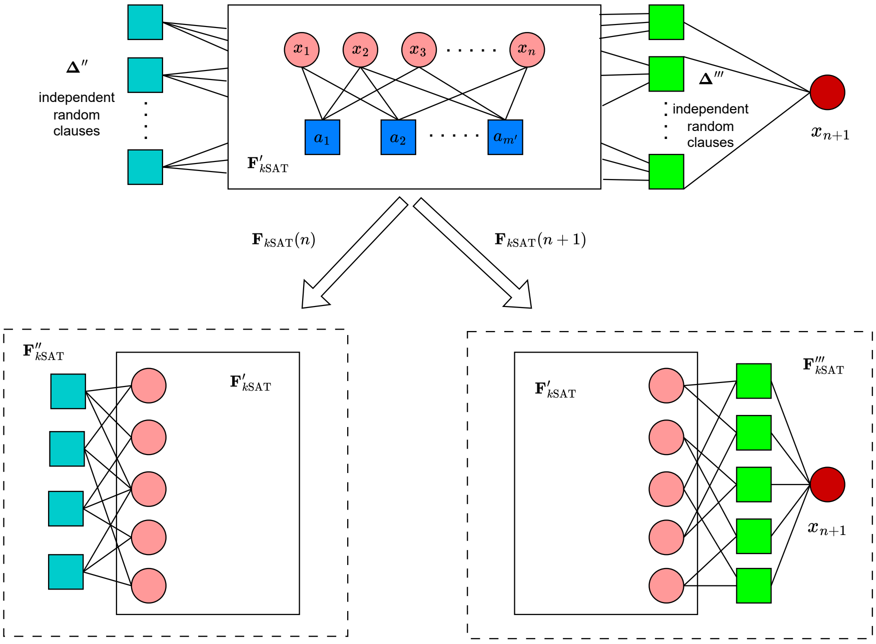


Figure 7.2: A graphical representation of coupling technique (Aizenmann-Sims-Starr scheme)

Finally, towards the proof of Theorem 7.2.1, the existence of the limit comes from Proposition 7.3.1 (for the detailed proof one can refer [21]) and the normalized log-partition function comes from the Aizenmann-Sims-Starr lower bound along with the 'PULP' algorithm we discussed in Chapter 3. Mathematically, the below equation (7.3.5) along with Corollary 7.3.2 implies the result.

$$\begin{aligned}\frac{1}{n}\mathbb{E}\left[\log(1\vee Z(\boldsymbol{F}_{k\mathrm{SAT}}(n)))\right] &= \frac{1}{n}\sum_{N=0}^{n-1}\left(\mathbb{E}\left[\log(1\vee Z(\boldsymbol{F}_{k\mathrm{SAT}}(N+1))\right]-\mathbb{E}\left[\log(1\vee Z(\boldsymbol{F}_{k\mathrm{SAT}}(N))\right]\right)\\ &= \mathfrak{B}_{d,k}(\pi_{d,k})+o(1)\ . \end{aligned} \tag{7.3.5}$$

7.3.4 Lower bound on Gibbs uniqueness threshold

Finally, we are left with the proof of the Gibbs uniqueness threshold lower bound stated in Theorem 7.2.2. An obvious challenge associated with the establishing the Gibbs uniqueness property discussed in Chapter 4 is to estimate the marginal of the root variable given any possible boundary conditions at a distance 2ℓ from the root r. But using the help of [2] we may confine ourselves to just a single, explicit boundary configuration $\boldsymbol{\tau}^+$ that satisfies

$$\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(r)=1\mid\mathbb{T},\forall x\in\partial^{2\ell}r:\boldsymbol{\tau}^{(\ell)}(x)=\boldsymbol{\tau}^+(x)\right]=\max_{\tau\in S(\mathbb{T}^{(\ell)})}\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(r)=1\mid\mathbb{T},\forall x\in\partial^{2\ell}r:\boldsymbol{\tau}^{(\ell)}(x)=\tau(x)\right]. \tag{7.3.6}$$

and define for any variable w at distance $2q$ from $\mathfrak{r}$ with parent clause a and grandparent variable u as

$$\boldsymbol{\tau}^+(w) = \text{sign}(w,a)\cdot\mathbb{1}\{\text{sign}(u,a) \neq \boldsymbol{\tau}^+(u)\} - \text{sign}(w,a)\cdot\mathbb{1}\{\text{sign}(u,a) = \boldsymbol{\tau}^+(u)\} \ . \tag{7.3.7}$$

Then we can say for any integer $\ell > 0$ the assignment $\boldsymbol{\tau}^+$ satisfies (7.3.6). Hence the Theorem 7.2.2 reduces to the following statement.

Proposition 7.3.8 (Proposition 2.13, [21])**.** *For $d < d_{\text{our}}(k)$ we have that*

$$\lim_{\ell\to\infty}\mathbb{E}\left[\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(r) = 1 \mid \mathbb{T}, \forall x \in \partial^{2\ell} r : \boldsymbol{\tau}^{(\ell)}(x) = \boldsymbol{\tau}^+(x)\right] - \mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(r) = 1 \mid \mathbb{T}\right]\right] = 0. \tag{7.3.8}$$

Although it seem delicate because the boundary condition $\boldsymbol{\tau}^+$ depends on the tree $\mathbb{T}^{(\ell)}$. To overcome this problem, we generalize another technique from the work [2] on random 2-SAT to $k \geq 3$ by introducing a quantity that allows us to prove (7.3.8) which behaves 'Markovian' as we pass up and down the tree.

Finally, by combining the mechanism of coupling and contraction with the treatment of both pure and mixed literals (the detailed proof can be found in [21]), we conclude that Theorem 7.2.2 directly follows from the Proposition 7.3.8 together with the triangle inequality.

8

The Last Chapter

"The important thing is not to stop questioning. Curiosity has its own reason for existing."

–Albert Einstein

In this concluding chapter we take a break from formulating and answering questions and instead summarize the main ideas of this thesis and an evaluation of the author's contribution to each paper is addressed. We also discuss some potential future directions and mention some relevant research areas that were out of scope for discussions in the previous chapters and in the papers (found in appendix).

8.1 Summary of the thesis

So far, this thesis has addressed several problems in the probabilistic analysis of random combinatorial structures, more particularly in random constraint satisfaction problems (CSPs). While Chapter 5,6,7 focused on more specific model, together they contribute to a unified understanding of how the combinatorial structure, randomness and statistical physics ideas interact in discrete probability.

Below we will give a very brief overview of our results stated in previous three chapters and a quick comparison with the previous results.

Random 2-SAT: Our first result concerns the random 2-SAT, where we establish a central limit theorem on the number of random 2-SAT solutions which exhibits log-normal fluctuations. This provides a precise distributional description of the solution space size and strengthen the clear picture of probabilistic argument of random satisfiability by analyzing the contraction property of $\mathtt{logBP}^{\otimes}_{d,t}$ which ensures the existence of a unique fixed point $\rho_{d,t}$ and involves the analysis of Belief Propagation on a Galton-Watson tree, which connects the fixed point to the BP marginals.

Finally, we derive our main result from the general martingale central limit theorem , which is a special case of ([55],Theorem 3.2).

Comparison with previous work:

Known results:

- In 1996, **Goerdt** [53] stated that $d = 2$ is the satisfiability threshold of a random 2-SAT formula. In other words, for any $\varepsilon > 0$, the probability of random 2-SAT is satisfiable tends to one if $d < 2-\varepsilon$ and tends to zero if $d > 2+\varepsilon$ as $n \to \infty$.

  Another obvious question is to find out the number of satisfying assignments in the satisfiable regime i.e., when $d < 2$.

- In 2021, **Achlioptas et.al.** [2] provides a first order approximation by stating the normalized partition function i.e., the logarithm of the number of solution (Z) of random 2-SAT formula $\boldsymbol{F}$ w.r.t. n (where n is the number of variables in $\boldsymbol{F}$) converges in probability to a constant (μ_d) which doesn't depend on n. Mathematically,

$$\frac{\log Z(\boldsymbol{F})}{n} \xrightarrow{p} \mu_d$$

Our results:

In this thesis, we say even more about the number of solutions of random 2-SAT. The number of random 2-SAT solutions exhibits fluctuations of order $\sqrt{n}$. More precisely,

$$\frac{\log Z(\boldsymbol{F}) - \mathbb{E}[\log Z(\boldsymbol{F})]}{\eta_d\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1)$$

where $\eta_d \geq 0$ is not depend on n. Along with the asymptotic normality we also evaluated the formula for variance effectively.

Note: By contrast, for other random CSPs the typical fluctuations of the logarithm of the number of solutions are bounded throughout all or most of the satisfiable regime.

Random k-XORSAT: In the context of the second result of this thesis, we analyzed the performance of Belief Propagation Guided Decimation, a statistical physics inspired algorithm on random k-XORSAT and our rigorous analysis mathematically verified the heuristic work of Ricci-Tersenghi and Semerjian [96]. Specifically, we derive an explicit threshold upto which the BPGD algorithm succeeds with a strictly positive probability $\Omega(1)$ and beyond which the algorithm fails with high probability. Due to its algebraic structure, BPGD is equivalent to the purely combinatorial algorithm called 'Unit Clause Propagation'(UCP), so the results work for BPGD should work for UCP as well for the same parameter values. In addition to this we analyze the '*Decimation process*' for which we identify a (non)-reconstruction and condensation phase transition.

Comparison with previous work:

Known results:

- In 2009, **Ricci-Tersenghi** and **Semerjian** [96] provide a heuristic on the satisfiability formula without giving a proof. They just mentioned the success probability of both BPGD and UCP without verified the proof for that.

  $$\mathbb{P}_{\text{succ}} = \exp\left(-\int_0^1 \frac{dt}{4(1-t)} \frac{f(t)^2}{(1-f(t))}\right)$$

  with $f(t) = \alpha k(k-1)t^{k-2}(1-t)$ where α is the clause-to-variable density.

- In a recent paper by **Yung** [102] establishes the '*Overlap Gap paradigm*' which only provides one-sided bound due to the moment computations (in physics jargon called '*annealed*' technique), which implies no positive result of the algorithm.

  For the above reason, his lower bounds on the clause length $k \geq 9$ for UCP and $k \geq 13$ for BPGD are not tight. Moreover according to his argument BPGD and Unit clause both algorithm fail for $d > d_{\text{core}}$.

Our results:

We mathematically verified the heuristic predictions of Ricci-Tersenghi and Semerjian [96] as to the performance of BPGD by providing an explicit formula for the success probability of the algorithm for precise clause-to-variable densities.
For $k \geq 3$, we proof the following:

- If $d < d_{\min}(k)$, then

  $$\lim_{n\to\infty} \mathbb{P}_{\text{succ}} = \exp\left(-\frac{d^2(k-1)^2}{4} \int_0^1 \frac{z^{2k-4}(1-z)}{1-d(k-1)z^{k-2}(1-z)}\,\mathrm{d}z\right).$$

- If $d_{\min}(k) < d < d_{\text{SAT}}(k)$, then

  $$\mathbb{P}_{\text{succ}} = o(1).$$

Moreover, in contrast to Yung's result, our technique relies on '*quenched*' argument which shows that for any $k \geq 3$, both BPGD and UCP find the satisfying assignment with strictly positive probability for $d < d_{\min}$ with $d_{\min} < d_{\text{core}}$.
We also analyze different phase transition of *decimation process* for which we pinpoint the (non)-reconstruction and condensation phase transition based on different d and θ values for any $k \geq 3$.

Random k-SAT: Turning to random k-SAT, we revisited the Gibbs Uniqueness threshold. In this result we prove that for any $k \geq 3$ for the clause-to-variable densities upto the Gibbs uniqueness threshold , the number of satisfying assignments of random k-SAT is given by physics inspired '*replica symmetric solution*'. Our result sharpen the understanding of the onset of long range correlations on random k-SAT by clarifying the transition of Gibbs measure from uniqueness to (non)-uniqueness regime. Below we will compare our result with the previous results in this context.

Comparison with previous work:

Known results:

- In 2004 **Panchenko** and **Talagrand** [90] proved a rigorous upper bound on the physics formula for the number of satisfying assignments using a proof technique called '*interpolation method*'.

- **Achlioptas et.al.** in [2] proved the physics formula in the case $k = 2$ which is much simpler than for any $k \geq 3$.

- The most important work in this regards by **Montanari** and **Shah** in 2007 provided a correct approximation on the number of '*good*' satisfying assignment all but $o(n)$ clauses for $k \geq 3$.

  However, it seems difficult to estimate the gap between the number of such 'good' assignments and the number of actual satisfying assignments.

Our results:

Our paper verifies the correct number of satisfying assignments given by physics method '*replica symmetry solution*' used in [77,78] for any $k \geq 3$ upto Gibbs uniqueness threshold.
Moreover we derive a lower bound on the Gibbs uniqueness threshold which improves significantly over the work of **Montanari** and **Shah** in [83] for small $k \geq 3$.
Below we provide the values of d_{giant} (giant component threshold of hypergraph induced by random k-CNF formula), d_{MS} (Montanari-Shah bound), d_{our} (our bound), d_{pure} (pure literal threshold), d_{SAT} (satisfiability threshold) for small k-values.

k	2	3	4	5
d_{giant}	1.0000	0.5000	0.3333	0.2500
d_{MS}	1.1625	0.8792	0.8695	0.9236
d_{our}	2.0000	1.3431	1.2451	1.2635
d_{pure}	2.0000	4.9108	6.1782	7.0178
d_{SAT}	2.0000	12.801	39.724	105.585

Taken together the results in this thesis highlight the estimation of partition function (in other words, the number of satisfying assignments) of random formulas and their phase transitions by providing a clear picture of the solution space geometry of different random satisfiability problems.

Therefore, it establishes a bridge between probabilistic combinatorics, random graphs and statistical physics. More specifically, it show how the tools such as local weak convergence, recursive fixed point distributional equations, different message passing algorithms (such as Belief Propagation, Warning Propagation, UCP), correlation decay methods can be leveraged to analyze the long range dependencies, phase transitions, solution space geometry and algorithmic thresholds in random structures.

There are several challenging and more interesting open problems still remain which we will encounter in the next section.

8.2 Future Directions

Paper 1 [23]: The number of random 2-SAT solutions is asymptotically log-normal

- Investigate whether the present method of considering correlated instances be extended to random optimization problems. Moreover, establishing a central limit theorems for random optimization problems are also very interesting.

  Cao [20] provided a general framework based on the 'objective method' [9]. Unfortunately, the conditions of Cao's theorem tend to be unwieldy for MAX-CSP problems with hard constraints. Recent work of Kreačič [58] and Glasgow, Kwan, Sah, Sawhney [52] on the matching number therefore instead resorts to the use of stochastic differential equations.

- Another most interesting question can be whether the log-normal fluctuations hold for counting the number of solutions in other models such as random Horn-SAT or any planted CSPs.

- Understand the precise behavior of variance near the satisfiability threshold and across the families of random CSPs whether the fluctuations emerge or not.

Paper 2 [22]: Belief Propagation Guided Decimation on random k-XORSAT

- Unlike random k-XORSAT, due to the lack of inherent symmetry in random k-SAT, BPGD algorithm provably fails to find the satisfying assignments on random k-SAT instances even below the threshold where the set of satisfying assignments shatters into well-separated clusters [1, 59]. So, a sophisticated message passing algorithm called '*Survey Propagation Guided Decimation*' has been suggested in [72, 96]. In random k-XORSAT, both BPGD and SPGD are equivalent but these two algorithms are substantially different in random k-SAT. Therefore investigating SPGD on random k-SAT can be one of the most interesting problem in this context and one might hope that SPGD outperforms BPGD on random k-SAT and finds satisfying assignments up to the aforementioned shattering transition. A negative result to the effect that Survey Propagation Guided Decimation fails asymptotically beyond the shattering transition point for large enough k exists [56]. Yet a complete analysis of SPGD/BPGD on random k-SAT like in random k-XORSAT used in this thesis remains an outstanding challenge.

- Investigating the interplay between the decimation process and the solution space geometry especially how the frozen variables emerge dynamically.

- Apply the technique used in this paper to some inference problems in graphical models (e.g., community detection, coding theory and planted problems).

- Finally, one of the most interesting open problem towards the performance of various types of algorithms such as greedy, message passing or local search that aim to find an assignment that violates the least possible number of clauses. A first step based on the heuristic 'dynamical cavity method' was recently undertaken by Maier, Behrens and Zdeborová [64].

Paper 3 [21]: The random k-SAT Gibbs Uniqueness threshold revisited

- Pinpoint the exact threshold and nature of the Gibbs uniqueness threshold for general $k \geq 3$ and its connection with statistical physics predictions.

- Detailed exploration of the other phase transitions beyond Gibbs uniqueness like- (non)-reconstruction / reconstruction, condensation and freezing phases and their interactions.

  Physics predicts a sequence of phase transitions [59]: **Replica Symmetric (non-reconstruction) → clustering/dynamic phase → condensation → freezing**

  whereas the reconstruction threshold coincide with the clustering threshold. Physics predictions exist for the reconstruction threshold in random k-SAT using the heuristic technique called '*cavity method*' (or, *replica symmetry*) and it happens when $\alpha_{\text{rec}} \approx \frac{2^k}{k} \log k$ for large k values. So, when it comes with the long range correlations in the Gibbs measure the exact rigorously mathematical proof of reconstruction threshold is still a challenging problem.

- Extending the results on Gibbs uniqueness to other random CSPs like hypergraph coloring, independent sets in random hypergraphs and spin glasses – making a promising and interesting research direction.

8.3 Contribution of the authors

In the previous chapters we discussed an overview of the results and the proofs along with the useful tools we have used to get our results. The full versions of the papers can be found in appendix. We conclude the thesis by providing a list of papers that are the backbone of the thesis and to which the author of this thesis contributed.

The first result of this thesis is from the paper titled "***The number of random* 2*-SAT solutions is asymptotically log-normal***" by Arnab Chatterjee, Amin Coja-Oghlan, Noela Müller, Connor Riddlesden, Maurice Rolvien, Pavel Zakharov, Haodong Zhu. The extended version of the paper has been published in the *'Theory of Computing'* (**TOC**) journal and the preliminary version appeared in

'Approximation, Randomization, and Combinatorial Optimization' (**APPROX/RANDOM'24**) – Leibniz International Proceedings in Informatics (LIPIcs), volume 317, 39 : 1 – 39 : 15, Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2024). This paper establishes a central limit theorem ('CLT') on the logarithm of the number of solutions of random 2-SAT formula throughout the satisfiable regime. Beside this the paper also calculate the variance effectively. The problem first raised when NM, CR and HZ visited TU Dortmund for a week and introduced the martingale central limit theorem from the book of Hall and Heyde [55] and AC, ACO, MR and PZ jointly discussed with them about the martingale CLT on the number of satisfying assignments of random 2-SAT and how to prove the finiteness of the variance. Later AC, ACO, MR and PZ discussed towards variance calculation by constructing two correlated formulas and how to obtained a pruned correlated formula using UCP. AC, ACO, MR and PZ jointly examined the effect of removing a single clause from the original formula on the number of solutions in the pruned formula. AC contributed towards creating the Galton-Watson tree which is the local limit of two correlated formulas. Beside this, AC, NM and HZ was involving in the contraction property of $\mathtt{logBP}^{\otimes}_{d,t}$ which ensures the existence of a unique fixed point $\rho_{d,t}$ and involves the analysis of belief propagation on a Galton-Watson tree, which connects this fixed point to the BP marginals and thus completing the proof of our main result.

The second result of this thesis is from the paper titled "*Belief Propagation Guided Decimation on random k-XORSAT*" by Arnab Chatterjee, Amin Coja-Oghlan, Mihyung Kang, Lena Krieg, Maurice Rolvien and Gregory Sorkin. The extended version of the paper has been submitted to *'Theory of Computing'* (**TOC**) journal and the conference version of this paper appeared in the *'52nd International Colloquium on Automata, Languages and Programming'* (**ICALP'25**) – Leibniz International Proceedings in Informatics (LIPIcs), volume 334, 47 : 1 – 47 : 21, Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2025). This paper analyze the performance of BPGD algorithm on random k-XORSAT formula and different phase transition of the decimation process. The problem was first raised when GS visited TU Dortmund and ACO introduced the problem by pointing out the paper by two physicists Ricci-Tersenghi and Semerjian [96] that their paper only provide a heuristic on the success probability of the algorithm. AC,MR and GS worked towards the tight analysis of BPGD by analyzing the function $\mathbf{\Phi}_{d,k}$ (Bethe-Free Entropy) and some threshold values of d. AC carried out the calculus part which leads to the proof of the behavior of the function w.r.t. different parameters (d,λ). Later, AC, ACO and LK jointly developed the lemmas involving the presence of so-called toxic cycles in a sub formula which create obstacles towards the success probability of the formula.

Coming to the last result of this thesis, it is from the paper titled "*The random k-SAT Gibbs uniqueness threshold revisited*" by Arnab Chatterjee, Amin Coja-Oghlan, Catherine Greenhill, Vincent Pfenninger, Maurice Rolvien, Pavel Zakharov and Konstantinos Zampetakis. The paper has been submitted to "*Combinatorics, Probability and Computing*" (**CPC**) journal. The aim of this paper is to determine the number of actual satisfying assignments of random k-SAT formula for clause-to-variable densities upto Gibbs uniqueness threshold. The problem was first raised when CG and VP visited TU Dortmund and we came up with an problem on counting the number of actual assignments of random k-SAT formula. Then we all looked at the paper by Montanari and Shah [83] in which they showed that for $k \geq 3$

certain clause/variable densities the ’replica symmetric solution’ from physics correctly approximates the number of ‘good’ assignments that satisfy all but $o(n)$ clauses. But unfortunately their bound was not tight even for $k = 2$ case. In this context AC, ACO, PZ and KZ jointly developed the pure literal pursuit (’PULP’) algorithm whose purpose is to trace the repercussions of setting a relatively small number of variables to specific truth values which constitutes the main technical challenge towards proof of the main result. AC and KZ was involved in constructing pure and mixed literal operator $\mathrm{LL}^{\star}_{k,d}$ and the contraction of the operator with respect to a metric related to W_1-Wasserstein distance and this summarizes the main step towards the proof of second theorem of the paper.

nainaṁ chhindanti śhhastrāṇi nainaṁ dahati pāvakaḥ
na chainaṁ kledayantyāpo na śhhoṣhayati mārutaḥ

The atma (soul) cannot be shattered by weapons, it cannot be burnt by fires,
it cannot be drenched by the waters and it cannot be rendered dry by the winds.
– Srimat Bhagavad Gita (2.23)

Bibliography

[1] D. Achlioptas and A. Coja-Oghlan. Algorithmic barriers from phase transitions. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008*, October 25–28, 2008, Philadelphia, PA, USA, IEEE Computer Society, pp. 793–802 (2008).

[2] D. Achlioptas, A. Coja-Oghlan, M. Hahn-Klimroth, J. Lee, N. Müller, M. Penschuck, G. Zhou: The number of satisfying assignments of random 2-SAT formulas. Random Structures and Algorithms 58, pp. 609–647, (2021).

[3] D. Achlioptas and E. Friedgut: A sharp threshold for k-colorability. *Random Structures & Algorithms*,14, pp. 63–70, (1999).

[4] D. Achlioptas, C. Moore: Random k-SAT: two moments suffice to cross a sharp threshold. *SIAM Journal on Computing* 36, 3, pp.740–762, (2006).

[5] D. Achlioptas and A. Naor: The Two Possible Values of the Chromatic Number of a Random Graph. *Annals of Mathematics*, Second Series, Vol. 162, 3, pp. 1335-1351, (2005).

[6] D. Achlioptas, A. Naor and Y. Peres: Rigorous location of phase transitions in hard optimization problems. *Nature 435*, 7043, pp.759–764, (2005).

[7] D. Achlioptas, Y. Peres: The threshold for random k-SAT is $2^k \log 2 - O(k)$. *Journal of the AMS 17*, pp.947–973, (2004).

[8] D. Achlioptas and F. Ricci-Tersenghi. On the solution-space geometry of random constraint satisfaction problems. In *Proceedings of 38th STOC*, 130-139, New York, NY, USA, (2006), ACM.

[9] D. Aldous, J. Steele: The objective method: probabilistic combinatorial optimization and local weak convergence. In: H. Kesten (ed.): Probability on Discrete Structures. Springer (2004).

[10] N. Alon and J. Spencer: The probabilistic method. *Willey*, 2nd edition, (2000).

[11] S. Arora and B. Barak. *Computational complexity: a modern approach.* Cambridge University Press, (2009).

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

[13] A.B.Babaev, A.K.Murtazaev and F.A. Kassan-Ogly: Ground State of an Antiferromagnetic Three-State Potts Model on a Triangular Lattice with Competing Interactions. *Journal of Experimental and Theoretical Physics*, Vol. 127, pp. 323–327, (2018).

[14] H. A. Bethe. Statistical physics of superlattices. *Proceedings of the Royal Society of London A*, 150:552–575, (1935).

[15] G. Biroli, R. Monasson and M. Weigt. A variational description of the ground state structure in random satisfiability problems. *The European Physics Journal B*, 14:551, (2000).

[16] R. Biswas, W. Chen, A. Sen: On the replica symmetric solution in general diluted spin glasses. arXiv:2410.15599 (2024).

[17] S.C. Brailsford, C.N. Potts and B.M. Smith. Constraint satisfaction problems: Algorithms and applications. *European journal of operational research*, 119(3):557–581, (1999).

[18] G. Bresler, B. Huang: The algorithmic phase transition of random k-sat for low degree polynomials. In *62nd Annual Symposium on Foundations of Computer Science, FOCS 2021*, IEEE Computer Society, Los Alamitos, CA, pp.298–309, (2022).

[19] A. Broder, A. Frieze, E. Upfal: On the satisfiability and maximum satisfiability of random 3-CNF formulas. Proc. 4th SODA, pp.322–330, (1993).

[20] S. Cao: Central limit theorems for combinatorial optimization problems on sparse Erdős-Rényi graphs. Annals of Applied Probability 31, pp.1687–1723, (2021).

[21] A. Chatterjee, A. Coja-Oghlan, C. Greenhill, V. Pfenninger, M. Rolvien, P. Zakharov, K. Zampetakis: The random k-SAT Gibbs Uniqueness Threshold revisited. *arXiv preprint arXiv:2506.01359*, (2025).

[22] A. Chatterjee, A. Coja-Oghlan, M. Kang, L. Krieg, M. Rolvien, G. Sorkin: Belief Propagation Guided Decimation on random k-XORSAT. In *Proceedings of the 52nd ICALP*, (2025).

[23] A. Chatterjee, A. Coja-Oghlan, N. Müller, C. Riddlesden, M. Rolvien, P. Zakharov, H. Zhu: The number of random 2-SAT solutions is asymptotically log-normal. In *Proceedings of 28th RANDOM*, 39, (2024).

[24] P. Cheeseman, B. Kanefsky, W. Taylor: Where the *really* hard problems are. In *Proceedings of the IJCAI*, pp.331–337, (1991).

[25] W.-K. Chen, P. Dey, D. Panchenko: Fluctuations of the free energy in the mixed p-spin models with external field. Probability Theory and Related Fields 168, pp.41–53, (2017).

[26] V. Chvátal, B.A. Reed: Mick gets some (the odds are on his side). In *33rd Annual Symposium on Foundations of Computer Science, Pittsburgh, Pennsylvania, USA*, 24–27 October 1992, IEEE Computer Society, pp.620–627.

[27] A. Coja-Oghlan. A better algorithm for random k-SAT. *SIAM Journal on Computing*, 39(7):2823–2864, (2010).

[28] A. Coja-Oghlan: Belief Propagation fails on random formulas. Journal of the ACM 63 (2017) #49.

[29] A. Coja-Oghlan, A. Ergür, P. Gao, S. Hetterich, M. Rolvien: The rank of sparse random matrices. Proc. 31st SODA, pp. 579–591, (2020).

[30] A. Coja-Oghlan, A. Haqshenas and S. Hetterich. Walksat stalls well below satisfiability. *SIAM Journal on Discrete Mathematics*, 31(2):1160–1173, (2017).

[31] A. Coja-Oghlan, T. Kapetanopoulos, N. Müller: The replica symmetric phase of random constraint satisfaction problems. *Combinatorics, Probability and Computing* 29, 3 , pp.346-422, (2020).

[32] A. Coja-Oghlan, F. Krzakala, W. Perkins, L. Zdeborová: Information-theoretic thresholds from the cavity method. Advances in Mathematics 333, pp.694–795, (2018).

[33] A. Coja-Oghlan, A. Pachon-Pinzon: The decimation process in random k-SAT. SIAM Journal on Discrete Mathematics 26, pp.1471–1509, (2012).

[34] A. Coja-Oghlan and K. Panagiotou. The asymptotic k-SAT threshold. *Advances in Mathematics*, 288, pp.985–1068, (2016).

[35] A. Coja-Oghlan, N. Wormald: The number of satisfying assignments of random regular k-SAT formulas. *Combinatorics, Probability and Computing* 27, pp.496–530, (2018).

[36] A. Crisanti, G. Paladin and H.-J.S.A. Vulpiani: Replica trick and fluctuations in disordered systems. *Journal de Physique I*, 2(7):1325–1332, (1992).

[37] A. Dembo, A. Montanari: Ising models on locally tree-like graphs. Annals of Applied Probability 20, pp.565–592, (2010).

[38] A. Dembo, A. Montanari: Gibbs measures and phase transitions on sparse random graphs. Brazilian Journal of Probability and Statistics 24, pp.137–211, (2010).

[39] A. Dembo, A. Montanari, A. Sly, N. Sun: The replica symmetric solution for Potts models on d-regular graphs. *Communications in Mathematical Physics*, 327, pp. 551–575, (2014).

[40] A. Dembo, A. Montanari, N. Sun: Factor models on locally tree-like graphs. Annals of Probability 41, pp.4162–4213, (2013).

[41] J. Ding, A. Sly and N. Sun. Proof of the satisfiability conjecture for large k. 20 *Annals of Mathematics 196*, pp.1–388, (2022).

[42] O. Dubois, J. Mandler: The 3-XORSAT threshold. In *43rd Annual IEEE Symposium on Foundations of Computer Science, FOCS*, pp. 769–778, (2002).

[43] A. Dylan: *PhD Thesis* : Gibbs Measures on Sparse Random Graphs. *Princeton University*, Computer Science Dept. Engineering Quadrangle Princeton, NJ, United States, (2022).

[44] C. Efthymiou: On sampling symmetric Gibbs distributions on sparse random graphs and hypergraphs. In *49th EATCS International Conference on Automata, Languages, and Programming*, vol. 229 of *LIPICs. Leibniz Int. Proc. Inform.* Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, (2022), page Art. No. 57, 16.

[45] J. Feigenbaum. The use of coding theory in computational complexity. In *Proceedings of Symposium in Applied Mathematics*, volume 50, pp.207–233, (1995).

[46] E.C. Freuder and A.K. Mackworth. Constraint satisfaction: An emerging paradigm. In *Foundations of Artificial Intelligence*, volume 2, 13–27. Elsevier, (2006).

[47] E. Friedgut. Sharp thresholds of graph properties, and the k-SAT problem. *Journal of the AMS*, 12:1017–1054, (1999).

[48] A. Frieze and M. Karoński: Introduction to Random Graphs. *Cambridge University Press*, (2015).

[49] R. G. Gallager: Low-density parity check codes. *IEEE Transaction on Information Theory*, 8:21–28, (1962).

[50] D. Gamarnik and M.Sudan. Performance of sequential local algorithms for the random NAE-k-SAT problem. *SIAM Journal on Computing*, 46(2):590–619, (2017).

[51] M.R. Garey and D.S. Johnson. Computers and intractability: *A guide to the theory of NP-completeness.* Freeman, San Francisco, (1979).

[52] M. Glasgow, M. Kwan, A. Sah, M. Sawhney: A central limit theorem for the matching number of a sparse random graph. arXiv:2402.05851 (2024).

[53] A. Goerdt: A threshold for unsatisfiability. *Journal of Computer and System Sciences*, 53, pp.469–486, (1996).

[54] J. Gu and R. Sosic. A parallel architecture for constraint satisfaction. In *International conference on industrial and engineering applications of artificial intelligence and expert systems*,pp. 229–237, (1991).

[55] P. Hall, C. Heyde: Martingale limit theory and its applications. Academic Press (1980).

[56] S. Hetterich: Analysing survey propagation guided decimation on random formulas. *arXiv preprint arXiv:1602.08519*, (2016).

[57] L.M. Kirousis, E. Kranakis, D. Krizanc and Y.C. Stamatiou: Approximating the unsatisfiability threshold of random formulas. *Random Structures & Algorithms*, 12(3):253–269, (1998).

[58] E. Kreačič: Some problems related to the Karp-Sipser algorithm on random graphs. Ph.D. thesis, University of Oxford, (2017).

[59] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian and L. Zdeborová: Gibbs states and the set of solutions of random constraint satisfaction problems. In *Proceedings of the National Academy of Sciences* 104 (25), pp. 10318–10323, (2007).

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

[61] V. Kumar. Algorithms for constraint-satisfaction problems: A survey. *AI magazine*, 13(1):32–32, (1992).

[62] L. Lovász: Large networks and graph limits. AMS (2012).

[63] T. Łuczakand J.C. Wierman: The chromatic number of random graphs at the double-jump threshold. *Combinatorica*, Volume 9, pages 39–49, (1989).

[64] A. Maier, F. Behrens, L. Zdeborová: Dynamical cavity method for hypergraphs and its application to quenches in the k-XOR-SAT problem. arxiv 2412.14794 (2024).

[65] S. Mertens, M. Mézard and R. Zecchina. Threshold values of random k-SAT from the cavity method. *Random Structures and Algorithms*, 28(3):340–373, (2006).

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

[67] M. Mézard and A. Montanari. Reconstruction on trees and spin glass transition. *Journal of Statistical Physics*, 124:1317-1350, (2006).

[68] M. Mézard and T. Mora. Constraint satisfaction problems and neural networks: A statistical physics perspective. *Journal of Physiology-Paris,* 103(1-2): 107–113, (2009).

[69] M. Mézard and G. Parisi. The bethe lattice spin glass revisited. *The European Physics Journal B*, 20:217, (2001).

[70] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse, and M. Virasoro: Replica symmetry breaking and the nature of the spin glass phase. *Journal de Physique*, 45(5), pp. 843–854, (1984).

[71] M. Mézard, G. Parisi and M. Virasoro: Spin glass theory and beyond: An introduction to the replica method and its applications, *World Scientific Publishing Company*, Vol. 9, (1987).

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

[73] D. Mitchell, B. Selman and H. Levesque. Hard and easy distributions of SAT problems. In *Proceedings of the 10th National Conference on Artificial Intelligence*, pp. 459-465, (1992).

[74] M. Molloy. Models for random constraint satisfaction problems. *SIAM Journal on Computing*, 32(4):935–949, (2003).

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

[76] M. Molloy. The freezing threshold for k-colorings of a random graph. In *Proceedings of the 44th symposium on Theory of Computing*, 921. ACM, (2012).

[77] R. Monasson, R. Zecchina. The entropy of the k-satisfiability problem. *Physics Review Letter*, 76, 3881, (1996).

[78] R. Monasson, R. Zecchina: Statistical mechanics of the random K-SAT model. Phys. Rev. E 56, pp.1357–1370, (1997).

[79] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman and L. Troyansky. 2+p-sat: Relation of typical-case complexity to the nature of the phase transition. *Random Structures and Algorithms*, 15:414, (1999).

[80] A. Montanari, R. Restrepo and P. Tetali. Reconstruction and clustering in random constraint satisfaction problems. *SIAM Journal on Discrete Mathematics*, 25(2):771–808, (2011).

[81] A. Montanari, F. Ricci-Tersenghi, G. Semerjian. Cluster of solutions and replica symmetry breaking in random k-satisfiability. *Journal of Statistical Mechanics*, P04004, (2008).

[82] A. Montanari and G. Semerjian: Rigorous inequalities between length and time scales in glassy systems. *Journal of Statistical Physics*, 125: 23, (2006).

[83] A. Montanari and D. Shah. Counting good truth assignments of random k-SAT formulae. In *Proceedings of the 18th SODA*, pp. 1255–1264, (2007).

[84] C. Moore and S. Mertens. *The Nature of Computation.* Oxford University Press, (2011).

[85] E. Mossel and Y. Peres. Information flow on trees. *Annals of Applied Probability*, 13(3):817–844, 08 (2003).

[86] N. Múller, R. Neininger, H. Zhu. Random 2-SAT: The set of atoms of the limiting empirical marginal distribution. arXiv:2410.17749 [math.PR], (2024).

[87] D. Panchenko: The Sherrington-Kirkpatrick model. Springer (2013).

[88] D. Panchenko: Spin glass models from the point of view of spin distributions. Annals of Probability 41, pp.1315–1361, (2013).

[89] D. Panchenko: On the replica symmetric solution of the K-sat model. Electron. J. Probab. 19 (2014) #67.

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

[91] C.H. Papadimitriou. *Computational complexity.* Addison-Wesley, (1994).

[92] J. Pearl: Probabilistic reasoning in intelligent systems : networks of plausible inference. *Morgan kaufmann Publishers Inc.*, San Francisco, CA, USA, (1988).

[93] B. Pittel, G.B. Sorkin: The satisfiability threshold for k-XORSAT. *Combinatorics, Probability and Computing* 25, 2, pp. 236–268, (2016).

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

[95] F. Rassmann : On the number of solutions in random graph k-coloring. *Combinatorics, Probability and Computing* 28, 1, pp. 130–158, (2019).

[96] F. Ricci-Tersenghi, G. Semerjian: On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms. *Journal of Statistical Mechanics*, P09001, (2009).

[97] R. Robinson, N. Wormald: Almost all regular graphs are Hamiltonian. Random Structures and Algorithms 5, pp. 363–374, (1994).

[98] A. Sly: Computational transition at the uniqueness threshold. Proc. 51st FOCS, pp. 287–296, (2010).

[99] E.P.K. Tsang. *Foundations of constraint satisfaction.* Academic press, (1993).

[100] L.G. Valiant. The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, 8(3):410–421, (1979).

[101] J.S. Yedidia, W.T. Freeman and Y. Weiss: Constructing free energy approximations and generalized belief propagation algorithms. *Technical report TR-2002–35, Mitsubishi Electrical Research Laboratories*, (2002).

[102] K. Yung: Limits of sequential local algorithms on the random k-XORSAT problem. Proc. 51st ICALP (2024) #123.

A
List of Papers

THE NUMBER OF RANDOM 2-SAT SOLUTIONS IS ASYMPTOTICALLY LOG-NORMAL

ARNAB CHATTERJEE, AMIN COJA-OGHLAN, NOËLA MÜLLER, CONNOR RIDDLESDEN, MAURICE ROLVIEN, PAVEL ZAKHAROV, HAODONG ZHU

ABSTRACT. We prove that throughout the satisfiable phase, the logarithm of the number of satisfying assignments of a random 2-SAT formula satisfies a central limit theorem. This implies that the log of the number of satisfying assignments exhibits fluctuations of order $\sqrt{n}$, with n the number of variables. The formula for the variance can be evaluated effectively. By contrast, for numerous other random constraint satisfaction problems the typical fluctuations of the logarithm of the number of solutions are *bounded* throughout all or most of the satisfiable regime. *MSc: 05C80, 60C05, 68Q87*

1. INTRODUCTION

1.1. **Background and motivation.** The quest for satisfiability thresholds has been a guiding theme of research into random constraint satisfaction problems [7, 17, 25]. But once the satisfiability threshold has been pinpointed a question of no less consequence is to determine the distribution of the number of satisfying assignments within the satisfiable phase [35]. Indeed, the number of solutions is intimately tied to phase transitions that affect the geometry of the solution space, which in turn impacts the computational nature of finding or sampling solutions [4, 18, 29]. However, few tools are currently available to count solutions of random problems. Where precise rigorous results exist (such as in random NAESAT or XORSAT), the proofs typically rely on the method of moments (e.g., [6, 27, 43, 44]). Yet a necessary condition for the success of this approach is that the problem in question exhibits certain symmetries, which are absent in many interesting cases [7, 21].

The aim of the present paper is to shed a closer light on the number of satisfying assignments in random 2-SAT, the simplest random CSP that lacks said symmetry properties. While the random 2-SAT satisfiability threshold has been known since the 1990s [20, 32], a first-order approximation to the number of satisfying assignments has been obtained only recently [5]. This timeline reflects the computational complexity of the respective questions. As is well known, deciding the satisfiability of a 2-CNF reduces to directed reachability, solvable in polynomial time [10].

By contrast, calculating the number of satisfying assignmets $Z(\Phi)$ of a 2-CNF Φ is a #P-hard task [48]. Nonetheless, Monasson and Zecchina [38] put forward a delicate physics-inspired conjecture as to the exponential order of the number of satisfying assignments of random 2-CNFs. Achlioptas et al. [5] recently proved this conjecture. Their theorem provides a first-order, law-of-large-numbers approximation of the logarithm of the number of satisfying assignments. The present paper contributes a much more precise result, namely a central limit theorem. We show that throughout the satisfiable phase the logarithm of the number of satisfying assignments, suitably shifted and scaled, converges to a Gaussian. This is the first central limit theorem of this type for any random CSP.

Let $\mathbf{\Phi} = \mathbf{\Phi}_{n,m}$ be a random 2-CNF on n Boolean variables $x_1, \ldots, x_n$ with m clauses, drawn independently and uniformly from all $4\binom{n}{2}$ possible 2-clauses. Suppose that $m \sim dn/2$ for a fixed real $d > 0$. Thus, d gauges the average number of clauses in which a variable x_i appears. The value $d = 2$ marks the satisfiability threshold; hence, $\mathbf{\Phi}$ is satisfiable with high probability ('w.h.p.') if $d < 2$, and unsatisfiable w.h.p. if $d > 2$ [20, 32]. Achlioptas et al. [5] determined a function $\phi(d) > 0$ such that for all $d < 2$, i.e., throughout the entire satisfiable phase we have

$$Z(\mathbf{\Phi}) = \exp(n\phi(d) + o(n)) \qquad \text{w.h.p.}, \tag{1.1}$$

thereby determining the leading exponential order of $Z(\mathbf{\Phi})$.

However, (1.1) fails to identify the limiting distribution of $Z(\mathbf{\Phi})$. To be precise, since (1.1) shows that $Z(\mathbf{\Phi})$ scales exponentially, we expect this random variable to exhibit *multiplicative* fluctuations. Therefore, the appropriate goal is to find the limiting distribution of the logarithm of this random variable, i.e., of $\log Z(\mathbf{\Phi})$. Indeed, physics intuition suggests that $\log Z(\mathbf{\Phi})$ should be asymptotically Gaussian [36]. The main result of the present paper confirms this hunch. Specifically, letting $\mathbf{\Gamma}_{\eta(d)}$ be a Gaussian with mean 0 and standard deviation $\eta(d) > 0$, we prove that for all $0 < d < 2$, $\log Z(\mathbf{\Phi})$ satisfies

$$\mathbb{P}\left[\log Z(\mathbf{\Phi}) - \mathbb{E}[\log Z(\mathbf{\Phi}) \mid Z(\mathbf{\Phi}) > 0] < z\sqrt{m}\right] \sim \mathbb{P}\left[\mathbf{\Gamma}_{\eta(d)} < z\right] \qquad (z \in \mathbb{R}). \tag{1.2}$$

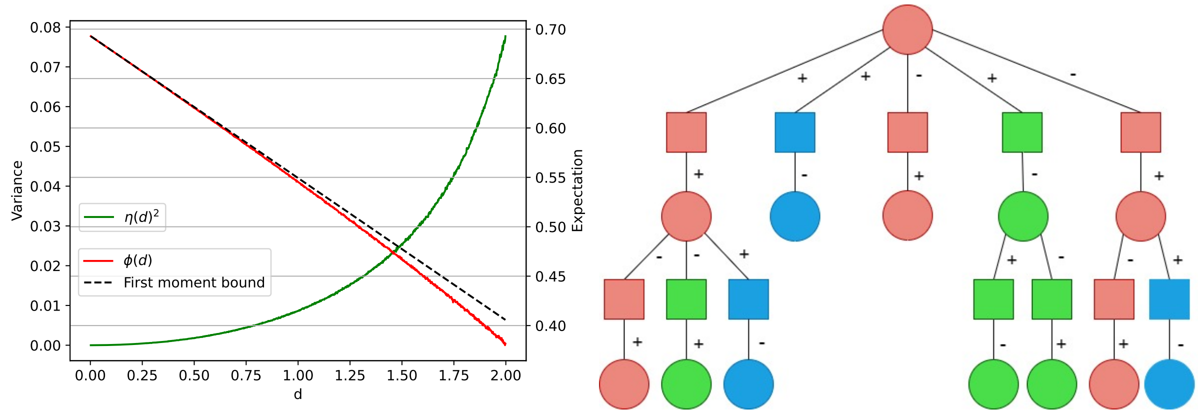


FIGURE 1. *Left:* Numerical approximations to the function $\phi(d)$ from (1.1) (red) and the variance $\eta(d)^2$ from (1.7) (green). The black dashed line is the first moment bound $d \mapsto \log(2) + \frac{d}{2}\log(3/4)$. *Right:* An illustration of the tree $\boldsymbol{T}^{\otimes}$ from Section 2.6.

The order $\Theta(\sqrt{n})$ of fluctuations confirmed by (1.2) sets random 2-SAT apart from a large family of other random constraint satisfaction problems. For example, for random graph q-colouring with $q \geq 3$ colours the log of the number of q-colourings *superconcentrates*, i.e., merely has *bounded* fluctuations throughout most of the regime where the random graph is q-colourable [12].[1] The same is true of random NAESAT, XORSAT and the symmetric perceptron [1, 11, 21, 43]. In each of these cases, certain fundamental symmetry properties (e.g., that the set of q-colourings remains invariant under permutations of the colours) enable the computation of the number of solutions via the method of moments. Random 2-SAT lacks the respective symmetry (as the set of satisfying assignments is not generally invariant under swapping 'true' and 'false'), and accordingly (1.2) establishes that the number of solutions fails to superconcentrate (for more details see [21]).

1.2. **The main result.** The formula for the standard deviation $\eta(d)$ from (1.2) comes in terms of a fixed point equation on a space of probability measures. Thus, let $\mathscr{P}(\mathbb{R}^2)$ be the set of all (Borel) probability measures on $\mathbb{R}^2$. For $0 < d < 2$ and $0 \leq t \leq 1$ we define an operator

$$\mathtt{logBP}^{\otimes}_{d,t} : \mathscr{P}\left(\mathbb{R}^2\right) \to \mathscr{P}\left(\mathbb{R}^2\right), \qquad \rho \mapsto \hat{\rho} = \mathtt{logBP}^{\otimes}_{d,t}(\rho), \tag{1.3}$$

as follows. Let

$$(\boldsymbol{\xi}_{\rho,i})_{i\geq 1}, (\boldsymbol{\xi}'_{\rho,i})_{i\geq 1}, (\boldsymbol{\xi}''_{\rho,i})_{i\geq 1}, \qquad \boldsymbol{\xi}_{\rho,i} = \begin{pmatrix} \boldsymbol{\xi}_{\rho,i,1} \\ \boldsymbol{\xi}_{\rho,i,2} \end{pmatrix}, \boldsymbol{\xi}'_{\rho,i} = \begin{pmatrix} \boldsymbol{\xi}'_{\rho,i,1} \\ \boldsymbol{\xi}'_{\rho,i,2} \end{pmatrix}, \boldsymbol{\xi}''_{\rho,i} = \begin{pmatrix} \boldsymbol{\xi}''_{\rho,i,1} \\ \boldsymbol{\xi}''_{\rho,i,2} \end{pmatrix}$$

be random vectors with distribution ρ, let $\boldsymbol{d} \stackrel{\text{dist}}{=} \mathrm{Po}(td)$, $\boldsymbol{d}', \boldsymbol{d}'' \stackrel{\text{dist}}{=} \mathrm{Po}((1-t)d)$ and let $\boldsymbol{s}_i, \boldsymbol{s}'_i, \boldsymbol{s}''_i, \boldsymbol{r}_i, \boldsymbol{r}'_i, \boldsymbol{r}''_i$ for $i \geq 1$ be uniformly random on $\{\pm 1\}$, all mutually independent. Then $\hat{\rho}$ is the distribution of the vector

$$\begin{pmatrix} \sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \log\left(\frac{1}{2}\left(1 + \boldsymbol{r}_i \tanh(\boldsymbol{\xi}_{\rho,i,1}/2)\right)\right) + \sum_{i=1}^{\boldsymbol{d}'} \boldsymbol{s}'_i \log\left(\frac{1}{2}\left(1 + \boldsymbol{r}'_i \tanh(\boldsymbol{\xi}'_{\rho,i,1}/2)\right)\right) \\ \sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \log\left(\frac{1}{2}\left(1 + \boldsymbol{r}_i \tanh(\boldsymbol{\xi}_{\rho,i,2}/2)\right)\right) + \sum_{i=1}^{\boldsymbol{d}''} \boldsymbol{s}''_i \log\left(\frac{1}{2}\left(1 + \boldsymbol{r}''_i \tanh(\boldsymbol{\xi}''_{\rho,i,2}/2)\right)\right) \end{pmatrix} \in \mathbb{R}^2 \ .$$

In addition, define a function $\mathscr{B}^{\otimes}_{d,t} : \mathscr{P}(\mathbb{R}^2) \to (0,\infty]$ by letting

$$\mathscr{B}^{\otimes}_{d,t}(\rho) = \mathbb{E}\left[\prod_{h=1}^{2} \log\left(1 - \frac{1}{4}(1 + \boldsymbol{r}_1 \tanh(\boldsymbol{\xi}_{\rho,1,h}/2))(1 + \boldsymbol{r}_2 \tanh(\boldsymbol{\xi}_{\rho,2,h}/2))\right)\right]. \tag{1.4}$$

[1] Formally, up to the so-called condensation threshold, which precedes the q-colourabiliy threshold by a small additive constant, the logarithm of the number of q-colurings minus its expectation converges in distribution to a random variable with bounded moments [12, 13, 21].

Theorem 1.1. *For any* $0<d<2$, $t\in[0,1]$ *there exists a unique probability measure* $\rho_{d,t}\in\mathscr{P}(\mathbb{R}^2)$ *such that*

$$\rho_{d,t}=\mathtt{logBP}^{\otimes}_{d,t}(\rho_{d,t}) \qquad \textit{and} \qquad \int_{\mathbb{R}^2}\|\xi\|_2^2\mathrm{d}\rho_{d,t}(\xi)<\infty. \tag{1.5}$$

Furthermore,

$$\lim_{n\to\infty}\frac{\log Z(\mathbf{\Phi})-\mathbb{E}[\log Z(\mathbf{\Phi})\mid Z(\mathbf{\Phi})>0]}{\sqrt{m}}=\mathbf{\Gamma}_{\eta(d)} \qquad \textit{in distribution, where} \tag{1.6}$$

$$\eta(d)^2=\int_0^1\mathscr{B}^{\otimes}_{d,t}(\rho_{d,t})\mathrm{d}t-\mathscr{B}^{\otimes}_{d,0}(\rho_{d,0})\in(0,\infty). \tag{1.7}$$

The conditioning on $\log Z(\mathbf{\Phi})>0$ is necessary in (1.6), because even for $d<2$ the formula $\mathbf{\Phi}$ is unsatisfiable with probability $\Omega(n^{-1})$, in which case $\log Z(\mathbf{\Phi})=-\infty$. Moreover, the L^2-bound from (1.5) ensures that the integral (1.7) is well-defined. Finally, (1.6) implies (1.2).

How can the formula (1.7) be evaluated? Because the proof of the uniqueness of the stochastic fixed point $\rho_{d,t}$ from (1.5) is based on the contraction method, a fixed point iteration will converge rapidly. In effect, for any d,t a discrete distribution that approximates $\rho_{d,t}$ arbitrarily well (in Wasserstein distance) can be computed via a randomised algorithm called *population dynamics* [36, Chapter 14]. Since $\mathscr{B}^{\otimes}_{d,t}(\rho_{d,t})$ varies continously in d and t, $\eta(d)^2$ can thus be approximated within any desired accuracy, see Figure 1.

2. Proof strategy

The main challenge towards the proof of Theorem 1.1 is to get a handle on the variance of $\log Z(\mathbf{\Phi})$ given satisfiability. The key idea, inspired by spin glass theory [19] but novel to random constraint satisfaction, is to count the *joint* number of satisfying assignments of two correlated random formulas. Once this is accomplished Theorem 1.1 will follow from the careful application of a general martingale central limit theorem. To get acclimatised we first revisit the method of moments, the reasons it fails on random 2-SAT and the combinatorial interpretation of the law of large numbers (1.1).

2.1. **The method of moments fails.** The default approach to estimating the number of solutions to a random CSP is the venerable second moment method [7]. Its thrust is to show that the second moment of the number of solutions is of the same order as the square of the expected number of solutions. If so then the moment computation together with small subgraph conditioning yields the precise limiting distribution of the number of solutions [24, 45]. However, this approach works only if the log of the number of solutions superconcentrates around the log of the expected number of solutions.

This necessary condition is not satisfied in random 2-SAT. In fact, a straightforward calculation yields

$$\frac{1}{n}\log\mathbb{E}[Z(\mathbf{\Phi})]\sim\log 2+\frac{d}{2}\log(3/4). \tag{2.1}$$

The formula on the r.h.s. is displayed as the black dashed line in Figure 1. As can be verified analytically, this line strictly exceeds the function $\phi(d)$ from (1.1) for any $0<d<2$. Consequently, (1.1) implies that $\log Z(\mathbf{\Phi})\le\log\mathbb{E}[Z(\mathbf{\Phi})]-\Omega(n)$ w.h.p. In other words, the expected number of solutions $\mathbb{E}[Z(\mathbf{\Phi})]$ overshoots the typical number of solutions by an exponential factor w.h.p. ; cf. the discussion in [6, 8].

2.2. **Belief Propagation.** Instead of the method of moments, the prescription of the physics-based work of Monasson and Zecchina [38] is to estimate $\log Z(\mathbf{\Phi})$ by way of the Belief Propagation (BP) message passing algorithm. This approach was vindicated rigorously by Achlioptas et al. [5].

As we will reuse certain elements of that analysis we dwell on BP briefly. For a clause a of a 2-CNF Φ let $\partial a=\partial_\Phi a$ be the set of variables that a contains. Moreover, for $x\in\partial a$ let $\mathrm{sign}_\Phi(x,a)=\mathrm{sign}(x,a)\in\{\pm1\}$ be the sign with which x appears in a. Analogously, let $\partial x=\partial_\Phi x$ be the set of clauses in which variable x appears. BP introduces 'messages' between clauses a and the variables $x\in\partial a$. More precisely, each such clause-variable pair a,x comes with two messages $\mu_{x\to a},\mu_{a\to x}$. The messages are probability distributions on 'true' and 'false', which we represent by ±1. Thus, $\mu_{x\to a}(\pm1),\mu_{a\to x}(\pm1)\ge0$ and $\mu_{x\to a}(1)+\mu_{x\to a}(-1)=\mu_{a\to x}(1)+\mu_{a\to x}(-1)=1$.

The messages get updated iteratively by an operator

$$\mathrm{BP}:(\mu_{x\to a},\mu_{a\to x})_{a,x\in\partial a}\mapsto(\hat\mu_{x\to a},\hat\mu_{a\to x})_{a,x\in\partial a}=\mathrm{BP}((\mu_{x\to a},\mu_{a\to x})_{a,x\in\partial a}). \tag{2.2}$$

For a clause a with adjacent variables $\partial a = \{x, y\}$ the updated messages $\hat{\mu}_{a\to x}(\pm 1)$ are defined by

$$\hat{\mu}_{a\to x}(\mathrm{sign}(x,a)) = \frac{1}{1+\mu_{y\to a}(\mathrm{sign}(y,a))}, \qquad \hat{\mu}_{a\to x}(-\mathrm{sign}(x,a)) = \frac{\mu_{y\to a}(\mathrm{sign}(y,a))}{1+\mu_{y\to a}(\mathrm{sign}(y,a))}. \tag{2.3}$$

Moreover, for a variable x and a clause $a \in \partial x$ we define[2]

$$\hat{\mu}_{x\to a}(s) = \frac{\prod_{b\in\partial x\setminus\{a\}}\mu_{b\to x}(s)}{\prod_{b\in\partial x\setminus\{a\}}\mu_{b\to x}(1)+\prod_{b\in\partial x\setminus\{a\}}\mu_{b\to x}(-1)} \qquad (s\in\{\pm1\}) \ . \tag{2.4}$$

The purpose of BP is to heuristically 'approximate' the marginal probabilities that a random satisfying assignment $\boldsymbol{\sigma} = \boldsymbol{\sigma}_\Phi$ of Φ will set a certain variable to a specific truth value. The 'approximation' given by the set $(\mu_{x\to a}, \mu_{a\to x})_{a,x\in\partial a}$ of messages reads

$$\mu_x(s) = \frac{\prod_{b\in\partial x}\mu_{b\to x}(s)}{\prod_{b\in\partial x}\mu_{b\to x}(1)+\prod_{b\in\partial x}\mu_{b\to x}(-1)} \qquad (s\in\{\pm1\}). \tag{2.5}$$

The BP 'ansatz' now asks that we iterate the BP operator until an (approximate) fixed point is reached, i.e., ideally until $\hat{\mu}_{a\to x} = \mu_{a\to x}$ and $\hat{\mu}_{x\to a} = \mu_{x\to a}$ for all a, x. Then we evaluate the BP marginals (2.5) and plug them into a generic formula called the *Bethe free entropy*, which yields the BP 'approximation' of $\log Z(\Phi)$; an excellent exposition can be found in [36]. The BP recipe provably yields the correct result if the bipartite graph induced by the clause-variable incidences of the 2-CNF Φ is acyclic, but may be totally off otherwise.

Of course, for $1 < d < 2$ the bipartite graph associated with the random formula $\boldsymbol{\Phi}$ contains cycles in abundance. Nonetheless, (1.1) confirms that the BP formula provides a valid approximation to within $o(n)$. The proof is based on two observations. First, that the local structure of the clause-variable incidence graph can be described by a Galton-Watson tree. Second, that the Galton-Watson tree enjoys a spatial mixing property called *Gibbs uniqueness*.

Since the proof of Theorem 1.1 also harnesses Gibbs uniqueness, let us elaborate. To mimic the local structure of $\boldsymbol{\Phi}$ consider a multitype Galton-Watson tree $\boldsymbol{T}$ whose types are *variable nodes* and *clause nodes* of four sub-types (s, s') with $s, s' \in \{\pm1\}$. The root o is a variable node. The offspring of any variable node is a $\mathrm{Po}(d/4)$ number of clause nodes of each of the four sub-types. Finally, the offspring of a clause node is a single variable node. The clause type (s, s') indicates that s is the sign with which the parent variable appears in the clause, while s' determines the sign of the child variable. Thus, the Galton-Watson tree $\boldsymbol{T}$ can be viewed as a (possibly infinite) 2-CNF. For an integer $\ell \ge 0$ let $\boldsymbol{T}^{(2\ell)}$ be the finite tree/2-CNF obtained by deleting all variables and clauses at a distance larger than 2ℓ from the root.

The tree $\boldsymbol{T}$ approximates $\boldsymbol{\Phi}$ locally in the sense that for any fixed ℓ and any given variable x_i the distribution of the depth-2ℓ neighbourhood of x_i in $\boldsymbol{\Phi}$ converges to $\boldsymbol{T}^{(2\ell)}$ as $n \to \infty$ (in the sense of local weak convergence). Moreover, Gibbs uniqueness posits that under random satisfying assignments of the tree-CNF $\boldsymbol{T}^{(2\ell)}$ the truth value $\boldsymbol{\sigma}_o$ of the root under a random satisfying assignment $\boldsymbol{\sigma}$ decouples from the values $\boldsymbol{\sigma}_{\boldsymbol{T},y}$ of variables $y \in \partial^{2\ell} o$ at distance precisely 2ℓ from o for large enough ℓ. Formally, with $S(\boldsymbol{T}^{(2\ell)})$ the set of satisfying assignments of the 2-CNF $\boldsymbol{T}^{(2\ell)}$, the following is true.

Proposition 2.1 ([5, Proposition 2.2])**.** *We have*

$$\lim_{\ell\to\infty}\mathbb{E}\left[\max_{\tau\in S(\boldsymbol{T}^{(2\ell)})}\left|\mathbb{P}\left[\boldsymbol{\sigma}_o = 1 \mid \boldsymbol{T}^{(2\ell)}, \boldsymbol{\sigma}_{\partial^{2\ell}o} = \tau_{\partial^{2\ell}o}\right] - \mathbb{P}\left[\boldsymbol{\sigma}_o = 1 \mid \boldsymbol{T}^{(2\ell)}\right]\right|\right] = 0. \tag{2.6}$$

2.3. **Approaching the variance.** The proof of the formula (1.1) combines the Gibbs uniqueness property and the local convergence to the Galton-Watson tree with a coupling argument called the 'Aizenman-Sims-Starr scheme' [5]. Unfortunately, this combination does not seem precise enough to get a handle on the limiting distribution of $\log Z(\boldsymbol{\Phi})$ by a long shot. Actually, it is anything but clear how even the *order* of the standard deviation of $\log Z(\boldsymbol{\Phi})$ could be derived along these lines. One specific problem is that the rate of convergence of (2.6) diminishes as d approaches the satisfiability threshold.

To tackle this challenge we devise a combinatorial interpretation of $\log^2 Z(\boldsymbol{\Phi})$. A key idea, which we borrow from spin glass theory [19], is to set up a family of correlated random formulas. Specifically, given integers $M, M' \ge 0$ we construct a correlated pair $(\boldsymbol{\Phi}_1(M, M'), \boldsymbol{\Phi}_2(M, M'))$ of formulas on the variable set $V_n = \{x_1, \ldots, x_n\}$ as follows. Let $(\boldsymbol{a}_i)_{i\ge1}$, $(\boldsymbol{a}'_i)_{i\ge1}$, $(\boldsymbol{a}''_i)_{i\ge1}$ be sequences of mutually independent uniformly random clauses on V_n. Then

$$\boldsymbol{\Phi}_1(M, M') = \boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_M \wedge \boldsymbol{a}'_1 \wedge \cdots \wedge \boldsymbol{a}'_{M'}, \qquad \boldsymbol{\Phi}_2(M, M') = \boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_M \wedge \boldsymbol{a}''_1 \wedge \cdots \wedge \boldsymbol{a}''_{M'}. \tag{2.7}$$

[2]For the sake of tidyness, if the above denominator vanishes we simply let $\hat{\mu}_{x\to a}(\pm1) = \frac{1}{2}$.

Thus, the two formulas share clauses $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M$. Additionally, each contains another M' independent clauses. In particular, $\boldsymbol{\Phi}_1(m,0)$, $\boldsymbol{\Phi}_2(m,0)$ are identical, while $\boldsymbol{\Phi}_1(0,m)$, $\boldsymbol{\Phi}_2(0,m)$ are independent.

Interpolating between these extreme cases offers a promising avenue for computing the variance: given that $\boldsymbol{\Phi}_1(M, m-M)$ and $\boldsymbol{\Phi}_2(M, m-M)$ are satisfiable for all M, we can write a telescoping sum

$$\begin{aligned}
&\log Z(\boldsymbol{\Phi}_1(m,0)) \cdot \log Z(\boldsymbol{\Phi}_2(m,0)) - \log Z(\boldsymbol{\Phi}_1(0,m)) \cdot \log Z(\boldsymbol{\Phi}_2(0,m)) \\
&\quad = \sum_{M=1}^{m} \log Z(\boldsymbol{\Phi}_1(M, m-M)) \cdot \log Z(\boldsymbol{\Phi}_2(M, m-M)) \\
&\qquad\qquad - \log Z(\boldsymbol{\Phi}_1(M-1, m-M+1)) \cdot \log Z(\boldsymbol{\Phi}_2(M-1, m-M+1)).
\end{aligned} \tag{2.8}$$

If we *could* take the expectation on the l.h.s. of (2.8), we would precisely obtain the variance of $\log Z(\boldsymbol{\Phi})$. Moreover, each summand on the r.h.s. amounts to a 'local' change of swapping a shared clause for a pair of independent clauses. Yet we cannot just take the expectation of (2.8), because some $\boldsymbol{\Phi}_h(M, m-M)$ may be unsatisfiable. To remedy this, we will replace $\log Z(\boldsymbol{\Phi})$ by a tamer random variable with the same limiting distribution. Its construction is based on the Unit Clause Propagation algorithm.

2.4. **Unit Clause Propagation.** Employed by all modern SAT solvers as a sub-routine, Unit Clause Propagation is a linear time algorithm that tracks the implications of partial assignments. The algorithm receives as input a 2-CNF Φ along with a set $\mathscr{L}$ of literals. These literals are deemed to be 'true'. The algorithm then pursues direct logical implications, thereby identifying additional 'implied' literals that need to be true so that no clause gets violated. This procedure is outlined in Steps 1–2 of Algorithm 1; the outcome of Steps 1–2 is independent of the order in which literals/clauses are processed.

Input: A 2-CNF Φ along with a set $\mathscr{L}$ of literals deemed true.

1 **while** there exists a clause $a \equiv l \vee \neg l'$ with $l' \in \mathscr{L}$ and $l \notin \mathscr{L}$ **do**

2   add literal l to $\mathscr{L}$;

3 For variables $x \in V(\Phi)$ such that $x \in \mathscr{L}$ or $\neg x \in \mathscr{L}$ let

$$\sigma_x = \begin{cases} 1 & \text{if } x \in \mathscr{L} \text{ and } \neg x \notin \mathscr{L}, \\ -1 & \text{if } \neg x \in \mathscr{L} \text{ and } x \notin \mathscr{L}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathscr{C}$ be the set of all clauses a such that $\sigma_x = 0$ for all $x \in \partial a$ and return $\mathscr{L}, \mathscr{C}, \sigma$;

Algorithm 1: Pessimistic Unit Clause Propagation ('PUC').

Clearly, trouble brews if PUC ends up placing both a literal l and its negation $\neg l$ into the set $\mathscr{L}$. Our 'pessimistic' Unit Clause variant makes no attempt at mitigating such contradictions. Instead, Step 3 just constructs a partial assignment where all conflicting literals are set to a dummy value zero. Additionally, PUC identifies the set $\mathscr{C}$ of *conflict clauses* that contain conflicted variables only.

Now consider a 2-CNF Φ on a set of variables $V(\Phi)$. For each possible literal $l \in \{x, \neg x : x \in V(\Phi)\}$ we run $\mathtt{PUC}(\Phi, \mathscr{L} = \{l\})$. Let $\mathscr{C}(\Phi, \{l\})$ be the set of conflict clauses returned by PUC. Obtain the *pruned formula* $\hat{\Phi}$ from Φ by removing all clauses in $\mathscr{C}(\Phi) = \bigcup_l \mathscr{C}(\Phi, \{l\})$. Then it is easy to verify the following.[3]

Fact 2.2. *For any 2-CNF Φ the pruned 2-CNF $\hat{\Phi}$ is satisfiable.*

Generally, the pruned formula $\hat{\Phi}$ could have far fewer clauses than the original formula Φ. Accordingly, even if Φ is satisfiable the number $Z(\hat{\Phi})$ of satisfying assignments of $\hat{\Phi}$ could dramatically exceed $Z(\Phi)$. However, the following proposition shows that on a random formula, the impact of pruning is modest.

Proposition 2.3. *With probability $1 - o(n^{-1/2})$ we have $|\log Z(\hat{\boldsymbol{\Phi}}) - \log Z(\boldsymbol{\Phi})| \le n^{1/3}$.*

[3]See Section 4.2 for a detailed proof.

2.5. **Variance redux.** The error bound from Proposition 2.3 is tight enough so that towards the proof of Theorem 1.1 it suffices to establish a central limit theorem for $\log Z(\hat{\mathbf{\Phi}})$, i.e., the log of the number of satisfying assignments of the pruned formula. Once again the pivotal task to this end is to compute the variance of $\log Z(\hat{\mathbf{\Phi}})$. Revisiting the telescoping sum (2.8), we obtain the following expression. Recalling (2.7), we write $\hat{\mathbf{\Phi}}_h(M,M') = \widehat{\mathbf{\Phi}_h(M,M')}$ for the formula obtained by pruning $\mathbf{\Phi}_h(M,M')$.

Lemma 2.4. *Let*

$$\Delta(M) = \mathbb{E}\left[\log\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M,m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\cdot\log\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M,m-M))}{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M))}\right)\right], \tag{2.9}$$

$$\Delta'(M) = \mathbb{E}\left[\log\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M+1))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\cdot\log\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M+1))}{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M))}\right)\right]. \tag{2.10}$$

Then $\mathrm{Var}\left[\log Z(\hat{\mathbf{\Phi}})\right] = \sum_{M=1}^{m} \Delta(M) - \Delta'(M)$.

Lemma 2.4 expresses the variance as a sum of local changes. For example, $\mathbf{\Phi}_1(M,m-M)$ is obtained from $\mathbf{\Phi}_1(M-1,m-M)$ by adding a single random clause, namely $\boldsymbol{a}_M$. Thus, $\Delta(M)$ equals the expected change upon addition of a single *shared* clause—modulo the effect of pruning, that is.

But fortunately, on random formulas only a few clauses get pruned w.h.p. In effect, we can express the impact of these random changes neatly in terms of random satisfying assignments of the 'small' formulas $\hat{\mathbf{\Phi}}_h(M-1,m-M)$ that appear in (2.9)–(2.10). Specifically, the quotients in (2.9)–(2.10) boil down to the probabilities that random satisfying assignments of the 'small' formulas survive the extra clause that gets added to obtain the 2-CNFs in the respective numerators. Thus, with $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_y)_{y\in V_n}$ denoting a random satisfying assignment of $\hat{\mathbf{\Phi}}_h(M-1,m-M)$, we obtain the following.

Proposition 2.5. *Let* $1 \le M \le m$. *W.h.p. we have*

$$\frac{Z(\hat{\mathbf{\Phi}}_h(M,m-M))}{Z(\hat{\mathbf{\Phi}}_h(M-1,m-M))} = 1 - \prod_{y\in\partial\boldsymbol{a}_M} \mathbb{P}\left[\boldsymbol{\sigma}_y \neq \mathrm{sign}(y,\boldsymbol{a}_M) \mid \hat{\mathbf{\Phi}}_h(M-1,m-M),\boldsymbol{a}_M\right] + o(1) \quad (h=1,2),$$

$$\frac{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M+1))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))} = 1 - \prod_{y\in\partial\boldsymbol{a}'_{m-M+1}} \mathbb{P}\left[\boldsymbol{\sigma}_y \neq \mathrm{sign}(y,\boldsymbol{a}'_{m-M+1}) \mid \hat{\mathbf{\Phi}}_1(M-1,m-M),\boldsymbol{a}'_{m-M+1}\right] + o(1),$$

$$\frac{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M+1))}{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M))} = 1 - \prod_{y\in\partial\boldsymbol{a}''_{m-M+1}} \mathbb{P}\left[\boldsymbol{\sigma}_y \neq \mathrm{sign}(y,\boldsymbol{a}''_{m-M+1}) \mid \hat{\mathbf{\Phi}}_2(M-1,m-M),\boldsymbol{a}'_{m-M+1}\right] + o(1).$$

2.6. **Local convergence in probability.** To evaluate the expressions from Proposition 2.5 we need to get a grip on the *joint* distribution of the truth values of y under random satisfying assignments of the two correlated formulas $\hat{\mathbf{\Phi}}_h(M-1,m-M)$. To this end we will devise a Galton-Watson tree $\boldsymbol{T}^{\otimes}$ that mimics the *joint* distribution of the local structure of $(\hat{\mathbf{\Phi}}_1(M-1,m-M),\hat{\mathbf{\Phi}}_2(M-1,m-M))$. Subsequently, we will establish Gibbs uniqueness for this Galton-Watson tree to compute the expressions from Proposition 2.5.

The Galton-Watson tree $\boldsymbol{T}$ from Section 2.2 that describes the local topology of the 'plain' random formula $\mathbf{\Phi}$ had one type of variable nodes and four types $(\pm1,\pm1)$ of clause nodes. To approach the correlated pair $(\hat{\mathbf{\Phi}}_1(M,m-M-1),\hat{\mathbf{\Phi}}_2(M,m-M-1))$ we need a Galton-Watson process with three types of variable nodes and a full dozen types of clause nodes. Specifically, there are *shared*, *1-distinct* and *2-distinct* variable nodes. The root o of $\boldsymbol{T}^{\otimes}$ is a shared variable node. The clause node types are (s,s')*-shared*, (s,s') *1-distinct* and (s,s') *2-distinct* for $s,s'\in\{\pm1\}$.

In addition to $d\in(0,2)$ the offspring distributions of $\boldsymbol{T}^{\otimes} = \boldsymbol{T}^{\otimes}_{d,t}$ involve a second parameter $t\in[0,1]$:

- A *shared variable* spawns $\mathrm{Po}(dt/4)$ shared clauses of type (s,s') as well as $\mathrm{Po}(d(1-t)/4)$ 1-distinct clauses of type (s,s') and $\mathrm{Po}(d(1-t)/4)$ 2-distinct clauses of type (s,s') for any $s,s'\in\{\pm1\}$.
- An h*-distinct variable* begets $\mathrm{Po}(d/4)$ h-distinct clauses of type (s,s') for any $s,s'\in\{\pm1\}$ $(h=1,2)$.
- A *shared clause* has precisely one shared variable as its offspring.
- An h*-distinct clause* spawns a single h-distinct variable $(h=1,2)$.

Figure 1 provides an illustration of the tree $\boldsymbol{T}^{\otimes}$. Shared variables/clauses are indicated in red, 1-distinct variables/clauses in green and 2-distinct ones in blue.

From $\boldsymbol{T}^{\otimes}$ we extract a pair $(\boldsymbol{T}_1,\boldsymbol{T}_2)$ of correlated random trees. Specifically, $\boldsymbol{T}_h$ is obtained from $\boldsymbol{T}^{\otimes}$ by deleting all $(3-h)$-distinct variables and clauses. Hence, the parameter t determines how 'similar' $\boldsymbol{T}_1,\boldsymbol{T}_2$ are. Specifically, if $t=1$ then no $\{1,2\}$-distinct clauses exist and thus $\boldsymbol{T}_1, \boldsymbol{T}_2$ are identical. By contrast, if $t=0$ then $\boldsymbol{T}_1, \boldsymbol{T}_2$ are independent copies of the tree $\boldsymbol{T}$ from Section 2.2.

For an integer $\ell \geq 0$ obtain $\boldsymbol{T}^{\otimes,(2\ell)}$, $\boldsymbol{T}_1^{(2\ell)}$, $\boldsymbol{T}_2^{(2\ell)}$ from $\boldsymbol{T}^{\otimes}, \boldsymbol{T}_1, \boldsymbol{T}_2$ by omitting all nodes at a distance greater than 2ℓ from the root o. As in Section 2.2, we can interpret these trees as 2-CNFs, with the type (s,s') of a clause indicating the signs of its parent and child variables. We say that two possible outcomes T, T' of $\boldsymbol{T}^{\otimes,(2\ell)}$ are *isomorphic* if there is a tree isomorphism that preserves the root o as well as all types.

Further, a variable $x \in V_n$ is called a 2ℓ*-instance* of T in $(\hat{\boldsymbol{\Phi}}_1(M,M'), \hat{\boldsymbol{\Phi}}_2(M,M'))$ if there exist isomorphisms ι_h of the 2-CNFs T_h obtained from T by deleting all $(3-h)$-distinct variables/clauses to the depth-2ℓ neighbourhoods $\partial^{\leq 2\ell}_{\hat{\boldsymbol{\Phi}}_h(M,M')} x$ of x in $\hat{\boldsymbol{\Phi}}_h(M,M')$ such that

- the root gets mapped to x, i.e., $\iota_1(o) = \iota_2(o) = x$,
- for any shared variable y of T_1, T_2 the image variables coincide, i.e., $\iota_1(y) = \iota_2(y)$,
- for any shared clauses a of T_1, T_2 the image $\iota_1(a) = \iota_2(a) \in \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M\}$ is a shared clause,
- for any 1-distinct clause a whose parent in T_1 is a shared variable, $\iota_1(a) \in \{\boldsymbol{a}'_1, \ldots, \boldsymbol{a}'_{M'}\}$, and
- for any 2-distinct clause a whose parent in T_2 is a shared variable, $\iota_1(a) \in \{\boldsymbol{a}''_1, \ldots, \boldsymbol{a}''_{M'}\}$.

Let $\boldsymbol{N}^{(2\ell)}(T, (\boldsymbol{\Phi}_1(M,M'), \boldsymbol{\Phi}_2(M,M')))$ be the number of 2ℓ-instances of T in $(\boldsymbol{\Phi}_1(M,M'), \boldsymbol{\Phi}_2(M,M'))$.The following proposition confirms that $\boldsymbol{T}^{\otimes}$ models the local structure of $(\hat{\boldsymbol{\Phi}}_1(M,M'), \hat{\boldsymbol{\Phi}}_2(M,M'))$ faithfully.

Proposition 2.6. *Let $\ell > 0$ be a fixed integer, let $t \in [0,1]$ and suppose that $M \sim tdn/2$ and $M' \sim (1-t)dn/2$. Then w.h.p. for all possible outcomes T of $\boldsymbol{T}^{\otimes,(2\ell)}$ we have $\boldsymbol{N}^{(2\ell)}(T, (\hat{\boldsymbol{\Phi}}_1(M,M'), \hat{\boldsymbol{\Phi}}_2(M,M'))) \sim n\mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell)} \cong T\right]$.*

2.7. **Correlated Belief Propagation.** Now that we have a branching process description of our pair of correlated formulas the next step is to run BP on the random trees $(\boldsymbol{T}_1, \boldsymbol{T}_2)$ to find the joint distribution of the truth values $\boldsymbol{\sigma}_{\boldsymbol{T}_1^{(2\ell)},o}, \boldsymbol{\sigma}_{\boldsymbol{T}_2^{(2\ell)},o}$ assigned to the root. Hence, let

$$\boldsymbol{\mu}^{(2\ell)} = \left(\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{T}_1^{(2\ell)},o} = 1 \mid \boldsymbol{T}^{\otimes}\right], \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{T}_2^{(2\ell)},o} = 1 \mid \boldsymbol{T}^{\otimes}\right]\right) \in (0,1)^2. \tag{2.11}$$

Since BP is exact on trees, we could calculate these marginals by iterating (2.2)–(2.4) for 2ℓ steps, starting from all-uniform messages. But our objective is not merely to calculate the marginals of a specific pair of trees, but the *distribution* of the vector (2.11) for a random $\boldsymbol{T}^{\otimes}$. Fortunately, due to the Markovian nature of the Galton-Watson tree $\boldsymbol{T}^{\otimes}$, the bottom-up BP computation on a random tree can be expressed by a fixed point iteration on the space of probability distributions on $\mathbb{R}^2$. The appropriate operator is the $\mathtt{logBP}^{\otimes}_{d,t}$-operator from (1.3). To be precise, that operator expresses the updates of the log-likelihood ratios of the BP messages from (2.3)–(2.4). Thus, let

$$\mathfrak{t} : (z_1, z_2) \in \mathbb{R}^2 \mapsto ((1+\tanh(z_1/2))/2, (1+\tanh(z_2/2))/2) \in (0,1)^2$$

be the function that maps log-likelihood ratios back to probabilities. Furthermore, for a probability measure $\rho \in \mathscr{P}(\mathbb{R}^2)$ let $\mathfrak{t}(\rho)$ be the pushforward probability measure on $(0,1)^2$.[4]

Proposition 2.7. *Let $\rho^{(0)}_{d,t} \in \mathscr{P}(\mathbb{R}^2)$ be the atom at the origin and let $\rho^{(\ell)}_{d,t} = \mathtt{logBP}^{\otimes}_{d,t}(\rho^{(\ell-1)}_{d,t})$. Then $\boldsymbol{\mu}^{(2\ell)}$ has distribution $\mathfrak{t}(\rho^{(\ell)}_{d,t})$.*

We employ the contraction method to show that the sequence $(\rho^{(\ell)}_{d,t})_{\ell \geq 1}$ of measures converges.

Proposition 2.8. *There exists a unique $\rho_{d,t} \in \mathscr{P}(\mathbb{R}^2)$ that satisfies (1.5) and $\lim_{\ell\to\infty} \rho^{(\ell)}_{d,t} = \rho_{d,t}$ weakly.*

Furthermore, the Gibbs uniqueness property (2.6) extends to $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$.

Corollary 2.9. *For all $t \in [0,1]$ and $h = 1,2$ we have*

$$\lim_{\ell\to\infty} \mathbb{E}\left[\max_{\tau \in S(\boldsymbol{T}_h^{(2\ell)})} \left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{T}_h^{(2\ell)},o} = 1 \mid \boldsymbol{T}^{\otimes}, \boldsymbol{\sigma}_{\boldsymbol{T}_h^{(2\ell)},\partial^{2\ell} o} = \tau_{\partial^{2\ell} o}\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{T}_h^{(2\ell)},o} = 1 \mid \boldsymbol{T}^{\otimes}\right]\right|\right] = 0. \tag{2.12}$$

Combining Propositions 2.7 and 2.8 and Corollary 2.9, we are now in a position to pinpoint the joint marginals of $\hat{\boldsymbol{\Phi}}_1(M,M'), \hat{\boldsymbol{\Phi}}_2(M,M')$. Formally, let

$$\pi_{\hat{\boldsymbol{\Phi}}_1(M,M'),\hat{\boldsymbol{\Phi}}_2(M,M')} = \frac{1}{n}\sum_{i=1}^{n} \delta_{(\mathbb{P}[\boldsymbol{\sigma}_{\hat{\boldsymbol{\Phi}}_1(M,M'),x_i}=1 \mid \hat{\boldsymbol{\Phi}}_1(M,M')], \mathbb{P}[\boldsymbol{\sigma}_{\hat{\boldsymbol{\Phi}}_2(M,M'),x_i}=1 \mid \hat{\boldsymbol{\Phi}}_2(M,M')])} \in \mathscr{P}([0,1]^2)$$

[4] That is, for a measurable $\mathfrak{A} \subseteq (0,1)^2$ we have $\mathfrak{t}(\rho)(\mathfrak{A}) = \rho(\mathfrak{t}^{-1}(\mathfrak{A}))$.

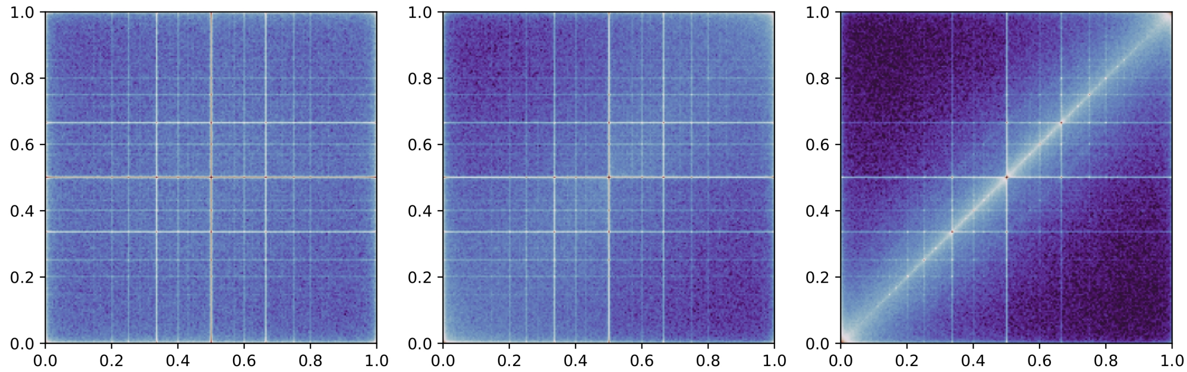


FIGURE 2. The distributions $\mathfrak{t}(\rho_{d,t})$ for $d = 1.9$ and $t = 0.1, 0.5, 0.9$.

be the empiricial distribution of the joint marginals of $\hat{\mathbf{\Phi}}_1(M, M')$ and $\hat{\mathbf{\Phi}}_2(M, M')$, which we need to know to evaluate the expressions from Proposition 2.5. Furthermore, denote by $W_1(\cdot,\cdot)$ the Wasserstein L^1-distance of two probability measures on $[0,1]^2$.

Corollary 2.10. *For any $t \in [0,1]$ and any $M \sim tnd/2$, $M' \sim (1-t)dn/2$ we have*

$$\mathbb{E}\left[W_1\left(\pi_{\hat{\mathbf{\Phi}}_1(M,M'),\hat{\mathbf{\Phi}}_2(M,M')}, \mathfrak{t}(\rho_{d,t})\right)\right] = o(1).$$

Finally, combining Proposition 2.5 with Corollary 2.10, we obtain the variance of $\log Z(\hat{\mathbf{\Phi}})$.

Corollary 2.11. *With $\eta(d)^2$ from (1.7) we have $\eta(d) > 0$ and $\mathrm{Var} \log Z(\hat{\mathbf{\Phi}}) \sim m\eta_d^2$.*

Because the proof of Proposition 2.8 is based on a contraction argument, for any d, t the distribution $\rho_{d,t}$ can be approximated effectively within any given accuracy via a fixed point iteration. Figure 2 displays approximations to $\mathfrak{t}(\rho_{d,t})$ for different values of t and shows how correlations between the two coordinates of the random vector increase with t (brighter diagonal).

2.8. **The central limit theorem.** With the variance computation done, we have now overcome the greatest hurdle en route to Theorem 1.1. Indeed, to obtain the desired asymptotic normality we just need to combine the techniques from the variance computation with a generic martingale central limit theorem.

To this end we set up a filtration $(\mathfrak{F}_{n,M})_{0 \le M \le m_n}$ by letting $\mathfrak{F}_{n,M}$ be the σ-algebra generated by $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M$. Hence, conditioning on $\mathfrak{F}_{n,M}$ amounts to conditioning on $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M$, while averaging on the remaining clauses $\boldsymbol{a}_{M+1}, \ldots, \boldsymbol{a}_m$. The conditional expectations

$$\boldsymbol{Z}_{n,M} = m^{-1/2}\mathbb{E}\left[\log Z(\hat{\mathbf{\Phi}}) \mid \mathfrak{F}_{n,M}\right] \tag{2.13}$$

then form a Doob martingale. Let $\boldsymbol{X}_{n,M} = \boldsymbol{Z}_{n,M} - \boldsymbol{Z}_{n,M-1}$ be the martingale differences.

Proposition 2.12. *For all $0 < d < 2$ the martingale (2.13) satisfies*

$$\lim_{n\to\infty} \mathbb{E}\left[\max_{1 \le M \le m} |\boldsymbol{X}_{n,M}|\right] = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \mathbb{E}\left|\eta(d)^2 - \sum_{M=1}^{m} \boldsymbol{X}_{n,M}^2\right| = 0. \tag{2.14}$$

Thanks to pruning, the first condition from (2.14) is easily checked. Furthermore, the steps that we pursued towards the proof of Corollary 2.11, i.e., the variance calculation, also imply the second condition without further ado. Finally, as (2.14) demonstrates that the marginal differences are small and that the variance process converges to a deterministic limit, Theorem 1.1 follows from the general martingale central limit theorem from [28].

3. DISCUSSION

The hunt for satisfiability thresholds of random constraint satisfaction problems was launched by the experimental work of Cheeseman, Kanefsky and Taylor [17]. The 2-SAT threshold was the first one to be caught [20, 32]. Subsequent successes include the 1-in-k-SAT threshold [3] and the k-XORSAT threshold [27, 43]. Furthermore, Friedgut [30] proved the existence of non-uniform (i.e., n-dependent) satisfiability thresholds in considerable generality. The plot thickened when physicists employed a compelling but non-rigorous technique called the cavity method to 'predict' the exact satisfiability thresholds of many further problems, including the k-SAT problem for $k \geq 3$ [37]. A line of rigorous work [6, 8, 23] culminated in the verification of this physics prediction for large k [25].

Even though the satisfiability threshold of random 2-SAT was determined already in the 1990s, the problem continued to receive considerable attention. For example, Bollobás, Borgs, Chayes, Kim and Wilson [15] investigated the scaling window around the satisfiability threshold, a point on which a recent contribution by Dovgal, de Panafieu and Ravelomanana elaborates [26]. Abbe and Montanari [2] made the first substantial step towards the study of the number of satisfying assignments that $\frac{1}{n}\log Z(\mathbf{\Phi})$ converges in probability to a deterministic limit $\varphi(d)$ for Lebesgue-almost all $d \in (0,2)$. However, their techniques do not reveal the value $\varphi(d)$. Moreover, Montanari and Shah [39] obtain a 'law-of-large-numbers' estimate of the number of assignments that satisfy all but $o(n)$ clauses for $d < 1.16$. Finally, the aforementioned article of Achlioptas et al. [5] verifies the prediction from [38] as to the number of satisfying assignments for all $d < 2$. The main result of the present paper refines these results considerably by establishing a central limit theorem.

For random k-CNFs with $k \geq 3$ an upper bound on the number of satisfying assignments can be obtained via the interpolation method from mathematical physics [42]. This bound matches the predictions of the cavity method [36]. However, no matching lower bound is currently known. The precise physics prediction called the 'replica symmetric solution' has only been verified for 'soft' versions of random k-SAT where unsatisfied clauses are penalised but not strictly forbidden, and for clause-to-variable ratios well below the satisfiability threshold [39, 41, 47].

Random CSPs such as random k-XORSAT or random k-NAESAT that exhibit stronger symmetry properties than random k-SAT tend to be amenable to the method of moments [6].[5] Therefore, more is known about their number of solutions. For example, due to the inherent connection to linear algebra, the number of satisfying assignments of random k-XORSAT formulas is known to concentrate on a single value right up to the satisfiability threshold [11, 27, 43]. Furthermore, in random k-NAESAT, random graph colouring and several related problems the logarithm of the number of solutions superconcentrates, i.e., has only bounded fluctuations for constraint densities up to the so-called condensation threshold, a phase transition that shortly precedes the satisfiability threshold [12, 21, 44]. The same is true of random k-SAT instances with regular literal degrees [24]. A further example is the symmetric perceptron [1], where the number of solutions superconcentrates but the limiting distribution is a log-normal with bounded variance. Going beyond the condensation transition, Sly, Sun and Zhang [46] proved that the number of satisfying assignments of random regular k-NAESAT formulas matches the '1-step replica symmetry breaking' prediction from physics.

Apart from the superconcentration results for symmetric problems from [12, 24, 21, 44], the limiting distribution of the logarithm of the number of solutions has not been known in any random constraint satisfaction problem. In particular, Theorem 1.1 is the first central limit theorem for this quantity in any random CSP. We expect that the technique developed in the present work, particularly the use of two correlated random instances in combination with spatial mixing, can be extended to other problems. The present use of correlated instances is inspired by the work of Chen, Dey and Panchenko [19] on the p-spin model from mathematical physics, a generalisation of the famous Sherrington-Kirkpatrick model. That said, on a technical level the present use of correlated instances is quite different from the approach from [19]. Specifically, while here we construct correlated 2-CNFs that share a specific fraction of their clauses and employ a martingale central limit theorem, Chen, Dey and Panchenko combine a continuous interpolation of two mixed p-spin Hamiltonians with Stein's method.

A further line of work deals with central limit theorems for random optimisation problems. Cao [16] provided a general framework based on the 'objective method' [9]. Unfortunately, the conditions of Cao's theorem tend to be unwieldy for MAX CSP problems with hard constraints. Recent work of Kreačič [34] and Glasgow, Kwan, Sah, Sawhney [31] on the matching number therefore instead resorts to the use of stochastic differential equations. A promising question for future work might be whether the present method of considering correlated instances might extend to random optimisation problems.

Organisation. In the rest of the paper we carry out the strategy from Section 2 in detail. After some preliminaries in Section 4, we prove Proposition 2.3 in Section 5. Subsequently Section 6 deals with the proof of Proposition 2.6. The proof of Proposition 2.5 follows in Section 7. Moreover, Section 8 contains the proof of Proposition 2.8. Further,

[5] Formally, by 'symmetry' we mean that the empirical distribution of the marginals of random solutions converges to an atom; cf. [22].

in Section 9 we prove Proposition 2.7 and Section 11 contains the proofs of Proposition 2.12 and Corollary 2.11. Finally, in Section 12 we complete the proof of Theorem 1.1.

4. Preliminaries and notation

4.1. **Boolean formulas.** A *2-SAT formula* or *2-CNF* Φ consists of a finite set $V(\Phi)$ of propositional variables and another set $F(\Phi)$ of clauses. Unless specified otherwise, we assume that each clause contains two distinct variables.

For a clause $a \in F(\Phi)$ we denote by $\partial a = \partial_\Phi a$ the set of variables that appear in clause a. Similarly, for a variable $x \in V(\Phi)$ let $\partial x = \partial_\Phi x$ signify the set of clauses in which x appears. Thus, the formula Φ induces a bipartite graph on variables and clauses, the so-called *incidence graph* of Φ. Further, the shortest path metric on the incidence graph induces a metric on the variables and clauses of Φ. Accordingly, for a variable or clause u let $\partial^\ell u = \partial^\ell_\Phi u$ be the set of all nodes at a distance precisely ℓ from u. Moreover, let $\partial^{\le \ell} u = \partial^{\le \ell}_\Phi u$ be the sub-formula of Φ obtained by deleting all clauses and variables at a distance greater than ℓ from u. In other words, $\partial^{\le \ell} u$ is the depth-ℓ neighbourhood of u.

We encode the Boolean values 'true' and 'false' as ± 1. Accordingly, let $S(\Phi) \subseteq \{\pm 1\}^{V(\Phi)}$ be the set of satisfying assignments of Φ and let $Z(\Phi) = |S(\Phi)|$. Further, $\mathrm{sign}(x,a) = \mathrm{sign}_\Phi(x,a) \in \{\pm 1\}$ denotes the sign with which variable x appears in clause a, i.e., $\mathrm{sign}(x,a) = 1$ if x appears in a positively and $\mathrm{sign}(x,a) = -1$ if a contains the negation $\neg x$. Finally, for a literal $l \in \{x, \neg x\}$ we let $|l| = x$ denote the underlying Boolean variable.

Assuming $S(\Phi) \ne \emptyset$ let

$$\mu_\Phi(\sigma) = \mathbb{1}\{\sigma \in S(\Phi)\}/Z(\Phi), \qquad (\sigma \in \{\pm 1\}^{V(\Phi)}) \tag{4.1}$$

be the uniform distribution on $S(\Phi)$. We write $\boldsymbol{\sigma} = \boldsymbol{\sigma}_\Phi = (\boldsymbol{\sigma}_{\Phi,x})_{x \in V(\Phi)} \in \{\pm 1\}^{V(\Phi)}$ for a sample from μ_Φ, i.e., a uniformly random satisfying assignment of Φ.

In contrast to k-SAT for $k \ge 3$, the 2-SAT problem can be solved in polynomial time. This is because a 2-SAT instance is unsatisfiable if and only if it contains a peculiar sub-formula called a bicycle. To be precise, let Φ be a CNF with clauses of length one or two. A *bicycle* of Φ is an alternating sequence $l_0, a_1, l_1, a_2, \dots, a_k, l_k$ of literals $l_0, \dots, l_k$ and clauses $a_1, \dots, a_k \in F(\Phi)$ such that

- **BIC1:** $l_0 = l_k$,
- **BIC2:** $l_i = \neg l_0$ for some $0 < i < k$ and
- **BIC3:** $a_i \equiv \neg l_{i-1} \vee l_i \equiv l_{i-1} \to l_i$.

(Observe that a clause a comprising only a single literal l is logically equivalent to $l \vee l \equiv \neg l \to l$.) Hence, the bicycle consists of clauses that are logically equivalent to a chain of implications $l_0 \to \neg l_0 \to l_0$.

Fact 4.1 ([10])**.** *A CNF Φ with clauses of lengths one or two is unsatisfiable iff Φ contains a bicycle.*

4.2. **Unit Clause Propagation.** The PUC algorithm (Algorithm 1) takes as input a CNF Φ along with an initial set $\mathscr{L}_0$ of literals. PUC outputs a set $\mathscr{L} = \mathscr{L}(\Phi, \mathscr{L}_0) \supseteq \mathscr{L}_0$ of literals. Let

$$\mathscr{V}(\Phi, \mathscr{L}_0) = \{|l| : l \in \mathscr{L}(\Phi, \mathscr{L}_0)\}$$

be the set of underlying variables. In addition to $\mathscr{L}(\Phi, \mathscr{L}_0)$, PUC also outputs a partial assignment

$$\sigma = \sigma_{\Phi, \mathscr{L}_0} : \mathscr{V}(\Phi, \mathscr{L}_0) \to \{0, \pm 1\}$$

that sets each $x \in \mathscr{V}$ either to a truth value ± 1 or to the dummy value 0. Let

$$\mathscr{V}_0(\Phi, \mathscr{L}_0) = \{x \in \mathscr{V}(\Phi, \mathscr{L}_0) : \sigma_{\Phi, \mathscr{L}_0, x} = 0\}$$

be the set of variables that receive the dummy value. Finally, the algorithm identifies a set $\mathscr{C}(\Phi, \mathscr{L}_0)$ of *conflict clauses*, i.e., clauses a such that $\partial a \subseteq \mathscr{V}_0(\Phi, \mathscr{L}_0)$.

We make a note of a few basic facts about PUC. These remarks apply to any CNF Φ with clauses of length *at most* two. To get started, we say that a literal l' is *implication-reachable* from another literal l if there exists an alternating sequence $l = l_0, a_1, l_1, \dots, a_k, l_k = l'$ of literals l_i and clauses a_i of Φ such that $a_i \equiv \neg l_{i-1} \vee l_i \equiv l_{i-1} \to l_i$ for all $1 \le i \le k$. We call this sequence an *implication chain* from l to l'. Observe that a unit clause (clause of length one) comprising a single literal l is equivalent to the implication $\neg l \to l \equiv l \vee l$. Furthermore, if l' is implication-reachable from l, then $\neg l$ is implication-reachable from $\neg l'$. Indeed, if $l = l_0, a_1, l_1, \dots, a_k, l_k = l'$ is an implication chain from l to l', then its contraposition

$$\neg l' = \neg l_k, a_k, \neg l_{k-1}, \dots, \neg l_1, a_1, \neg l$$

is an implication chain from $\neg l'$ to $\neg l$.

Lemma 4.2. *Let Φ be a CNF with clauses of length at most two and let $\mathscr{L}_0$ be a set of literals of Φ. Then $\mathscr{L}(\Phi, \mathscr{L}_0)$ is the set of all literals l' that are implication-reachable from a literal $l \in \mathscr{L}_0$.*

Proof. This is an easy induction on the length of the shortest implication chain from l to l'. □

An immediate consequence of Lemma 4.2 is that the order in which PUC proceeds is irrelevant.

Finally, for the sake of completeness, we carry out the proof of Fact 2.2.

Proof of Fact 2.2. Fix some order $l_1, \ldots, l_v$ of the literals $\{x, \neg x : x \in V(\Phi)\}$. Let σ_i be the assignments produced by PUC on input $(\Phi, \{l_i\})$. We construct an assignment $\sigma : V(\Phi) \to \{0, \pm 1\}$ by proceeding as follows for $i = 1, \ldots, v$. For each variable x such that $\{x, \neg x\} \cap \mathscr{L}(\Phi, l_i) \setminus \bigcup_{1 \le h < i} \mathscr{L}(\Phi, l_h) \neq \emptyset$ let $\sigma_x = \sigma_i(x)$. We claim that for each clause $a \in \hat{\Phi}$ there is a variable $x \in \partial a$ such that $\sigma_x = \mathrm{sign}(x, a)$. Indeed, it is not possible that $\sigma_x = 0$ for all $x \in \partial a$; for otherwise $a \in \mathscr{C}(\Phi, l_i)$ for some i, and thus a would not be present in $\hat{\Phi}$. Thus, there exists $x \in \partial a$ such that $\sigma_x \in \{\pm 1\}$. If $\sigma_x \neq \mathrm{sign}(x, a)$, then PUC would have included the second literal k that appears in a into the set $\mathscr{L}$. Hence, $\sigma_{|k|} = \mathrm{sign}(|k|, a)$, because otherwise a would have been a contradiction and therefore omitted from $\hat{\Phi}$. □

4.3. **Random 2-SAT.** Recall from Section 1.1 that $\mathbf{\Phi}$ denotes the random 2-CNF formula with variables $V_n = \{x_1, \ldots, x_n\}$ and clauses $F_m = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}$, where $m = m_n \sim dn/2$ for a fixed $d > 0$. ***We tacitly assume that*** $0 < d < 2$***, i.e., that we are in the satisfiable regime.***

In the following sections we will need estimates of the sizes of the sets $|\mathscr{V}(\mathbf{\Phi}, \mathscr{L})|$, $|\mathscr{C}(\mathbf{\Phi}, \mathscr{L})|$ produced by PUC on the random formula $\mathbf{\Phi}$ for singletons $\mathscr{L}$. Thus, suppose we start the PUC algorithm from an initial literal $\mathscr{L} = \{l\}$. Since the ensuing chain of implications traced by PUC is stochastically dominated by a sub-critical branching process (for $d < 2$), we obtain the following bound.

Lemma 4.3 ([5, Claim 6.8])**.** *For any literal l and every $t > 8/(2-d)$ we have*

$$\mathbb{P}\left[|\mathscr{V}(\mathbf{\Phi}, \{l\})| > t\right] \le (2 + o(1)) \exp(-dt/40).$$

Corollary 4.4. *With probability $1 - o(n^{-2})$ we have*

$$\max_{1 \le i \le n} |\mathscr{V}(\mathbf{\Phi}, \{x_i\})| + |\mathscr{V}(\mathbf{\Phi}, \{\neg x_i\})| \le \log^2 n.$$

Proof. This is an immediate consequence of Lemma 4.3. □

Finally, the following statement estimates the probability that a random formula is unsatisfiable.

Lemma 4.5. *We have $\mathbb{P}\left[\mathbf{\Phi} \text{ is unsatisfiable}\right] \le n^{o(1)-1}$.*

Proof. This follows from Fact 4.1 and [5, Claim 6.9]. □

Recall the Galton-Watson tree $\boldsymbol{T}$ from Section 2.2. The following lemma shows that $\boldsymbol{T}$ mimics the local structure of the 'plain' random formula with n' variables and m' independent random 2-clauses. Also recall that $\partial_{\Phi}^{\le 2\ell} x$ denotes the sub-formula of Φ comprising all clauses and variables at distance at most 2ℓ from x.

Lemma 4.6 ([5, p. 15])**.** *Let $\ell \ge 0$ be an integer and let T be a possible outcome of $\boldsymbol{T}^{(2\ell)}$. Let $\mathbf{\Phi}_0$ be a random 2-CNF with $n' \sim n$ variables and $m' \sim dn/2$ clauses. Then w.h.p. the number $\boldsymbol{N}^{(2\ell)}(T, \mathbf{\Phi}_0)$ of variables x_i of $\mathbf{\Phi}_0$ such that $\partial_{\mathbf{\Phi}_0}^{\le 2\ell} x_i \cong T^{2\ell}$ satisfies*

$$\boldsymbol{N}^{(2\ell)}(T, \mathbf{\Phi}_0) = n\mathbb{P}\left[\boldsymbol{T}^{(2\ell)} \cong T\right] + o(n).$$

As a final preparation we need an upper bound on the maximum variable degree.

Lemma 4.7. *With probability $1 - o(n^{-10})$ the degree of any variable node x_i, $i = 1, \ldots, n$ in $\mathbf{\Phi}$ is bounded by $\log^2 n$.*

Proof. The number of clauses that contain a given variable x_i has distribution $\mathrm{Bin}(m, 2/n)$. Therefore, the assertion follows from the Chernoff bound. □

Corollary 4.8. *With probability $1 - o(n^{-10})$ the degree of any variable node x_i, $i = 1, \ldots, n$ in $(\mathbf{\Phi}_1, \mathbf{\Phi}_2)$ is bounded by $\log^2 n$.*

Proof. Since $\mathbf{\Phi}_1, \mathbf{\Phi}_2$ separately are distributed as $\mathbf{\Phi}$, the assertion follows from Lemma 4.7 and the union bound. □

4.4. **Convergence of probability measures.** For a measurable subset Ω of Euclidean space $\mathbb{R}^k$ we let $\mathscr{P}(\Omega)$ denote the space of all probability distributions on Ω equipped with the Borel σ-algebra. Moreover, for $p \geq 1$ we define $\mathscr{W}_p(\Omega)$ to be the set of all $\mu \in \mathscr{P}(\Omega)$ such that $\int_\Omega \|x\|_2^p \mathrm{d}\mu(x) < \infty$. We equip $\mathscr{W}_p(\Omega)$ with the Wasserstein metric

$$W_p(\mu, \mu') = \inf_{\boldsymbol{X},\boldsymbol{X}'} \mathbb{E}\left[\|\boldsymbol{X} - \boldsymbol{X}'\|_2^p\right]^{1/p} \qquad (\mu, \mu' \in \mathscr{W}_p(\Omega)), \tag{4.2}$$

where the infimum is taken over all pairs of random variables $\boldsymbol{X}, \boldsymbol{X}'$ that are defined on some common probability space such that $\boldsymbol{X}$ has distribution μ and $\boldsymbol{X}'$ has distribution μ'.

The infimum in (4.2) is attained for any μ, μ'. Random vectors $\boldsymbol{X}, \boldsymbol{X}'$ for which the infimum is attained are called *optimal couplings*. Such optimal couplings exist for all μ, μ' [14].

The spaces $(\mathscr{W}_p(\Omega), W_p)$ are complete metric spaces [14]. Finally, convergence in $(\mathscr{W}_p(\Omega), W_p)$ implies weak convergence of the corresponding probability measures.

For a measure $\rho \in \mathscr{P}(\Omega)$ and a measurable function $f : \Omega \to \Omega'$ from Ω to another probability space Ω' we denote by $f(\rho)$ the pushforward measure of ρ. Thus, the measure $f(\rho)$ that assigns mass $\rho(f^{-1}(A))$ to measurable $A \subseteq \Omega'$.

Throughout the paper we let

$$\mathfrak{t} : \mathbb{R}^2 \to (0,1)^2, \mathfrak{t}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1+\tanh(x_1/2)}{2} \\ \frac{1+\tanh(x_2/2)}{2} \end{pmatrix}, \qquad \mathfrak{l} : (0,1)^2 \to \mathbb{R}^2, \mathfrak{l}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \log \frac{x_1}{1-x_1} \\ \log \frac{x_2}{1-x_2} \end{pmatrix}.$$

5. Proof of Proposition 2.3

In this section we estimate the difference between the number of satisfying assignments of the pruned random formula $\hat{\boldsymbol{\Phi}}$ and the original formula $\boldsymbol{\Phi}$. We begin with a basic observation about the Unit Clause Propagation algorithm, and then estimate the number of clauses that the pruning process removes. Apart from proving Proposition 2.3, the considerations in this section also pave the way for the proof of the variance formula in Section 11.

5.1. **Tracing Unit Clause Propagation.** For a 2-CNF Φ and a set of literals $\mathscr{L}_0$ consider a set of conflict clauses $\mathscr{C} = \mathscr{C}(\Phi, \mathscr{L}_0)$ that PUC produces along with a set $\mathscr{V} = \mathscr{V}(\Phi, \mathscr{L}_0)$ of conflict variables. Let $\Phi - \mathscr{C}$ be the formula obtained from Φ by deleting the clauses from $\mathscr{C}$. Clearly $Z(\Phi) \leq Z(\Phi - \mathscr{C})$. Conversely, the following lemma puts a bound on how much bigger $Z(\Phi - \mathscr{C})$ may be.

Lemma 5.1. *Assume that Φ is a satisfiable 2-CNF. For any set $\mathscr{L}_0$ of literals we have*

$$Z(\Phi - \mathscr{C}(\Phi, \mathscr{L}_0)) \leq 2^{|\mathscr{V}(\Phi, \mathscr{L}_0)| \cdot \mathbb{1}\{\mathscr{C}(\Phi, \mathscr{L}_0) \neq \emptyset\}} Z(\Phi). \tag{5.1}$$

Towards the proof of Lemma 5.1 let $\mathscr{L} = \mathscr{L}(\Phi, \mathscr{L}_0)$ be the final set of literals that PUC produces. Moreover, let $\sigma : \mathscr{V} \to \{0, \pm 1\}$ be the function that PUC outputs and let $\mathscr{V}_0 = \{x \in \mathscr{V} : \sigma_x = 0\}$. Further, let Φ_0 be a CNF with variable set $\mathscr{V}$ that contains the following clauses:

(i) any clause $a \in F(\Phi)$ with $\partial a \subseteq \mathscr{V}$,
(ii) a unit clause l for every literal l with $|l| \in \mathscr{V}$ such that Φ contains a clause $a \equiv l \vee l'$ with $|l'| \not\subseteq \mathscr{V}$.

Thus, Φ_0 contains clauses of length one or two.

Claim 5.2. *The formula Φ_0 possesses a satisfying assignment τ such that $\tau_x = \sigma_x$ for all $x \in \mathscr{V} \setminus \mathscr{V}_0$.*

Proof. Obtain a formula Φ_1 by adding to Φ_0 a unit clause x for every variable $x \in \mathscr{V}$ with $\sigma_x = 1$ and a unit clause $\neg x$ for every $x \in \mathscr{V}$ with $\sigma_x = -1$. Then we just need to show that Φ_1 is satisfiable.

Assume otherwise. Then by Fact 4.1 Φ_1 contains a bicycle $l_0, a_1, l_1, a_2, \ldots, a_k, l_k$. This bicycle is logically equivalent to an implication chain

$$l_0 \to l_1 \to \cdots \to \neg l_0 \to \cdots \to l_{k-1} \to l_k = l_0. \tag{5.2}$$

The contraposition of this chain reads

$$\neg l_0 = \neg l_k \to \neg l_{k-1} \to \cdots \to l_0 \to \cdots \to \neg l_1 \to \neg l_0. \tag{5.3}$$

Since Φ is satisfiable, Fact 4.1 shows that the bicycle (5.2) cannot be contained in Φ. Therefore, the bicycle contains a unit clause $l_i \in F(\Phi_1) \setminus F(\Phi)$ for some $1 \leq i \leq k$. Hence, the constructions of Φ_0 and Φ_1 ensure that $l_i \in \mathscr{L}(\Phi, \mathscr{L}_0)$. Indeed, letting $\mathscr{U}$ be the the set of all literals l_i that appear in (5.2) as unit clauses, we obtain $\mathscr{U} \subseteq \mathscr{L}(\Phi, \mathscr{L}_0)$.

We claim that in fact $l_0, \ldots, l_k \in \mathscr{L}(\Phi, \mathscr{L}_0)$. To see this, pick any $0 \leq j \leq k$ such that l_j does not appear as a unit clause in Φ_1. Define $l_{-i} = l_{k-i}$ for $0 \leq i < k$ and let $1 - k \leq i < j$ be the largest index such that $l_i \in \mathscr{U}$. Then

Φ contains the implication chain $l_i \to \cdots \to l_j$. Therefore, Lemma 4.2 implies that $l_j \in \mathscr{L}(\Phi, \mathscr{L}_0)$. Analogously, considering the contraposition (5.3), we conclude that the negations of the literals $l_0, \ldots, l_k$ belong to $\mathscr{L}(\Phi, \mathscr{L}_0)$. In summary,

$$l_0, \neg l_0, \ldots, l_k, \neg l_k \in \mathscr{L}(\Phi, \mathscr{L}_0). \tag{5.4}$$

But (5.4) implies that $|l_0|, \ldots, |l_k| \in \mathscr{V}_0$. Consequently, none of these literals belongs to a unit clause $u \in F(\Phi_1) \setminus F(\Phi_0)$. Furthermore, none of the literals $l_i, \neg l_i$ belongs to a unit clause $a \in F(\Phi_0) \setminus F(\Phi)$. This is because if Φ contains a clause $l_i \vee l'$ or $\neg l_i \vee l'$ and $l_i, \neg l_i \in \mathscr{L}(\Phi, \mathscr{L}_0)$, then PUC added l' to $\mathscr{L}(\Phi, \mathscr{C}_0)$ as well. Thus, we conclude that the bicycle (5.2) consists of clauses of Φ only. But by Fact 4.1 this contradicts the fact that Φ is satisfiable. □

Proof of Lemma 5.1. Clearly, if $\mathscr{C}(\Phi, \mathscr{L}_0) = \emptyset$, the statement is true. Hence, assume that $\mathscr{C}(\Phi, \mathscr{L}_0) \neq \emptyset$ and let τ be a satisfying assignment of Φ_0 from Claim 5.2. Consider a satisfying assignment χ of $\Phi - \mathscr{C}$ and let $\chi' : V(\Phi) \setminus \mathscr{V} \to \{\pm 1\}$ be the restriction of χ to $V(\Phi) \setminus \mathscr{V}$. We extend χ' to a satisfying assignment χ'' of Φ by letting

$$\chi''_x = \mathbb{1}\{x \in \mathscr{V}\}\tau_x + \mathbb{1}\{x \notin \mathscr{V}\}\chi'_x;$$

clearly, χ'' satisfies all clauses a such that $\partial a \cap \mathscr{V} = \emptyset$, because all these clauses are contained in $\Phi - \mathscr{C}$. Moreover, χ'' satisfies all a such that $\partial a \subseteq \mathscr{V}$, because these clauses belong to Φ_0. Further, if $a = l \vee l'$ is a clause such that $|l| \in \mathscr{V}$ but $|l'| \not\subseteq \mathscr{V}$, then $\tau_{|l|} = \sigma_{|l|}$. Since $|l'| \not\subseteq \mathscr{V}$, this means that $\sigma_x = \mathrm{sign}(|l|, a)$, as otherwise PUC would have added l' to $\mathscr{L}$. Therefore, $\chi''_{|l|} = \tau_{|l|} = \sigma_{|l|}$ satisfies a. Since the map $\chi \mapsto \chi'$ only discards the values of the variables in $\mathscr{V}$, we obtain the bound (5.1). □

5.2. **Cycles in random formulas.** To prove Proposition 2.3 we need a good estimate of the total number of clauses that will be removed from $\boldsymbol{\Phi}$ to obtain $\hat{\boldsymbol{\Phi}}$. This estimate is provided by the following lemma.

Lemma 5.3. *Fix any $\delta > 0$. With probability $1 - o(n^{-1})$ the number of literals l such that $\mathscr{C}(\boldsymbol{\Phi}, \{l\}) \neq \emptyset$ is smaller than n^{δ}.*

Proof. Let $\boldsymbol{N}$ be the number of literals l such that $\mathscr{C}(\boldsymbol{\Phi}, \{l\}) \neq \emptyset$. We are going to show that for any fixed (i.e., n-independent) $\ell \geq 1$ for large enough n we have,

$$\mathbb{E}\left[\prod_{i=1}^{\ell}(\boldsymbol{N} - i + 1)\right] \leq \left(\ell \log^3 n\right)^{\ell}. \tag{5.5}$$

Providing $\ell \geq 2/\delta$ and n is sufficiently large, Markov's inequality then shows that

$$\mathbb{P}\left[\boldsymbol{N} \geq n^{\delta}\right] \leq \mathbb{P}\left[\prod_{i=1}^{\ell}(\boldsymbol{N} - i + 1) \geq (n^{\delta}/2)^{\ell}\right] \leq \left(\frac{2\ell \log^3 n}{n^{\delta}}\right)^{\ell} = o(n^{-1}),$$

which implies the assertion.

Thus, we are left to prove (5.5). By symmetry it suffices to bound the probability of the event

$$\mathfrak{E} = \bigcap_{i=1}^{\ell}\{\mathscr{C}(\boldsymbol{\Phi}, \{x_i\}) \neq \emptyset\}$$

that PUC will produce at least one conflict clause from each of the literals $x_1, \ldots, x_\ell$; then

$$\mathbb{E}\left[\prod_{i=1}^{\ell}(\boldsymbol{N} - i + 1)\right] \leq (2n)^{\ell}\,\mathbb{P}\left[\mathfrak{E}\right]. \tag{5.6}$$

In order to estimate the probability of $\mathfrak{E}$ we are going to launch PUC from the initial set $\mathscr{L} = \{x_1, \ldots, x_\ell\}$. While the order in which the literals and clauses are processed does not affect the ultimate outcome of PUC, for the present analysis we assume that PUC processes the literals one at a time, each time pursuing *all* the clauses $l \vee \neg l'$ that contain the negation of a specific l'. We also presume that the literals are processed in the same order as they get inserted into the set $\mathscr{L}$. In other words, PUC proceeds in breadth-first-search order. Let $\mathfrak{H}_t$ be the history of the execution of PUC up to and including the point where the first t literals and their adjacent clauses have been explored. Formally, $\mathfrak{H}_t$ is the σ-algebra generated by these first t literals that get added to $\mathscr{L}$ and their adjacent clauses.

Lemma 4.3 implies that with probability $1-o(n^{-1})$ the set $\mathscr{L}$ returned by PUC has size at most $L = \ell \log^2 n$. Let $\mathfrak{E}_t$ be the event that at time t we explored a clause that contains two variables from $\mathscr{L}$ and $|\mathscr{L}| \le L$. Moreover, let $\boldsymbol{S} = \sum_t \mathbb{1}\{\mathfrak{E}_t\}$. Let $0 < t_1 < \ldots < t_\ell \le L$ be distinct time steps. Then

$$\mathbb{P}[\mathfrak{E}] \le \mathbb{P}[\boldsymbol{S} \ge \ell] \le \sum_{t_1,\ldots,t_\ell} \mathbb{P}\left[\bigcap_{i=1}^{\ell} \mathfrak{E}_{t_i}\right] = \sum_{t_1,\ldots,t_\ell} \prod_{i=1}^{\ell} \mathbb{P}\left[\mathfrak{E}_{t_i} \middle| \bigcap_{j=1}^{i-1} \mathfrak{E}_{t_j}\right]. \tag{5.7}$$

To bound the r.h.s. of (5.7) we will estimate the probability of $\mathfrak{E}_{t+1}$ given the history $\mathfrak{H}_t$ of the process up to time t, showing that for all $t \ge 0$,

$$\mathbb{P}[\mathfrak{E}_{t+1}|\mathfrak{H}_t] \le \frac{L}{n}. \tag{5.8}$$

In fact the probability that in step $t+1$ we will run into already discovered variable is bounded by the probability that the literal explored during that step shares a clause with an already explored variable, which is bounded by $Lm/n \le L/n$; for if more than L literals have been already explored the event $\mathfrak{E}_{t+1}$ does not occur by definition. Finally, because the event $\mathfrak{E}_t$ is $\mathfrak{H}_t$-measurable, (5.8) implies

$$\prod_{i=1}^{\ell} \mathbb{P}\left[\mathfrak{E}_{t_i} \middle| \bigcap_{j=1}^{i-1} \mathfrak{E}_{t_j}\right] \le \left(\frac{L}{n}\right)^{\ell}.$$

Thus (5.7) gives $\mathbb{P}[\mathfrak{E}] \le \binom{L}{\ell}\frac{L^\ell}{n^\ell} \le \left(\frac{eL^2}{\ell n}\right)^{\ell}$, which together with (5.6) implies (5.5). $\square$

Proof of Proposition 2.3. Lemma 4.5 shows that $\mathbb{P}[Z(\boldsymbol{\Phi}) = 0] = o(n^{-1/2})$. Thus, we may condition on the event that $\boldsymbol{\Phi}$ is satisfiable. Furthermore, Lemma 5.1 shows that given that $\boldsymbol{\Phi}$ is satisfiable we have

$$\log Z(\hat{\boldsymbol{\Phi}}) - \log Z(\boldsymbol{\Phi}) \le \sum_{i=1}^{n} \mathbb{1}\{\mathscr{C}(\boldsymbol{\Phi},\{x_i\}) \ne \emptyset\}|\mathscr{V}(\boldsymbol{\Phi},\{x_i\})| + \mathbb{1}\{\mathscr{C}(\boldsymbol{\Phi},\{\neg x_i\}) \ne \emptyset\}|\mathscr{V}(\boldsymbol{\Phi},\{\neg x_i\})|. \tag{5.9}$$

Finally, Corollary 4.4 and Lemma 5.3 (applied with $\delta < 1/3$) imply that with probability $1-o(n^{-1/2})$,

$$\begin{aligned}
&\sum_{i=1}^{n} \mathbb{1}\{\mathscr{C}(\boldsymbol{\Phi},\{x_i\}) \ne \emptyset\}|\mathscr{V}(\boldsymbol{\Phi},\{x_i\})| + \mathbb{1}\{\mathscr{C}(\boldsymbol{\Phi},\{\neg x_i\}) \ne \emptyset\}|\mathscr{V}(\boldsymbol{\Phi},\{\neg x_i\})| \qquad (5.10)\\
&\quad \le (|\{x \in V_n : \mathscr{C}(\boldsymbol{\Phi},\{x\}) \ne \emptyset\}| + |\{x \in V_n : \mathscr{C}(\boldsymbol{\Phi},\{\neg x\}) \ne \emptyset\}|) \max_{1 \le i \le n} |\mathscr{V}(\boldsymbol{\Phi},\{x_i\})| + |\mathscr{V}(\boldsymbol{\Phi},\{\neg x_i\})| = o(n^{1/3}).
\end{aligned}$$

Thus, the assertion follows from (5.9)–(5.10). $\square$

6. Proof of Proposition 2.6

The proof of Proposition 2.6 is based on a combination of a coupling and a second moment argument. As a first step we observe that we do not need to worry about trees of very high maximum degree.

Lemma 6.1. *For any $\varepsilon > 0$, $\ell \ge 0$ there exists $L > 0$ such that for all $t \in [0,1]$ with probability at least $1-\varepsilon$ the tree $\boldsymbol{T}^{\otimes,(2\ell)}$ has maximum degree less than L.*

Proof. The construction of the tree $\boldsymbol{T}^{\otimes,(2\ell)}$ in Section 2.6 ensures that every variable node has a Poisson number of clauses as offspring. The mean of this Poisson variable is always bounded by $2d$. Hence, Bennett's inequality shows that for any $L > 2d$ the probability that a specific variable has more than L offspring is bounded by $\exp(-L^2/(4d+L))$. Thus, choosing L sufficiently large so that $\varepsilon > L^{2\ell}\exp(-L^2/(4d+L))$ and applying the union bound, we obtain the assertion (combined with the chain rule starting from the root). $\square$

Thus, in the following we confine ourselves to trees T with a maximum degree bounded by a large enough number L. First we are going to count the number of copies of such trees T in $(\boldsymbol{\Phi}_1(M,M'),\boldsymbol{\Phi}_2(M,M'))$ via the method of moments. The following lemma estimates the first moment.

Lemma 6.2. *For any fixed integers L,ℓ, any possible outcome T of $\boldsymbol{T}^{\otimes,(2\ell)}$ of maximum degree at most L and any $M \sim tdn/2$, $M' \sim (1-t)dn/2$ we have*

$$\mathbb{E}[\boldsymbol{N}^{(2\ell)}(T,(\boldsymbol{\Phi}_1(M,M'),\boldsymbol{\Phi}_2(M,M')))] \sim n\mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell)} \cong T\right].$$

Proof. We proceed by induction on ℓ. In the case $\ell = 0$ the tree T consists of nothing but the root, so that there is nothing to show. Hence, let $\ell \geq 1$. Let λ_{0,s_1,s_2} be the number of shared children of the root o of T where o appears with sign $s_1 \in \{-1,+1\}$ and the other variable appears with sign $s_2 \in \{-1,+1\}$. Also let λ_{h,s_1,s_2} be the number of h-distinct children of o ($h = 1,2$), where o appears with sign $s_1 \in \{-1,+1\}$ and the other variable appears with sign $s_2 \in \{-1,+1\}$.

Consider the event $\mathfrak{E}$ that variable x_1 is a 2ℓ-instance of T. Further, consider the event $\mathfrak{R}$ that x_1 occurs in precisely λ_{0,s_1,s_2} clauses among $\boldsymbol{a}_1,\dots,\boldsymbol{a}_M$, where the sign of x_1 is s_1 and the sign of the other variable is s_2, precisely in λ_{1,s_1,s_2} clauses among $\boldsymbol{a}'_1,\dots,\boldsymbol{a}'_{M'}$, where the sign of x_1 is s_1 and the sign of the other variable is s_2 and precisely in λ_{2,s_1,s_2} clauses among $\boldsymbol{a}''_1,\dots,\boldsymbol{a}''_{M'}$, where the sign of x_1 is s_1 and the sign of the other variable is s_2. Since $M \sim dtn/2$ and $\lambda_{h,\pm1,\pm1} \leq L$ for $h \in \{0,1,2\}$ we have

$$\begin{aligned}\mathbb{P}[\mathfrak{R}] &\sim \prod_{s_1,s_2\in\{\pm1\}} \mathbb{P}\left[\mathrm{Bin}(M,(2n)^{-1}) = \lambda_{0,s_1,s_2}\right]\mathbb{P}\left[\mathrm{Bin}(m-M,(2n)^{-1}) = \lambda_{1,s_1,s_2}\right]\mathbb{P}\left[\mathrm{Bin}(m-M,(2n)^{-1}) = \lambda_{2,s_1,s_2}\right]\\ &\sim \prod_{s_1,s_2\in\{\pm1\}} \mathbb{P}\left[\mathrm{Po}(dt/4) = \lambda_{0,s_1,s_2}\right]\mathbb{P}\left[\mathrm{Po}(d(1-t)/4) = \lambda_{1,s_1,s_2}\right]\mathbb{P}\left[\mathrm{Po}(d(1-t)/4) = \lambda_{2,s_1,s_2}\right]. \end{aligned}\tag{6.1}$$

Let $\lambda_h = \lambda_{h,-1,-1}+\lambda_{h,-1,+1}+\lambda_{h,+1,-1}+\lambda_{h,+1,+1}$ for $h \in \{0,1,2\}$. Given $\mathfrak{R}$ let $(\boldsymbol{v}_{0,i})_{1\leq i\leq\lambda_0}$ be the second variables (other than x_1) contained in neighbours of x_1 among $\boldsymbol{a}_1,\dots,\boldsymbol{a}_M$. Analogously, let $(\boldsymbol{v}_{1,i})_{1\leq i\leq\lambda_h}$ be the second variables contained in neighbours of x_1 among $\boldsymbol{a}'_1,\dots,\boldsymbol{a}'_M$ and $(\boldsymbol{v}_{2,i})_{1\leq i\leq\lambda_h}$ be the second variables contained in neighbours of x_1 among $\boldsymbol{a}''_1,\dots,\boldsymbol{a}''_M$. By $\boldsymbol{\Phi}^-_h$, $h = 1,2$ define a random formula obtained from $\boldsymbol{\Phi}_h(M,M')$ by deleting x_1 and its adjacent clauses. Let $\mathfrak{F}$ be the event that the distance between any two of $\boldsymbol{v}_{0,1},\dots,\boldsymbol{v}_{2,\lambda_2}$ in both $\boldsymbol{\Phi}^-_1$ and $\boldsymbol{\Phi}^-_2$ is at least 2ℓ. A routine union bound argument shows that

$$\mathbb{P}[\mathfrak{F}] = 1-o(1).\tag{6.2}$$

Further, let $T_{0,i}$ be the sub-tree obtained from T comprising the i-th shared grandchild of o and its descendants. Consider the event $\mathfrak{H}_0$ that $\mathfrak{R}$ and $\mathfrak{F}$ occur and $\boldsymbol{v}_{0,i}$ is a $(2\ell-2)$-instance of $T_{0,i}$ in $(\boldsymbol{\Phi}^-_1,\boldsymbol{\Phi}^-_2)$ for any $i = 1,\dots,\lambda_0$. Since the depth and the maximum degree of T are bounded, by induction we obtain

$$\mathbb{P}\left[\boldsymbol{v}_{0,i} \text{ is a } (2\ell-2)\text{-instance of } T_{0,i} \text{ in } (\boldsymbol{\Phi}^-_1,\boldsymbol{\Phi}^-_2)\right] = \mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell-2)} \cong T_{0,i}\right] + o(1).$$

for $i = 1,\dots,\lambda_0$. Thus

$$\mathbb{P}[\mathfrak{H}_0|\mathfrak{F}\cap\mathfrak{R}] = \mathbb{P}\left[\bigcap_{i=1}^{\lambda_0} \boldsymbol{v}_{0,i} \text{ is a } (2\ell-2)\text{-instance of } T_{0,i} \text{ in } (\boldsymbol{\Phi}^-_1,\boldsymbol{\Phi}^-_2)\right] = \prod_{i=0}^{\lambda_0}\mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell-2)} \cong T_{0,i}\right] + o(1).\tag{6.3}$$

Analogously, let $T_{h,i}$ be the sub-tree of T pending on the i-th h-distinct grandchild of the root. Consider the events $\mathfrak{H}_h$ that $\mathfrak{F}$ and $\mathfrak{R}$ occur and that the depth $(2\ell-2)$-neighbourhood of $\boldsymbol{v}_{h,i}$ is isomorphic to $T_{h,i}$ in $\boldsymbol{\Phi}^-_h$ for any $i = 1,\dots,\lambda_h$, $h = 1,2$. Since $\boldsymbol{v}_{1,i}$ and $\boldsymbol{v}_{2,j}$ are chosen independently for all i and j, using the same embedding process as above in combination with Lemma 4.6 we obtain

$$\mathbb{P}[\mathfrak{H}_1|\mathfrak{F}\cap\mathfrak{R}\cap\mathfrak{H}_0] = \prod_{i=1}^{\lambda_1}\mathbb{P}\left[\boldsymbol{T}^{(2\ell-2)} \cong T_{1,i}\right] + o(1)\tag{6.4}$$

$$\mathbb{P}[\mathfrak{H}_2|\mathfrak{F}\cap\mathfrak{R}\cap\mathfrak{H}_0\cap\mathfrak{H}_1] = \prod_{i=1}^{\lambda_2}\mathbb{P}\left[\boldsymbol{T}^{(2\ell-2)} \cong T_{2,i}\right] + o(1).\tag{6.5}$$

Finally, combining (6.1)–(6.5) we obtain

$$\mathbb{P}[\mathfrak{E}] \sim \mathbb{P}[\mathfrak{R}]\prod_{i=1}^{\lambda_0}\mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell-2)} \cong T_{0,i}\right]\prod_{i=1}^{\lambda_1}\mathbb{P}\left[\boldsymbol{T}^{(2\ell-2)} \cong T_{1,i}\right]\prod_{i=1}^{\lambda_2}\mathbb{P}\left[\boldsymbol{T}^{(2\ell-2)} \cong T_{2,i}\right] \sim \mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell-2)} \cong T\right].$$

As $\mathbb{E}[\boldsymbol{N}^{(2\ell)}(T,(\boldsymbol{\Phi}_1(M,M'),\boldsymbol{\Phi}_2(M,M')))] = n\mathbb{P}[\mathfrak{E}]$ the assertion follows from the linearity of expectation. □

We also need an estimate of the second moment of $\boldsymbol{N}^{(2\ell)}(T,(\boldsymbol{\Phi}_1(M,M'),\boldsymbol{\Phi}_2(M,M')))$.

Lemma 6.3. *For any fixed integers L,ℓ and any possible outcome T of $\boldsymbol{T}^{(2\ell)}$ of maximum degree at most L and any $M \sim tdn/2$, $M' \sim (1-t)dn/2$ we have*

$$\mathbb{E}[\boldsymbol{N}^{(2\ell)}(T,(\boldsymbol{\Phi}_1(M,M'),\boldsymbol{\Phi}_2(M,M')))^2] \sim n^2\mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell)} \cong T\right]^2.$$

Proof. Consider the event $\mathfrak{E}_{ij}$ that both variables x_i and x_j are 2ℓ-instances of T for $i, j = 1, \ldots, n$. Now we can rewrite the second moment as follows.

$$\mathbb{E}\left[\boldsymbol{N}^{(2\ell)}(T, (\boldsymbol{\Phi}_1(M, M'), \boldsymbol{\Phi}_2(M, M')))^2\right] = \sum_{i,j=1}^{n} \mathbb{P}\left[\mathfrak{E}_{ij}\right] = n\mathbb{P}\left[\mathfrak{E}_{11}\right] + n(n-1)\mathbb{P}\left[\mathfrak{E}_{12}\right]. \tag{6.6}$$

From Lemma 6.2 we know that $\mathbb{P}\left[\mathfrak{E}_{11}\right] = \mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell)} \cong T\right] + o(1)$, so we only need to estimate $\mathbb{P}\left[\mathfrak{E}_{12}\right]$. Let $\mathfrak{F}$ be such event that the distance between x_1 and x_2 is at least 2ℓ. A routine union bound argument shows that

$$\mathbb{P}\left[\mathfrak{F}\right] = 1 - o(1).$$

This fact completes the proof of Lemma 6.3.

$$\begin{aligned}\mathbb{P}\left[\mathfrak{E}_{12}\right] &= \mathbb{P}\left[\mathfrak{E}_{12} \mid \mathfrak{F}\right] + o(1) = \mathbb{P}\left[x_1 \text{ is a } 2\ell\text{-instance of } T \mid \mathfrak{F}\right] \cdot \mathbb{P}\left[x_2 \text{ is a } 2\ell\text{-instance of } T \mid \mathfrak{F}\right] + o(1)\\ &= \mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell)} \cong T\right]^2 + o(1). \end{aligned}\tag{6.7}$$

Thus the assertion follows from (6.6) and (6.7). □

Proof of Proposition 2.6. From Lemmas 6.1–6.3 in combination with Chebyshev's inequality it follows that for any $\ell \geq 0$, T w.h.p.

$$\boldsymbol{N}^{(2\ell)}(T, (\boldsymbol{\Phi}_1(M, M'), \boldsymbol{\Phi}_2(M, M'))) \sim n\mathbb{P}\left[\boldsymbol{T}^{\otimes,(2\ell)} \cong T\right]. \tag{6.8}$$

We need to extend this to the pruned formulas $(\hat{\boldsymbol{\Phi}}_1(M, M'), \hat{\boldsymbol{\Phi}}_2(M, M'))$. Let $\boldsymbol{N}^{(2\ell),+}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2))$ be the number of variable nodes x such that x is an 2ℓ-instance of T in $(\hat{\boldsymbol{\Phi}}_1, \hat{\boldsymbol{\Phi}}_2)$ but not in $(\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)$. Similarly, let $\boldsymbol{N}^{(2\ell),-}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2))$ be the number of variable nodes x such that they are 2ℓ-instances of T in $(\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)$ but not in $(\hat{\boldsymbol{\Phi}}_1, \hat{\boldsymbol{\Phi}}_2)$. Then

$$\boldsymbol{N}^{(2\ell)}(T, (\hat{\boldsymbol{\Phi}}_1, \hat{\boldsymbol{\Phi}}_2)) = \boldsymbol{N}^{(2\ell)}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)) + \boldsymbol{N}^{(2\ell),+}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)) - \boldsymbol{N}^{(2\ell),-}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)). \tag{6.9}$$

Note that both $\boldsymbol{N}^{(2\ell),+}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2))$ and $\boldsymbol{N}^{(2\ell),-}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2))$ do not exceed the number of variable nodes x whose depth-2ℓ neighbourhood in $(\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)$ contains at least one clause from $\bigcup_{l \in \{x_i, \neg x_i,\ 1 \leq i \leq n\}} \mathscr{C}(\boldsymbol{\Phi}, \{l\})$. Moreover, Lemmas 4.4 and 5.3 show that w.h.p.

$$\left|\bigcup_{l \in \{x_i, \neg x_i,\ 1 \leq i \leq n\}} \mathscr{C}(\boldsymbol{\Phi}, \{l\})\right| \leq n^{0.1}. \tag{6.10}$$

It follows from Lemma 4.8 that w.h.p. the 2ℓ-depth neighbourhood of each vertex consists of no more then $\log^{4\ell+4} n$ vertices. Combining this fact with (6.10) we conclude that

$$\boldsymbol{N}^{(2\ell),+}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)) \leq n^{0.1} \log^{4\ell+4} n, \qquad \boldsymbol{N}^{(2\ell),-}(T, (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)) \leq n^{0.1} \log^{4\ell+4} n \tag{6.11}$$

w.h.p. Finally, the assertion follows from (6.8), (6.9) and (6.11). □

7. PROOF OF PROPOSITION 2.5

We will deal with $\frac{Z(\hat{\boldsymbol{\Phi}}_h(M, m-M))}{Z(\hat{\boldsymbol{\Phi}}_h(M-1, m-M))}$ in detail; the arguments for the other two quotients are similar.

Lemma 7.1. *Let $h \in \{1, 2\}$. W.h.p. $\hat{\boldsymbol{\Phi}}_h(M, m-M)$ is obtained from $\hat{\boldsymbol{\Phi}}_h(M-1, m-M)$ by adding a clause $\boldsymbol{a}_M$.*

Proof. Let $\boldsymbol{l}, \boldsymbol{l}'$ be the constituent literals of $\boldsymbol{a}_M$, i.e., $\boldsymbol{a}_M = \boldsymbol{l} \vee \boldsymbol{l}'$. Moreover, let $\mathfrak{Q}$ be the event that $\hat{\boldsymbol{\Phi}}_h(M, m-M)$ does *not* result from $\hat{\boldsymbol{\Phi}}_h(M-1, m-M)$ by adding clause $\boldsymbol{a}_M$. Thus, on the event $\mathfrak{Q}$ the additional clause $\boldsymbol{a}_M$ triggers the pruning of clauses that do not get pruned from $\hat{\boldsymbol{\Phi}}(M-1, m-M)$ (including potentially $\boldsymbol{a}_M$ itself).

We are going to construct events $\mathfrak{E}, \mathfrak{E}'$ whose probabilities are easy to estimate such that

$$\mathfrak{Q} \subseteq \mathfrak{E} \cup \mathfrak{E}'. \tag{7.1}$$

To this end, for a literal l let $\mathscr{L}_l = \mathscr{L}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\})$ be the final set of literals that $\mathtt{PUC}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\})$ produces. Call l a *trigger* of $\neg \boldsymbol{l}$ if $\neg \boldsymbol{l} \in \mathscr{L}_l$. Further, let $\mathfrak{E}$ be the event that there exists a trigger l of $\neg \boldsymbol{l}$ such that

E1: $\mathscr{C}(\boldsymbol{\Phi}(M-1, m-M), \{l\}) \cup \mathscr{C}(\boldsymbol{\Phi}(M-1, m-M), \{l, \boldsymbol{l}'\}) \neq \emptyset$, or

E2: $\neg \boldsymbol{l}' \in \bigcup_{\lambda \in \{l, \boldsymbol{l}'\}} \mathscr{L}_\lambda$.

Define $\mathfrak{E}'$ analogously with the roles of $\boldsymbol{l}, \boldsymbol{l}'$ swapped.

We claim that these events $\mathfrak{E}, \mathfrak{E}'$ satisfy (7.1). To see this, assume that neither $\mathfrak{E}$ nor $\mathfrak{E}'$ occurs. We claim that then

$$\mathscr{C}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\}) = \mathscr{C}(\boldsymbol{\Phi}_h(M, m-M), \{l\}) \tag{7.2}$$

for all literals l; if so, then clearly $\mathfrak{Q}$ does not occur either.

Thus, assume that (7.2) is false and that l is a literal such that

$$\mathscr{C}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\}) \neq \mathscr{C}(\boldsymbol{\Phi}_h(M, m-M), \{l\}). \tag{7.3}$$

Then l must be a trigger of $\neg \boldsymbol{l}$ or of $\neg \boldsymbol{l}'$; for otherwise the presence of the extra clause $\boldsymbol{a}_M$ has no impact on the set of conflict clauses. Hence, suppose that l is a trigger of $\neg \boldsymbol{l}$. Then the presence of clause $\boldsymbol{a}_M$ in $\boldsymbol{\Phi}_h(M, m-M)$ causes PUC to add $\boldsymbol{l}'$ to $\mathscr{L}(\boldsymbol{\Phi}_h(M, m-M), \{l\})$. Since the event $\mathfrak{E}$ does not occur, neither does **E1** and we conclude that $\mathscr{C}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\}) = \mathscr{C}(\boldsymbol{\Phi}_h(M-1, m-M), \{l, \boldsymbol{l}'\}) = \emptyset$. Hence, none of the clauses $a \in F(\boldsymbol{\Phi}_h(M-1, m-M))$ is a conflict clause and thus (7.3) implies that

$$\{\boldsymbol{a}_M\} = \mathscr{C}(\boldsymbol{\Phi}_h(M, m-M), \{l\}) \setminus \mathscr{C}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\}).$$

But this is not possible either. For if $\boldsymbol{a}_M \in \mathscr{C}(\boldsymbol{\Phi}_h(M, m-M), \{l\})$, then Lemma 4.2 shows that one of $l, \boldsymbol{l}, \boldsymbol{l}'$ is a trigger of $\neg \boldsymbol{l}'$, and thus **E2** occurs. Thus, we obtain (7.2).

To complete the proof we are going to show that

$$\mathbb{P}[\mathfrak{E}], \mathbb{P}\left[\mathfrak{E}'\right] = o(1). \tag{7.4}$$

Indeed, Lemma 5.3 shows that the number of literals l such that $\mathscr{C}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\}) \neq \emptyset$ can be bounded by $n^{0.1}$ w.h.p. Furthermore, Corollary 4.4 shows that $|\mathscr{V}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\})| \leq \log^2 n$ w.h.p. for all l. Hence, w.h.p. the total number of literals λ that have a trigger l such that $\mathscr{C}(\boldsymbol{\Phi}_h(M-1, m-M), \{l\}) \neq \emptyset$ is bounded by $O(n^{0.1} \log^2 n)$. Consequently, the probability that the random literal $\neg \boldsymbol{l}$ possesses such a trigger is bounded by $O(n^{-0.9} \log^2 n)$. Moreover, since $\boldsymbol{l}'$ is a random literal as well, Lemma 5.3 shows that $\mathbb{P}\left[\mathscr{C}(\boldsymbol{\Phi}(M-1, m-M), \{\boldsymbol{l}'\}) = \emptyset\right] = 1 - O(n^{-0.9})$. Additionally, w.h.p. for any trigger l of $\neg \boldsymbol{l}$ we have $\mathscr{V}(\boldsymbol{\Phi}(M-1, m-M), \{l\}) \cap \mathscr{V}(\boldsymbol{\Phi}(M-1, m-M), \{\boldsymbol{l}'\}) = \emptyset$, because $\boldsymbol{l}, \boldsymbol{l}'$ are drawn independently of $\boldsymbol{\Phi}(M-1, m-M)$. Similarly,

$$\mathbb{P}\left[\neg \boldsymbol{l}' \in \mathscr{L}(\boldsymbol{\Phi}(M-1, m-M), \{l\})\right] = o(1) \qquad \text{and} \qquad \mathbb{P}\left[\neg \boldsymbol{l}' \in \mathscr{L}(\boldsymbol{\Phi}(M-1, m-M), \{\boldsymbol{l}'\})\right] = o(1).$$

Combining these estimates, we conclude that $\mathbb{P}[\mathfrak{E}] = o(1)$. By symmetry, the same estimate holds for $\mathfrak{E}'$. Thus, we obtain (7.4). Finally, the assertion follows from (7.1) and (7.4). □

Corollary 7.2. *Let $h \in \{1, 2\}$. W.h.p. we have*

$$\frac{Z(\hat{\boldsymbol{\Phi}}_h(M, m-M))}{Z(\hat{\boldsymbol{\Phi}}_h(M-1, m-M))} = \mu_{\hat{\boldsymbol{\Phi}}_h(M-1, m-M)}(\{\boldsymbol{\sigma} \models \boldsymbol{a}_M\}).$$

Proof. From Lemma 7.1 we know that w.h.p.

$$Z(\hat{\boldsymbol{\Phi}}_h(M, m-M)) = Z(\hat{\boldsymbol{\Phi}}(M-1, m-M) + \boldsymbol{a}_M). \tag{7.5}$$

Assuming that (7.5) is correct, $Z(\hat{\boldsymbol{\Phi}}_h(M, m-M))$ equals the number of satisfying assignments of $\hat{\boldsymbol{\Phi}}_h(M-1, m-M)$ that also happen to satisfy $\boldsymbol{a}_M$. □

Additionally, we need the following asymptotic independence property, known as 'replica symmetry' in physics parlance.

Lemma 7.3. *Let $h \in \{1, 2\}$. For all $s, s' \in \{\pm 1\}$ we have*

$$\frac{1}{n^2} \sum_{i,j=1}^{n} \mathbb{E}\left|\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_i} = s, \boldsymbol{\sigma}_{x_j} = s'\}) - \mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_i} = s\})\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_j} = s'\})\right| = o(1).$$

Proof. We adapt an argument from [40] to the present setting. By exchangeability it suffices to prove that

$$\mathbb{E}\left|\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_1} = s, \boldsymbol{\sigma}_{x_2} = s'\}) - \mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_1} = s\})\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_2} = s'\})\right| = o(1).$$

The proof rests on the Gibbs uniqueness property. Indeed, Proposition 2.6 shows that for any fixed ℓ the depth-2ℓ neighbourhood $\partial^{\leq 2\ell} x_i$ of x_i in $\hat{\boldsymbol{\Phi}}_h(M-1, m-M)$ is within total variation distance $o(1)$ of the Galton-Watson tree $\boldsymbol{T}_h^{(2\ell)}$. Furthermore, the distribution of $\boldsymbol{T}_h^{(2\ell)}$ by itself is identical to the distribution of the Galton-Watson tree $\boldsymbol{T}^{(2\ell)}$.

Additionally, Proposition 2.1 shows that $\boldsymbol{T}^{(2\ell)}$ enjoys the Gibbs uniqueness property (2.6). Consequently, taking $\ell=\ell(n)\to\infty$ sufficiently slowly as $n\to\infty$, we see that w.h.p.

$$\sum_{s\in\{\pm1\}}\max_{\kappa\in S(\hat{\boldsymbol{\Phi}}_h(M-1,m-M))}\left|\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_1}=s\mid\boldsymbol{\sigma}_{\partial^\ell x_1}=\kappa_{\partial^{2\ell}x_1}\})-\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_1}=s\})\right|=o(1). \tag{7.6}$$

Furthermore, providing $\ell=\ell(n)\to\infty$ slowly enough, the distance between x_1,x_2 exceeds 4ℓ w.h.p. In this case, (7.6) gives

$$\begin{aligned}
&\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{x_1}=s,\boldsymbol{\sigma}_{x_2}=s'\})=\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{x_1}=s\mid\boldsymbol{\sigma}_{x_2}=s'\})\cdot\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{x_2}=s'\})\\
&=\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{x_2}=s'\})\cdot\sum_{\kappa\in\{\pm1\}^{\partial^{2\ell}x_1}}\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{x_1}=s\mid\boldsymbol{\sigma}_{\partial^{2\ell}x_1}=\kappa,\boldsymbol{\sigma}_{x_2}=s'\})\\
&\qquad\qquad\cdot\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{\partial^{2\ell}x_1}=\kappa\mid\boldsymbol{\sigma}_{x_2}=s'\})\\
&=\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{x_2}=s'\})\cdot\sum_{\kappa\in\{\pm1\}^{\partial^{2\ell}x_1}}\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{x_1}=s\mid\boldsymbol{\sigma}_{\partial^{2\ell}x_1}=\kappa\})\\
&\qquad\qquad\cdot\mu_{\hat{\boldsymbol{\Phi}}_h(M,m-M)}(\{\boldsymbol{\sigma}_{\partial^{2\ell}x_1}=\kappa\mid\boldsymbol{\sigma}_{x_2}=s'\})\\
&=\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_1}=s\})\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}(\{\boldsymbol{\sigma}_{x_2}=s'\})\cdot(1+o(1)),
\end{aligned}\tag{7.7}$$

as claimed. □

Proof of Proposition 2.5. The proposition follows from Corollary 7.2 and Lemma 7.3. □

8. Proof of Proposition 2.8

In this section, we prove Proposition 2.8 via a contraction argument. For this, recall the operator $\mathtt{logBP}^{\otimes}_{d,t}$ from (1.3). For notational convenience we let

$$\boldsymbol{V}=\begin{pmatrix}\sum_{i=1}^{\boldsymbol{d}}\boldsymbol{s}_i\log\left(\frac{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\rho,i,1}/2)}{2}\right)+\sum_{i=1}^{\boldsymbol{d}'}\boldsymbol{s}'_i\log\left(\frac{1+\boldsymbol{r}'_i\tanh(\boldsymbol{\xi}'_{\rho,i,1}/2)}{2}\right)\\ \sum_{i=1}^{\boldsymbol{d}}\boldsymbol{s}_i\log\left(\frac{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\rho,i,2}/2)}{2}\right)+\sum_{i=1}^{\boldsymbol{d}''}\boldsymbol{s}''_i\log\left(\frac{1+\boldsymbol{r}''_i\tanh(\boldsymbol{\xi}''_{\rho,i,2}/2)}{2}\right)\end{pmatrix}.$$

The main step towards Proposition 2.8 is the following lemma:

Lemma 8.1. $\mathtt{logBP}^{\otimes}_{d,t}$ *is a contraction on the space* $(\mathscr{W}_2(\mathbb{R}^2),W_2)$ *for all* $0<d<2$ *and* $0\le t\le 1$.

Indeed, it immediately follows from Lemma 8.1 and Banach's fixed point theorem that for every $d\in(0,2)$ and $t\in[0,1]$, there is a unique $\rho_{d,t}\in\mathscr{W}_2(\mathbb{R}^2)$ with $\rho_{d,t}=\mathtt{logBP}^{\otimes}_{d,t}(\rho_{d,t})$, and that for any $\rho\in\mathscr{W}_2(\mathbb{R}^2)$ and $\ell\to\infty$, the ℓ-fold application of $\mathtt{logBP}^{\otimes}_{d,t}$ to ρ converges to $\rho_{d,t}$ in Wasserstein distance.

We prove Lemma 8.1 in the following subsection, and conclude the section with the proof of Proposition 2.8.

8.1. **Proof of Lemma 8.1.**

We first check that the operator $\mathtt{logBP}^{\otimes}_{d,t}$ is well-defined in the sense that it maps the space $(\mathscr{W}_2(\mathbb{R}^2),W_2)$ to itself.

Claim 8.2. *The operator* $\mathtt{logBP}^{\otimes}_{d,t}$ *maps the space* $(\mathscr{W}_2(\mathbb{R}^2),W_2)$ *to itself.*

Proof. Let $\rho\in(\mathscr{W}_2(\mathbb{R}^2),W_2)$ and $\boldsymbol{V}$ be a random vector with distribution $\mathtt{logBP}^{\otimes}_{d,t}(\rho)$. By the definition of $\mathtt{logBP}^{\otimes}_{d,t}$,

$$\begin{aligned}\mathbb{E}[\|\boldsymbol{V}\|_2^2]=\mathbb{E}\Bigg[&\left(\sum_{i=1}^{\boldsymbol{d}}\boldsymbol{s}_i\log\left(\frac{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\rho,i,1}/2)}{2}\right)+\sum_{i=1}^{\boldsymbol{d}'}\boldsymbol{s}'_i\log\left(\frac{1+\boldsymbol{r}'_i\tanh(\boldsymbol{\xi}'_{\rho,i,1}/2)}{2}\right)\right)^2\\
&+\left(\sum_{i=1}^{\boldsymbol{d}}\boldsymbol{s}_i\log\left(\frac{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\rho,i,2}/2)}{2}\right)+\sum_{i=1}^{\boldsymbol{d}''}\boldsymbol{s}''_i\log\left(\frac{1+\boldsymbol{r}''_i\tanh(\boldsymbol{\xi}''_{\rho,i,2}/2)}{2}\right)\right)^2\Bigg].\end{aligned}\tag{8.1}$$

By the independence of the random variables $(\boldsymbol{s}_i)_{i\ge1}$, $(\boldsymbol{s}'_i)_{i\ge1}$ and $(\boldsymbol{s}''_i)_{i\ge1}$ from everything else, all cross-terms in the evaluation of the squares in (8.1) vanish, e.g. for $i\neq j$,

$$\begin{aligned}&\mathbb{E}\left[\boldsymbol{s}_i\boldsymbol{s}_j\log\left(\frac{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\rho,i,1}/2)}{2}\right)\log\left(\frac{1+\boldsymbol{r}_j\tanh(\boldsymbol{\xi}_{\rho,j,1}/2)}{2}\right)\right]\\
&=\mathbb{E}[\boldsymbol{s}_i]\,\mathbb{E}\left[\boldsymbol{s}_j\log\left(\frac{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\rho,i,1}/2)}{2}\right)\log\left(\frac{1+\boldsymbol{r}_j\tanh(\boldsymbol{\xi}_{\rho,j,1}/2)}{2}\right)\right]=0.\end{aligned}$$

As a consequence, (8.1) in combination with the independence of the Poisson random variables gives that

$$
\begin{aligned}
\mathbb{E}\left[\|\boldsymbol{V}\|_2^2\right] =&\mathbb{E}\Bigg[td\log^2\left(\frac{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\rho,1,1}/2)}{2}\right)+(1-t)d\log^2\left(\frac{1+\boldsymbol{r}'_1\tanh(\boldsymbol{\xi}'_{\rho,1,1}/2)}{2}\right)\\
&+td\log^2\left(\frac{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\rho,1,2}/2)}{2}\right)+(1-t)d\log^2\left(\frac{1+\boldsymbol{r}''_1\tanh(\boldsymbol{\xi}''_{\rho,1,2}/2)}{2}\right)\Bigg].
\end{aligned}
\tag{8.2}
$$

Finally, conditioning on the value of $\boldsymbol{r}_1$ and an application of the fundamental theorem of calculus, followed by the Cauchy-Schwarz inequality give

$$
\begin{aligned}
\mathbb{E}\left[\log^2\left(\frac{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\rho,1,1}/2)}{2}\right)\right] &=\frac{1}{2}\mathbb{E}\left[\log^2\left(\frac{1+\tanh(\boldsymbol{\xi}_{\rho,1,1}/2)}{2}\right)+\log^2\left(\frac{1-\tanh(\boldsymbol{\xi}_{\rho,1,1}/2)}{2}\right)\right]\\
&=\frac{1}{2}\mathbb{E}\left[\left(\int_0^{\boldsymbol{\xi}_{\rho,1,1}}\frac{1-\tanh(x/2)}{2}\,\mathrm{d}x-\log 2\right)^2+\left(\int_0^{\boldsymbol{\xi}_{\rho,1,1}}\frac{1+\tanh(x/2)}{2}\,\mathrm{d}x+\log 2\right)^2\right]\\
&\le 2\mathbb{E}\left[\log^2 2+\boldsymbol{\xi}_{\rho,1,1}^2\right].
\end{aligned}
$$

Analogous bounds can be derived for the remaining three terms in (8.2). Therefore, for any vector $\boldsymbol{\xi}\in\mathbb{R}^2$ with distribution ρ,

$$
\mathbb{E}\left[\|\boldsymbol{V}\|_2^2\right]\le\mathbb{E}\left[2td\left(\boldsymbol{\xi}_{\rho,1,1}^2+\boldsymbol{\xi}_{\rho,1,2}^2\right)+2(1-t)d\left(\boldsymbol{\xi}'^2_{\rho,1,1}+\boldsymbol{\xi}''^2_{\rho,1,2}\right)\right]+4d\log^2 2=2d\mathbb{E}\left[\|\boldsymbol{\xi}\|_2^2\right]+4d\log^2 2<\infty.
$$

□

Proof of Lemma 8.1. Let $\rho,\nu\in\mathscr{W}_2(\mathbb{R}^2)$ be arbitrary. To show contraction, consider three independent sequences of optimally coupled pairs $(\boldsymbol{\xi}_{\rho,i},\boldsymbol{\xi}_{\nu,i})_{i\ge1}$, $(\boldsymbol{\xi}'_{\rho,i},\boldsymbol{\xi}'_{\nu,i})_{i\ge1}$ and $(\boldsymbol{\xi}''_{\rho,i},\boldsymbol{\xi}''_{\nu,i})_{i\ge1}$ such that for each $\boldsymbol{\zeta}=\boldsymbol{\xi},\boldsymbol{\xi}',\boldsymbol{\xi}''$, the $\boldsymbol{\zeta}_{\rho,i}=(\boldsymbol{\zeta}_{\rho,i,1},\boldsymbol{\zeta}_{\rho,i,2})\in\mathbb{R}^2$ have distribution ρ, the $\boldsymbol{\zeta}_{\nu,i}=(\boldsymbol{\zeta}_{\nu,i,1},\boldsymbol{\zeta}_{\nu,i,2})\in\mathbb{R}^2$ have distribution ν and

$$
W_2(\rho,\nu)=\mathbb{E}\left[\|\boldsymbol{\zeta}_{\rho,i}-\boldsymbol{\zeta}_{\nu,i}\|_2^2\right]^{1/2}.
\tag{8.3}
$$

Let $\boldsymbol{d}\sim\mathrm{Po}(td)$ and $\boldsymbol{d}',\boldsymbol{d}''\sim\mathrm{Po}((1-t)d)$ all be independent. Moreover, let $(\boldsymbol{s}_i)_{i\ge1}$, $(\boldsymbol{r}_i)_{i\ge1}$, $(\boldsymbol{s}'_i)_{i\ge1}$, $(\boldsymbol{r}'_i)_{i\ge1}$, $(\boldsymbol{s}''_i)_{i\ge1}$ and $(\boldsymbol{r}''_i)_{i\ge1}$ be independent sequences of i.i.d. Rademacher random variables with parameter $1/2$; all not explicitly coupled random variables are assumed to be independent. Then with $\hat\rho=\mathtt{logBP}^{\otimes}_{d,t}(\rho)$ and $\hat\nu=\mathtt{logBP}^{\otimes}_{d,t}(\nu)$ we obtain

$$
\begin{aligned}
W_2(\hat\rho,\hat\nu)^2\le\mathbb{E}&\left[\left(\sum_{i=1}^{\boldsymbol{d}}\boldsymbol{s}_i\log\left(\frac{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\rho,i,1}/2)}{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\nu,i,1}/2)}\right)+\sum_{i=1}^{\boldsymbol{d}'}\boldsymbol{s}'_i\log\left(\frac{1+\boldsymbol{r}'_i\tanh(\boldsymbol{\xi}'_{\rho,i,1}/2)}{1+\boldsymbol{r}'_i\tanh(\boldsymbol{\xi}'_{\nu,i,1}/2)}\right)\right)^2\right]\\
&+\mathbb{E}\left[\left(\sum_{i=1}^{\boldsymbol{d}}\boldsymbol{s}_i\log\left(\frac{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\rho,i,2}/2)}{1+\boldsymbol{r}_i\tanh(\boldsymbol{\xi}_{\nu,i,2}/2)}\right)+\sum_{i=1}^{\boldsymbol{d}''}\boldsymbol{s}''_i\log\left(\frac{1+\boldsymbol{r}''_i\tanh(\boldsymbol{\xi}''_{\rho,i,2}/2)}{1+\boldsymbol{r}''_i\tanh(\boldsymbol{\xi}''_{\nu,i,2}/2)}\right)\right)^2\right].
\end{aligned}
$$

Analogous to the derivation of (8.2), by the independence of the random signs, the expectations of the cross-terms cancel. Combined with the independence of the Poisson random variables, we conclude that

$$
\begin{aligned}
W_2(\hat\rho,\hat\nu)^2\le td\mathbb{E}&\left[\log^2\left(\frac{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\rho,1,1}/2)}{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\nu,1,1}/2)}\right)\right]+(1-t)d\mathbb{E}\left[\log^2\left(\frac{1+\boldsymbol{r}'_1\tanh(\boldsymbol{\xi}'_{\rho,1,1}/2)}{1+\boldsymbol{r}'_1\tanh(\boldsymbol{\xi}'_{\nu,1,1}/2)}\right)\right]\\
&+td\mathbb{E}\left[\log^2\left(\frac{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\rho,1,2}/2)}{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\nu,1,2}/2)}\right)\right]+(1-t)d\mathbb{E}\left[\log^2\left(\frac{1+\boldsymbol{r}''_1\tanh(\boldsymbol{\xi}''_{\rho,1,2}/2)}{1+\boldsymbol{r}''_1\tanh(\boldsymbol{\xi}''_{\nu,1,2}/2)}\right)\right].
\end{aligned}
\tag{8.4}
$$

Moreover, conditioning on the value of $\boldsymbol{r}_1$ and an application of the fundamental theorem of calculus yield

$$
\log^2\frac{1+\tanh(\boldsymbol{\xi}_{\rho,1,1}/2)}{1+\tanh(\boldsymbol{\xi}_{\nu,1,1}/2)}=\left[\int_{\boldsymbol{\xi}_{\nu,1,1}}^{\boldsymbol{\xi}_{\rho,1,1}}\frac{\partial\log(1+\tanh(z/2))}{\partial z}\mathrm{d}z\right]^2=\left[\int_{\boldsymbol{\xi}_{\rho,1,1}\wedge\boldsymbol{\xi}_{\nu,1,1}}^{\boldsymbol{\xi}_{\rho,1,1}\vee\boldsymbol{\xi}_{\nu,1,1}}\frac{1-\tanh(z/2)}{2}dz\right]^2,
\tag{8.5}
$$

$$
\log^2\frac{1-\tanh(\boldsymbol{\xi}_{\rho,1,1}/2)}{1-\tanh(\boldsymbol{\xi}_{\nu,1,1}/2)}=\left[\int_{\boldsymbol{\xi}_{\nu,1,1}}^{\boldsymbol{\xi}_{\rho,1,1}}\frac{\partial\log(1-\tanh(z/2))}{\partial z}dz\right]^2=\left[\int_{\boldsymbol{\xi}_{\rho,1,1}\wedge\boldsymbol{\xi}_{\nu,1,1}}^{\boldsymbol{\xi}_{\rho,1,1}\vee\boldsymbol{\xi}_{\nu,1,1}}\frac{1+\tanh(z/2)}{2}dz\right]^2.
\tag{8.6}
$$

Combining (8.5) and (8.6) and applying the Cauchy-Schwarz inequality, we obtain

$$\mathbb{E}\left[\log^2\left(\frac{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\rho,1,1}/2)}{1+\boldsymbol{r}_1\tanh(\boldsymbol{\xi}_{\nu,1,1}/2)}\right)\right]\leq\frac{1}{2}\mathbb{E}\left[\left(\boldsymbol{\xi}_{\rho,1,1}-\boldsymbol{\xi}_{\nu,1,1}\right)^2\right]. \tag{8.7}$$

An identical argument can be made for $\boldsymbol{\xi}_{\rho,1,2}$, $\boldsymbol{\xi}_{\nu,1,2}$, $\boldsymbol{\xi}'_{\rho,1,1}$, $\boldsymbol{\xi}'_{\nu,1,1}$ and $\boldsymbol{\xi}''_{\rho,1,2}$, $\boldsymbol{\xi}''_{\nu,1,2}$. Finally, (8.3),(8.4) and (8.7) yield

$$\begin{aligned}W_2(\hat{\rho},\hat{\nu})^2&\leq\frac{td}{2}\mathbb{E}\left[\left(\boldsymbol{\xi}_{\rho,1,1}-\boldsymbol{\xi}_{\nu,1,1}\right)^2+\left(\boldsymbol{\xi}_{\rho,1,2}-\boldsymbol{\xi}_{\nu,1,2}\right)^2\right]+\frac{(1-t)d}{2}\mathbb{E}\left[\left(\boldsymbol{\xi}'_{\rho,1,1}-\boldsymbol{\xi}'_{\nu,1,1}\right)^2+\left(\boldsymbol{\xi}''_{\rho,1,2}-\boldsymbol{\xi}''_{\nu,1,2}\right)^2\right]\\&=\frac{td}{2}W_2(\rho,\nu)^2+\frac{(1-t)d}{2}W_2(\rho,\nu)^2=\frac{d}{2}W_2(\rho,\nu)^2,\end{aligned} \tag{8.8}$$

which implies contraction because $d<2$. □

8.2. **Proof of Proposition 2.8.** The uniqueness of $\rho_{d,t}\in\mathscr{W}_2(\mathbb{R}^2)$ with $\mathtt{logBP}^{\otimes}_{d,t}(\rho_{d,t})=\rho_{d,t}$ (which yields (1.5)) follows from Lemma 8.1 and the Banach fixed point theorem. As the Dirac measure in zero is an element of $\mathscr{W}_2(\mathbb{R}^2)$, Lemma 8.1 also implies the weak convergence of $(\rho^{(\ell)}_{d,t})_{\ell\geq 0}$ to $\rho_{d,t}$.

9. Proof of Proposition 2.7

As a first step toward the proof of Proposition 2.7 we are going to introduce an operator on probability distributions on the unit square that resembles the Belief Propagation update equations (2.3)–(2.4). We will see that this operator is closely related to the operator from (1.3). Specifically, (1.3) is the log-likelihood version of the new operator. Subsequently, we will show that the Belief Propagation operator correctly implements marginal computations on the Galton-Watson trees $(\boldsymbol{T}_1,\boldsymbol{T}_2)$.

9.1. **Density evolution.** Recall that $\mathscr{P}((0,1)^2)$ is the space of all Borel probability measures on the unit square $(0,1)^2$. We define an operator

$$\mathrm{BP}^{\otimes}_{d,t}:\mathscr{P}((0,1)^2)\to\mathscr{P}((0,1)^2),\qquad\qquad \pi\mapsto\hat{\pi}=\mathrm{BP}^{\otimes}_{d,t}(\pi) \tag{9.1}$$

as follows. For $s\in\{\pm1\}$ let

$$(\boldsymbol{\mu}_{\pi,s,i,1},\boldsymbol{\mu}_{\pi,s,i,2})_{i\geq1},(\boldsymbol{\mu}'_{\pi,s,i,1},\boldsymbol{\mu}'_{\pi,s,i,2})_{i\geq1},(\boldsymbol{\mu}''_{\pi,s,i,1},\boldsymbol{\mu}''_{\pi,s,i,2})_{i\geq1}$$

be three sequences of random vectors with distribution π. Further, let $(\boldsymbol{d}_s,\boldsymbol{d}'_s,\boldsymbol{d}''_s)_{s\in\{\pm1\}}$ be Poisson variables with $\mathbb{E}[\boldsymbol{d}_s]=td/2$ and $\mathbb{E}[\boldsymbol{d}'_s]=\mathbb{E}[\boldsymbol{d}''_s]=(1-t)d/2$. Finally, let $((\boldsymbol{r}_{s,i},\boldsymbol{r}'_{s,i},\boldsymbol{r}''_{s,i}))_{s\in\{\pm1\},i\geq1}$ be uniformly distributed on $\{\pm1\}^3$. All of these random variables are mutually independent. Then $\hat{\pi}\in\mathscr{P}((0,1)^2)$ is the distribution of the random vector

$$\begin{aligned}&\left(\frac{2^{-\boldsymbol{d}_{-1}-\boldsymbol{d}'_{-1}}\prod_{i=1}^{\boldsymbol{d}_{-1}}(1+\boldsymbol{r}_{-1,i}(2\boldsymbol{\mu}_{\pi,-1,i,1}-1))\prod_{i=1}^{\boldsymbol{d}'_{-1}}(1+\boldsymbol{r}'_{-1,i}(2\boldsymbol{\mu}'_{\pi,-1,i,1}-1))}{\sum_{s\in\{\pm1\}}2^{-\boldsymbol{d}_s-\boldsymbol{d}'_s}\prod_{i=1}^{\boldsymbol{d}_s}(1+\boldsymbol{r}_{s,i}(2\boldsymbol{\mu}_{\pi,s,i,1}-1))\prod_{i=1}^{\boldsymbol{d}'_s}(1+\boldsymbol{r}'_{s,i}(2\boldsymbol{\mu}'_{\pi,s,i,1}-1))},\right.\\&\qquad\left.\frac{2^{-\boldsymbol{d}_{-1}-\boldsymbol{d}''_{-1}}\prod_{i=1}^{\boldsymbol{d}_{-1}}(1+\boldsymbol{r}_{-1,i}(2\boldsymbol{\mu}_{\pi,-1,i,2}-1))\prod_{i=1}^{\boldsymbol{d}''_{-1}}(1+\boldsymbol{r}''_{-1,i}(2\boldsymbol{\mu}''_{\pi,-1,i,2}-1))}{\sum_{s\in\{\pm1\}}2^{-\boldsymbol{d}_s-\boldsymbol{d}''_s}\prod_{i=1}^{\boldsymbol{d}_s}(1+\boldsymbol{r}_{s,i}(2\boldsymbol{\mu}_{\pi,s,i,2}-1))\prod_{i=1}^{\boldsymbol{d}''_s}(1+\boldsymbol{r}''_{s,i}(2\boldsymbol{\mu}''_{\pi,s,i,2}-1))}\right)\in(0,1)^2.\end{aligned} \tag{9.2}$$

Let $\mathrm{u}^{\otimes}\in\mathscr{P}((0,1)^2)$ denote the atom on the centre $(\frac{1}{2},\frac{1}{2})$ of the unit square. We write $\mathrm{BP}^{\otimes(\ell)}_{d,t}$ for the ℓ-fold application of the operator $\mathrm{BP}^{\otimes}_{d,t}$. We are going to perform a fixed point iteration using the operator $\mathrm{BP}^{\otimes}_{d,t}$, starting from $\mathrm{u}^{\otimes}$. This fixed point iteration is known as *density evolution* in physics jargon [36]. Let

$$\pi^{(\ell)}_{d,t}=\mathrm{BP}^{\otimes(\ell)}_{d,t}(\mathrm{u}^{\otimes}).$$

Lemma 9.1. *Let $d\in(0,2)$, $t\in[0,1]$ and set $\pi_{d,t}=\mathfrak{t}(\rho_{d,t})$, where $\rho_{d,t}$ is the unique fixed point of $\mathtt{logBP}^{\otimes}_{d,t}$ from Proposition 2.8.*

Then $\pi_{d,t}$ is a fixed point of $\mathrm{BP}^{\otimes}_{d,t}$, and

$$\pi_{d,t}=\lim_{\ell\to\infty}\pi^{(\ell)}_{d,t}.$$

Proof. For $s \in \{\pm 1\}$ let

$$(\boldsymbol{\xi}_{\rho,s,i,1}, \boldsymbol{\xi}_{\rho,s,i,2})_{i\geq 1}, (\boldsymbol{\xi}'_{\rho,s,i,1}, \boldsymbol{\xi}'_{\rho,s,i,2})_{i\geq 1}, (\boldsymbol{\xi}''_{\rho,s,i,1}, \boldsymbol{\xi}''_{\rho,s,i,2})_{i\geq 1} \tag{9.3}$$

be three sequences of random vectors with distribution $\rho_{d,t}$. Further, let $(\boldsymbol{d}_s, \boldsymbol{d}'_s, \boldsymbol{d}''_s)_{s\in\{\pm 1\}}$ be Poisson variables with $\mathbb{E}[\boldsymbol{d}_s] = td/2$ and $\mathbb{E}[\boldsymbol{d}'_s] = \mathbb{E}[\boldsymbol{d}''_s] = (1-t)d/2$ and let $(\boldsymbol{d}, \boldsymbol{d}', \boldsymbol{d}'')$ be Poisson variables with $\mathbb{E}[\boldsymbol{d}] = td$ and $\mathbb{E}[\boldsymbol{d}'] = \mathbb{E}[\boldsymbol{d}''] = (1-t)d$. Finally, let $((\boldsymbol{r}_{s,i}, \boldsymbol{r}'_{s,i}, \boldsymbol{r}''_{s,i}))_{s\in\{\pm1\}, i\geq 1}$, $((\boldsymbol{s}_i, \boldsymbol{s}'_i, \boldsymbol{s}''_i))_{i\geq 1}$ and $((\boldsymbol{r}_i, \boldsymbol{r}'_i, \boldsymbol{r}''_i))_{i\geq 1}$ all be uniformly distributed on $\{\pm 1\}^3$. All of these random variables are mutually independent. Throughout the proof, we write $\mathfrak{l} = (\mathfrak{l}_1, \mathfrak{l}_2)$ and $\mathfrak{t} = (\mathfrak{t}_1, \mathfrak{t}_2)$ for $\mathfrak{l}_h : (0,1) \to \mathbb{R}$, $\mathfrak{l}_h(x) = \log \frac{x}{1-x}$, and $\mathfrak{t}_h : \mathbb{R} \to (0,1)$, $\mathfrak{t}_h(x) = (1 + \tanh(x/2))/2$, where $h \in \{1,2\}$.

Then, since $\mathfrak{t}(\boldsymbol{\xi}_{\rho,s,1,1}, \boldsymbol{\xi}_{\rho,s,1,2}) = (\mathfrak{t}_2(\boldsymbol{\xi}_{\rho,s,1,1}), \mathfrak{t}_2(\boldsymbol{\xi}_{\rho,s,1,2}))$ has distribution $\mathfrak{t}(\rho_{d,t})$, using the definitions of $\mathfrak{l}$ and $\mathfrak{t}$,

$$\begin{aligned}
\mathrm{BP}^{\otimes}_{d,t}(\mathfrak{t}(\rho_{d,t})) &\overset{\text{dist}}{=} \Bigg(\frac{2^{-\boldsymbol{d}_{-1} - \boldsymbol{d}'_{-1}} \prod_{i=1}^{\boldsymbol{d}_{-1}} (1 + \boldsymbol{r}_{-1,i}(2\mathfrak{t}_1(\boldsymbol{\xi}_{\rho,-1,i,1}) - 1)) \prod_{i=1}^{\boldsymbol{d}'_{-1}} (1 + \boldsymbol{r}'_{-1,i}(2\mathfrak{t}_1(\boldsymbol{\xi}'_{\rho,-1,i,1}) - 1))}{\sum_{s\in\{\pm1\}} 2^{-\boldsymbol{d}_s - \boldsymbol{d}'_s} \prod_{i=1}^{\boldsymbol{d}_s} (1 + \boldsymbol{r}_{s,i}(2\mathfrak{t}_1(\boldsymbol{\xi}_{\rho,s,i,1}) - 1)) \prod_{i=1}^{\boldsymbol{d}'_s} (1 + \boldsymbol{r}'_{s,i}(2\mathfrak{t}_1(\boldsymbol{\xi}'_{\rho,s,i,1}) - 1))}, \\
&\qquad \frac{2^{-\boldsymbol{d}_{-1} - \boldsymbol{d}''_{-1}} \prod_{i=1}^{\boldsymbol{d}_{-1}} (1 + \boldsymbol{r}_{-1,i}(2\mathfrak{t}_1(\boldsymbol{\xi}_{\rho,-1,i,2}) - 1)) \prod_{i=1}^{\boldsymbol{d}''_{-1}} (1 + \boldsymbol{r}''_{-1,i}(2\mathfrak{t}_1(\boldsymbol{\xi}''_{\rho,-1,i,2}) - 1))}{\sum_{s\in\{\pm1\}} 2^{-\boldsymbol{d}_s - \boldsymbol{d}''_s} \prod_{i=1}^{\boldsymbol{d}_s} (1 + \boldsymbol{r}_{s,i}(2\mathfrak{t}_1(\boldsymbol{\xi}_{\rho,s,i,2}) - 1)) \prod_{i=1}^{\boldsymbol{d}''_s} (1 + \boldsymbol{r}''_{s,i}(2\mathfrak{t}_1(\boldsymbol{\xi}''_{\rho,s,i,2}) - 1))} \Bigg) \\
&\overset{\text{dist}}{=} \mathfrak{t} \begin{pmatrix} \sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \log \frac{1 + \boldsymbol{r}_i(2\mathfrak{t}_1(\boldsymbol{\xi}_{\rho,i,1}) - 1)}{2} + \sum_{i=1}^{\boldsymbol{d}'} \boldsymbol{s}'_i \log \frac{1 + \boldsymbol{r}'_i(2\mathfrak{t}_1(\boldsymbol{\xi}'_{\rho,i,1}) - 1)}{2} \\ \sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \log \frac{1 + \boldsymbol{r}_i(2\mathfrak{t}_1(\boldsymbol{\xi}_{\rho,i,2}) - 1)}{2} + \sum_{i=1}^{\boldsymbol{d}''} \boldsymbol{s}''_i \log \frac{1 + \boldsymbol{r}''_i(2\mathfrak{t}_1(\boldsymbol{\xi}''_{\rho,i,2}) - 1)}{2} \end{pmatrix} \\
&\overset{\text{dist}}{=} \mathfrak{t} \begin{pmatrix} \sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \log \frac{1 + \boldsymbol{r}_i \tanh(\boldsymbol{\xi}_{\rho,i,1}/2)}{2} + \sum_{i=1}^{\boldsymbol{d}'} \boldsymbol{s}'_i \log \frac{1 + \boldsymbol{r}'_i \tanh(\boldsymbol{\xi}'_{\rho,i,1}/2)}{2} \\ \sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \log \frac{1 + \boldsymbol{r}_i \tanh(\boldsymbol{\xi}_{\rho,i,2})/2)}{2} + \sum_{i=1}^{\boldsymbol{d}''} \boldsymbol{s}''_i \log \frac{1 + \boldsymbol{r}''_i \tanh(\boldsymbol{\xi}''_{\rho,i,2}/2)}{2} \end{pmatrix}.
\end{aligned}$$

Since $\rho_{d,t}$ is a fixed point of $\mathtt{logBP}^{\otimes}_{d,t}$, the last argument vector of $\mathfrak{t}$ has distribution $\rho_{d,t}$, and we get that

$$\mathrm{BP}^{\otimes}_{d,t}(\mathfrak{t}(\rho_{d,t})) = \mathfrak{t}(\rho_{d,t}).$$

So $\mathfrak{t}(\rho_{d,t})$ is a fixed point of $\mathrm{BP}^{\otimes}_{d,t}$.

Next, let $\mathtt{n}^{\otimes} \in \mathscr{P}(\mathbb{R}^2)$ denote the atom in $(0,0)$. Then since $\mathfrak{t}(0,0) = (\frac{1}{2}, \frac{1}{2})$, we have $\mathtt{u}^{\otimes} = \mathfrak{t}(\mathtt{n}^{\otimes})$. By a computation analogous to the first part of the proof, one can show inductively that for all $\ell \geq 1$, $\pi^{(\ell)}_{d,t} = \mathrm{BP}^{\otimes}_{d,t}(\mathfrak{t}(\mathtt{n}^{\otimes})) = \mathfrak{t}(\mathtt{logBP}^{\otimes\,(\ell)}_{d,t}(\mathtt{n}^{\otimes}))$. As

$$\rho_{d,t} = \lim_{\ell\to\infty} \mathtt{logBP}^{\otimes\,(\ell)}_{d,t}(\mathtt{n}^{\otimes}),$$

the second part of the claim now follows from the continuous mapping theorem. $\square$

9.2. **Belief Propagation on the Galton-Watson tree.** The proof of Proposition 2.7 relies on the fact that Belief Propagation is 'exact' on trees. The following fact, which is a direct consequence of [36, Theorem 14.1], furnishes the precise statement that we will use.

Fact 9.2. *Assume that the bipartite graph associated with the 2-CNF Φ is a (finite) tree. Let $z \in V(\Phi)$ be a variable and let $\ell \geq 1$ be an integer such that no variable or clause of Φ has distance greater than 2ℓ from z. Let*

$$\mu^{(0)}_{x\to a}(s) = \mu^{(0)}_{a\to x}(s) = \frac{1}{2} \qquad\qquad \textit{for all } x \in V(\Phi),\ a \in \partial x,\ s \in \{\pm 1\}.$$

Furthermore, obtain the messages $(\mu^{(i+1)}_{x\to a}(s), \mu^{(i+1)}_{a\to x}(s))_{x,a,s}$ by applying the BP operator (2.2) to $(\mu^{(i)}_{x\to a}(s), \mu^{(i)}_{a\to x}(s))_{x,a,s}$. Then for all $i \geq 2\ell$ we have

$$\mu_{\Phi}(\{\boldsymbol{\sigma}_z = s\}) = \frac{\prod_{a\in\partial z} \mu^{(i)}_{a\to z}(s)}{\prod_{a\in\partial z} \mu^{(i)}_{a\to z}(1) + \prod_{a\in\partial z} \mu^{(i)}_{a\to z}(-1)}. \tag{9.4}$$

As a preparation toward the proof of Proposition 2.7 we establish the following 'univariate' variant of the proposition.

Lemma 9.3. *Let $h = 1,2$. Let $\pi^{(\ell)}_{d,t,h}$ be the distribution of the h-th component of a random vector with distribution $\pi^{(\ell)}_{d,t}$. Then $\mu_{\boldsymbol{T}^{(2\ell)}_h}(\{\boldsymbol{\sigma}_o = 1\})$ has distribution $\pi^{(\ell)}_{d,t,h}$.*

Proof. We proceed by induction on ℓ. For $\ell = 0$ there is nothing to show because both $\mu_{\boldsymbol{T}_h^{(0)}}(\{\boldsymbol{\sigma}_o = 1\}) = \frac{1}{2}$ and $\pi_{d,t,h}^{(\ell)}$ is the atom on $1/2$. To go from $\ell - 1$ to $\ell \geq 1$ we exploit the fact that $\boldsymbol{T}_h$ by itself has the same distribution as the 'plain' Galton-Watson tree $\boldsymbol{T}$ from Section 2.2, after all distinctions between different types of clauses and variables are dropped. In effect, the tree $\boldsymbol{T}_{h,x}$ pending on any grandchild $x \in \partial^2 o$ of the root has the same distribution as $\boldsymbol{T}$ itself, and these trees are mutually independent for all $x \in \partial^2 o$. Consequently, by induction we know that the marginal $\mu_{\boldsymbol{T}_{h,x}^{(2(\ell-1))}}(\{\boldsymbol{\sigma}_x = 1\})$ of x in $\boldsymbol{T}_{h,x}^{(2(\ell-1))}$ has distribution $\pi_{d,t,h}^{(\ell-1)}$.

Now let $a_x \in \partial x \cap \partial o$ be the clause that links x and o. Fact 9.2 implies that the marginals $\mu_{\boldsymbol{T}_{h,x}^{(2(\ell-1))}}(\{\boldsymbol{\sigma}_x = s\})$ coincide with the messages $\mu_{\boldsymbol{T}^{(2\ell)},x\to a_x}^{(2(\ell-1))}(s)$. Indeed, the marginal formula (9.4) for the tree $\boldsymbol{T}_{h,x}^{(2(\ell-1))}$ coincides with the message update formula (2.4), because clause a_x is not part of $\boldsymbol{T}_{h,x}$. Furthermore, because the trees pending on the different grandchildren of the root o are mutually independent, the incoming messages $(\mu_{\boldsymbol{T}_{h,x}^{(2\ell)},x\to a_x}^{(2(\ell-1))}(\pm 1))_{x\in\partial^2 o}$ are mutually independent. Moreover, the quotient from (9.4), which, upon substituting in the update equation (2.3), can be rewritten as

$$\mu_{\boldsymbol{T}_h^{(2\ell)}}(\{\boldsymbol{\sigma}_o = 1\}) = \frac{\prod_{x\in\partial^2 o} \mathbb{1}\{\mathrm{sign}(o,a_x)=1\} + \mathbb{1}\{\mathrm{sign}(o,a_x)=-1\}\mu_{\boldsymbol{T}_{h,x}^{(2\ell)},x\to a_x}^{(2(\ell-1))}(\mathrm{sign}(x,a_x))}{\sum_{s\in\{\pm1\}}\prod_{x\in\partial^2 o}\mathbb{1}\{\mathrm{sign}(o,a_x)=s\} + \mathbb{1}\{\mathrm{sign}(o,a_x)=-s\}\mu_{\boldsymbol{T}_{h,x}^{(2\ell)},x\to a_x}^{(2(\ell-1))}(\mathrm{sign}(x,a_x))}. \tag{9.5}$$

Also recall that o has $\mathrm{Po}(d)$ children. Hence, comparing (9.5) with the h-component of (9.2), we conclude that $\mu_{\boldsymbol{T}_h^{(2\ell)}}(\{\boldsymbol{\sigma}_o = 1\})$ has distribution $\pi_{d,t,h}^{(\ell)}$. □

Lemma 9.4. *Let $t \in [0,1]$. Then $\boldsymbol{\mu}^{(2\ell)}$ has distribution $\pi_{d,t}^{(\ell)}$.*

Proof. As in the proof of Lemma 9.3 we proceed by induction on ℓ. For $\ell = 0$ there is nothing to show. To go from $\ell - 1$ to $\ell \geq 1$ we reuse the rewritten update equation (9.5). In each of the correlated trees $\boldsymbol{T}_1, \boldsymbol{T}_2$ the root o has $\mathrm{Po}((1-t)d)$ $\{1,2\}$-distinct grandchildren. By construction, the trees pending on these grandchildren are mutually independent copies of the tree $\boldsymbol{T}$. Hence, the same consideration as in the proof of Lemma 9.3 shows that the messages that the $\{1,2\}$-distinct grandchildren pass up are independent with distribution $\pi_{d,t,h}^{(\ell-1)}$. The same is true of the messages $\boldsymbol{\mu}'_{\pi_{d,t}^{(\ell-1)},\pm1,i,h}, \boldsymbol{\mu}''_{\pi_{d,t}^{(\ell-1)},\pm1,i,h}$ from (9.2). Consequently, the contribution of the $\{1,2\}$-distinct children in (9.2) matches the corresponding contribution to the update equation (9.5).

With respect to the shared grandchildren, we apply induction as in the proof of Lemma 9.3. Indeed, the trees pending on the shared grandchildren $x \in \partial^2_{\boldsymbol{T}^\otimes} o$ have the same distribution as the original tree $\boldsymbol{T}^\otimes$ and are mutually independent. Therefore, by the induction hypothesis, the pair of messages $(\mu_{\boldsymbol{T}_h^{(2\ell-2)},x\to a_x}^{(2(\ell-1))}(1))_{h=1,2}$ that a shared grandchild x sends towards the root has distribution $\pi_{d,t}^{(\ell-1)}$. Finally, since o has $\mathrm{Po}(dt)$ shared grandchildren, matching the expressions (9.5) and (9.2) completes the proof. □

Proof of Proposition 2.7. The assertion follows from Lemmas 9.1 and 9.4. □

10. Proof that the variance is finite

The main goal of this section is to show that both the evaluation of the functional $\mathscr{B}_{d,t}^\otimes$ on $\rho_{d,t}$ as well as the integration to obtain $\eta(d)^2$ yield finite values for any $d \in (0,2)$ and $t \in [0,1]$.

Lemma 10.1. *For any $d \in (0,2)$ and $t \in [0,1]$, $\mathscr{B}_{d,t}^\otimes(\rho_{d,t}) < \infty$. Moreover, for any $d \in (0,2)$, $\eta(d)^2 < \infty$.*

10.1. **Proof of Lemma 10.1.** Let $\rho_{d,t}^{(\ell)} \in \mathscr{W}_2(\mathbb{R}^2)$ be the result of ℓ iterations of $\mathtt{logBP}_{d,t}^\otimes$ launched from $\mathtt{n}^\otimes$, the atom at $(0,0)$. In the proof of Lemma 10.1, the following properties of the fixed point $\pi_{d,t}$ will be used:

Claim 10.2. *Let $\pi_{d,t} = \mathfrak{t}(\rho_{d,t})$ and $\boldsymbol{\mu}_{\pi_{d,t}} = (\boldsymbol{\mu}_{\pi_{d,t},1}, \boldsymbol{\mu}_{\pi_{d,t},2})$ be a random vector with distribution $\pi_{d,t}$. Then*

$$\boldsymbol{\mu}_{\pi_{d,t},1} \overset{\mathrm{dist}}{=} \boldsymbol{\mu}_{\pi_{d,t},2} \qquad \textit{and} \qquad \boldsymbol{\mu}_{\pi_{d,t}1} \overset{\mathrm{dist}}{=} 1 - \boldsymbol{\mu}_{\pi_{d,t},1}.$$

Proof. Recall the definition of $\mathrm{BP}_{d,t}^\otimes$ from (9.2). The first claim then follows from the following limiting argument: By Lemma 9.1, $\pi_{d,t} = \lim_{\ell\to\infty} \mathrm{BP}_{d,t}^{\otimes(\ell)}(\mathrm{u}^\otimes)$, where $\mathrm{u}^\otimes$ is the Dirac measure on $(1/2,1/2)$. As the marginal distributions of the initial distribution $\mathrm{u}^{\otimes}$ are identical, inspection of the update rule (9.2) yields that also the marginal distributions of $\mathrm{BP}_{d,t}^{\otimes(1)}(\mathrm{u}^{\otimes})$ are identical. Analogously, it is immediate from (9.2) that any π with two identical marginals will be mapped to a measure $\mathrm{BP}_{d,t}^{\otimes}(\pi)$ with two identical marginals, such that the marginal distributions of $\mathrm{BP}_{d,t}^{\otimes(\ell)}(\mathrm{u}^{\otimes})$ for any $\ell \geq 0$ are identical. Hence, also in the limit,

$$\boldsymbol{\mu}_{\pi_{d,t},1} \stackrel{\text{dist}}{=} \boldsymbol{\mu}_{\pi_{d,t},2}.$$

On the other hand, Lemma 9.1 also implies that $\pi_{d,t} = \mathrm{BP}_{d,t}^{\otimes}(\pi_{d,t})$, so that the distribution of $\boldsymbol{\mu}_{\pi_{d,t},1}$ is the same as the distribution of

$$\frac{2^{-\boldsymbol{d}_{-1}-\boldsymbol{d}'_{-1}} \prod_{i=1}^{\boldsymbol{d}_{-1}}(1+\boldsymbol{r}_{-1,i}(2\boldsymbol{\mu}_{\pi_{d,t},-1,i,1}-1)) \prod_{i=1}^{\boldsymbol{d}'_{-1}}(1+\boldsymbol{r}'_{-1,i}(2\boldsymbol{\mu}'_{\pi_{d,t},-1,i,1}-1))}{\sum_{s\in\{\pm1\}} 2^{-\boldsymbol{d}_{-s}-\boldsymbol{d}'_{-s}} \prod_{i=1}^{\boldsymbol{d}_{-s}}(1+\boldsymbol{r}_{-s,i}(2\boldsymbol{\mu}_{\pi_{d,t},-s,i,1}-1)) \prod_{i=1}^{\boldsymbol{d}'_{-s}}(1+\boldsymbol{r}'_{-s,i}(2\boldsymbol{\mu}'_{\pi_{d,t},-s,i,1}-1))} \tag{10.1}$$

while the distribution of $1-\boldsymbol{\mu}_{\pi_{d,t},1}$ is the same as the distribution of

$$\frac{2^{-\boldsymbol{d}_{1}-\boldsymbol{d}'_{1}} \prod_{i=1}^{\boldsymbol{d}_{1}}(1+\boldsymbol{r}_{1,i}(2\boldsymbol{\mu}_{\pi_{d,t},1,i,1}-1)) \prod_{i=1}^{\boldsymbol{d}'_{1}}(1+\boldsymbol{r}'_{1,i}(2\boldsymbol{\mu}'_{\pi_{d,t},1,i,1}-1))}{\sum_{s\in\{\pm1\}} 2^{-\boldsymbol{d}_{-s}-\boldsymbol{d}'_{-s}} \prod_{i=1}^{\boldsymbol{d}_{-s}}(1+\boldsymbol{r}_{-s,i}(2\boldsymbol{\mu}_{\pi_{d,t},-s,i,1}-1)) \prod_{i=1}^{\boldsymbol{d}'_{-s}}(1+\boldsymbol{r}'_{-s,i}(2\boldsymbol{\mu}'_{\pi_{d,t},-s,i,1}-1))}. \tag{10.2}$$

This immediately shows that (10.1) and (10.2) have the same distribution. As a consequence, the second claim holds as well. $\square$

Proof of Lemma 10.1. Recall that $\pi_{d,t} = \mathfrak{t}(\rho_{d,t})$. Let

$$\boldsymbol{\mu}_{\pi_{d,t},1} = (\boldsymbol{\mu}_{\pi_{d,t},1,1}, \boldsymbol{\mu}_{\pi_{d,t},1,2}) = \mathfrak{t}(\boldsymbol{\xi}_{\rho_{d,t},1,1}, \boldsymbol{\xi}_{\rho_{d,t},1,2}),$$

and

$$\boldsymbol{\mu}_{\pi_{d,t},2} = (\boldsymbol{\mu}_{\pi_{d,t},2,1}, \boldsymbol{\mu}_{\pi_{d,t},2,2}) = \mathfrak{t}(\boldsymbol{\xi}_{\rho_{d,t},2,1}, \boldsymbol{\xi}_{\rho_{d,t},2,2}).$$

Then they are independent random vectors with distribution $\pi_{d,t}$. Let $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ be independent Rademacher random variables with parameter $1/2$, independent of $\boldsymbol{\mu}_{\pi_{d,t},1}$ and $\boldsymbol{\mu}_{\pi_{d,t},2}$. Conditioning on the values of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ yields the upper bound

$$\begin{aligned}
|\mathscr{B}_{d,t}^{\otimes}(\rho_{d,t})| \leq{}& \frac{1}{4}\mathbb{E}\left[\left|\log\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\log\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,2}\boldsymbol{\mu}_{\pi_{d,t},2,2}\right)\right|\right] \\
&+\frac{1}{4}\mathbb{E}\left[\left|\log\left(1-\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\log\left(1-\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,2}\right)\boldsymbol{\mu}_{\pi_{d,t},2,2}\right)\right|\right] \\
&+\frac{1}{4}\mathbb{E}\left[\left|\log\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\left(1-\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\right)\log\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,2}\left(1-\boldsymbol{\mu}_{\pi_{d,t},2,2}\right)\right)\right|\right] \\
&+\frac{1}{4}\mathbb{E}\left[\left|\log\left(1-\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)\left(1-\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\right)\log\left(1-\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,2}\right)\left(1-\boldsymbol{\mu}_{\pi_{d,t},2,2}\right)\right)\right|\right].
\end{aligned}$$

The Cauchy-Schwarz inequality further gives that

$$\begin{aligned}
|\mathscr{B}_{d,t}^{\otimes}(\rho_{d,t})| \leq{}& \frac{1}{4}\mathbb{E}\left[\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\right]^{1/2}\mathbb{E}\left[\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,2}\boldsymbol{\mu}_{\pi_{d,t},2,2}\right)\right]^{1/2} \\
&+\frac{1}{4}\mathbb{E}\left[\log^2\left(1-\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\right]^{1/2}\mathbb{E}\left[\log^2\left(1-\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,2}\right)\boldsymbol{\mu}_{\pi_{d,t},2,2}\right)\right]^{1/2} \\
&+\frac{1}{4}\mathbb{E}\left[\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\left(1-\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\right)\right]^{1/2}\mathbb{E}\left[\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,2}\left(1-\boldsymbol{\mu}_{\pi_{d,t},2,2}\right)\right)\right]^{1/2} \\
&+\frac{1}{4}\mathbb{E}\left[\log^2\left(1-\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)\left(1-\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\right)\right]^{1/2}\mathbb{E}\left[\log^2\left(1-\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,2}\right)\left(1-\boldsymbol{\mu}_{\pi_{d,t},2,2}\right)\right)\right]^{1/2}.
\end{aligned}$$

As $\boldsymbol{\mu}_{\pi_{d,t},1,1} \stackrel{\text{dist}}{=} \boldsymbol{\mu}_{\pi_{d,t},1,2}$ and $\boldsymbol{\mu}_{\pi_{d,t},1,1} \stackrel{\text{dist}}{=} 1-\boldsymbol{\mu}_{\pi_{d,t},1,1}$ thanks to Claim 10.2, we further get

$$|\mathscr{B}_{d,t}^{\otimes}(\rho_{d,t})| \leq \mathbb{E}\left[\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\right].$$

Next, recalling (9.3),

$$\begin{aligned}
&\mathbb{E}\left[\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\boldsymbol{\mu}_{\pi_{d,t},2,1}\right)\right] \le \mathbb{E}\left[\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)\right]\\
&\le \mathbb{E}\left[\mathbb{1}\left\{\boldsymbol{\mu}_{\pi_{d,t},1,1} \le \frac{1}{2}\right\}\left|\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)\right|\right] + \mathbb{E}\left[\mathbb{1}\left\{\boldsymbol{\mu}_{\pi_{d,t},1,1} > \frac{1}{2}\right\}\left|\log^2\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)\right|\right]\\
&\le \mathbb{E}\left[\mathbb{1}\left\{\boldsymbol{\mu}_{\pi_{d,t},1,1} \le \frac{1}{2}\right\}\log^2 2\right] + \mathbb{E}\left[\mathbb{1}\left\{\boldsymbol{\mu}_{\pi_{d,t},1,1} > \frac{1}{2}\right\}\left(\log\frac{\boldsymbol{\mu}_{\pi_{d,t},1,1}}{\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)} - \log\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)^2\right]\\
&\le \log^2 2 + \mathbb{E}\left[\log^2\left(\frac{\boldsymbol{\mu}_{\pi_{d,t},1,1}}{1-\boldsymbol{\mu}_{\pi_{d,t},1,1}}\right)\right] - 2\mathbb{E}\left[\mathbb{1}\left\{\boldsymbol{\mu}_{\pi_{d,t},1,1} > \frac{1}{2}\right\}\log\frac{\boldsymbol{\mu}_{\pi_{d,t},1,1}}{\left(1-\boldsymbol{\mu}_{\pi_{d,t},1,1}\right)}\log\boldsymbol{\mu}_{\pi_{d,t},1,1}\right]\\
&\le \log^2 2 + \mathbb{E}\left[\log^2\left(\frac{\boldsymbol{\mu}_{\pi_{d,t},1,1}}{1-\boldsymbol{\mu}_{\pi_{d,t},1,1}}\right)\right] = \log^2 2 + \mathbb{E}\left[\boldsymbol{\xi}^2_{\rho_{d,t},1,1}\right].
\end{aligned}$$

However $\mathbb{E}[\boldsymbol{\xi}^2_{\rho_{d,t},1,1}] < \infty$, since $\rho_{d,t} \in \mathscr{W}_2(\mathbb{R}^2)$. Thus, $\mathscr{B}^{\otimes}_{d,t}(\rho_{d,t}) < \infty$.

Moreover, it is easy to see that the distribution of the marginal sample $\boldsymbol{\xi}_{\rho_{d,t},1,1}$ is independent of for any $t \in [0,1]$. Call this distribution ρ_d and let $\boldsymbol{\xi}_{\rho_d}$ be a sample from ρ_d. Then the previous upper bound yields

$$\eta(d)^2 \le \int_0^1 |\mathscr{B}^{\otimes}_{d,t}(\rho_{d,t})|\mathrm{d}t + |\mathscr{B}^{\otimes}_{d}(\rho_{d,0})| \le \int_0^1 \log^2 2 + \mathbb{E}\left[\boldsymbol{\xi}^2_{\rho_d}\right]\mathrm{d}t + \log^2 2 + \mathbb{E}\left[\boldsymbol{\xi}^2_{\rho_d}\right] \le 2\left(\log^2 2 + \mathbb{E}\left[\boldsymbol{\xi}^2_{\rho_d}\right]\right) < \infty.$$

□

11. Proof of Proposition 2.12

We combine the results from the previous sections in order to analyse the variance process. As a first step we derive a rough upper bound on the potential change in the number of satisfying assignments upon insertion of a single clause (Lemmas 11.1 and 11.6). Subsequently we derive a combinatorial formula for the squared martingale difference (Lemma 11.9), which easily implies Lemma 2.4. The combinatorial formula puts us in a position to obtain an L^2-bound on the squared martingale difference (Lemma 11.10). With these ingredients in place, we complete the proof of Proposition 2.12 and of Theorem 1.1 in Section 11.5.

11.1. **A pessimistic estimate.** Let Φ, Ψ be two 2-CNFs on the same set of variables such that Ψ is obtained from Φ by adding a single clause e. We are going to need a baseline estimate of the difference $|\log Z(\hat\Phi) - \log Z(\hat\Psi)|$. The principal difficulty here is to assess the impact of the additional clause on the pruning operation. The issue is that the extra clause may also cause additional pruning. Indeed, while clearly

$$F(\hat\Psi) \setminus \{e\} \subseteq F(\hat\Phi),$$

i.e., any clause $a \neq e$ that survives pruning on Ψ also remains present in $\hat\Phi$, the pruned formula $\hat\Psi$ may ironically end up having strictly fewer clauses than $\hat\Phi$.

To get a handle on the potential repercussions of pruning, let $\{v, v'\} = \partial e$ be the variables that appear in clause e. Let $\mathscr{N}(\Phi, v)$ be the set of all literals l such that $\{v, \neg v\} \cap \mathscr{L}(\Phi, \{l\}) \neq \emptyset$. Thus, PUC may reach v or $\neg v$ once l is deemed true. Observe that $v \in \mathscr{N}(\Phi, v)$ and $\neg v \in \mathscr{N}(\Phi, v)$. Define $\mathscr{N}(\Phi, v')$ analogously. Further, let

$$\mathscr{N}(\Phi, e) = \bigcup_{l \in \mathscr{N}(\Phi,v)\cup\mathscr{N}(\Phi,v')} \mathscr{L}(\Phi, \{l\}). \tag{11.1}$$

Thus, $\mathscr{N}(\Phi, e)$ contains all literals that PUC can reach by tracing the implications of a literal from $\mathscr{N}(\Phi, v) \cup \mathscr{N}(\Phi, v')$. The definition of the sets $\mathscr{N}(\Phi, v)$, $\mathscr{N}(\Phi, v')$ ensures that

$$v, \neg v \in \mathscr{N}(\Phi, v), \qquad\qquad v', \neg v' \in \mathscr{N}(\Phi, v'). \tag{11.2}$$

Lemma 11.1. *Let Φ be a 2-CNF formula. Suppose that Ψ is obtained from Φ by adding a single clause e. Then*

$$\left|\log(Z(\hat\Phi)) - \log(Z(\hat\Psi))\right| \le |\mathscr{N}(\Phi, e)| \log 2. \tag{11.3}$$

As a first step towards the proof of Lemma 11.1 we observe that (11.1) can be rewritten as follows.

Claim 11.2. *We have*

$$\mathcal{N}(\Phi, e) = \bigcup_{l \in \mathcal{N}(\Phi,v) \cup \mathcal{N}(\Phi,v')} \mathcal{L}(\Psi, \{l\}). \tag{11.4}$$

Proof. Clearly $\mathcal{L}(\Psi,\{l\}) \supseteq \mathcal{L}(\Phi,\{l\})$ for every literal l. Hence, we just need to show that

$$\mathcal{L}(\Psi,\{l\}) \subseteq \mathcal{N}(\Phi,e) \qquad\qquad \text{for all } l \in \mathcal{N}(\Phi,v) \cup \mathcal{N}(\Phi,v'). \tag{11.5}$$

Hence, let $l \in \mathcal{N}(\Phi,v) \cup \mathcal{N}(\Phi,v')$ and let $l' \in \mathcal{L}(\Psi,\{l\})$. Then Lemma 4.2 shows that there exists an implication chain

$$l = l_0, a_1, l_1, \dots, a_k, l_k = l' \tag{11.6}$$

comprising literals l_i and clauses $a_i \in F(\Psi)$ such that $a_i \equiv l_{i-1} \to l_i$ for all $1 \le i \le k$. If $a_i \neq e$ for all i, then the chain (11.6) is contained in Φ and thus $l' \in \mathcal{L}(\Phi,\{l\}) \subseteq \mathcal{N}(\Phi,e)$. Otherwise let $1 \le j \le k$ be the largest index such that $a_j = e$. Then l_j is one of the constituent literals of a_j and thus $l_j \in \{v, \neg v, v', \neg v'\}$. Furthermore, the implication chain $l_j, a_{j+1}, l_{j+1}, \dots, a_k, l_k = l'$ from l_j to l' is contained in Φ. Therefore, (11.2) shows in combination with Lemma 4.2 and (11.1) that $l' \in \mathcal{L}(\Phi,\{l_j\}) \subseteq \mathcal{N}(\Phi,e)$. □

We proceed to show that $\mathcal{N}(\Phi,e)$ contains the variables of all clauses $a \in F(\Phi)$ on which the pruning processes run on Φ, Ψ differ.

Claim 11.3. *For any clause $a \in F(\hat\Phi) \setminus F(\hat\Psi)$ we have $\partial a \subseteq \mathcal{N}(\Phi,e)$.*

Proof. Consider a clause $a \in F(\Phi)$ that was removed by pruning applied to Ψ but not by pruning applied to Φ. Let w, w' be the constituent literals of a, i.e., $a \equiv w \vee w'$. Then $\mathtt{PUC}(\Psi,\{l\})$ added a to the set $\mathcal{C}(\Psi,\{l\})$ of conflict clauses for some literal l. Hence,

$$w, \neg w, w', \neg w' \in \mathcal{L}(\Psi,\{l\}). \tag{11.7}$$

Consequently, Lemma 4.2 shows that for each literal $k \in \{\neg w, \neg w'\}$, $\mathtt{PUC}(\Psi,\{l\})$ traverses an implication chain

$$l_{0,k} = l, a_{0,k}, l_{1,k}, a_{1,k}, \dots, l_{j_k,k} = k$$

of literals $l_{i,k} \in \mathcal{L}(\Psi,\{l\})$ and clauses $a_{i,k} \equiv \neg l_{i,k} \vee l_{i+1,k} \equiv l_{i,k} \to l_{i+1,k}$ for $0 \le i < j_k$. Because $a \notin \mathcal{C}(\Phi,l)$, at least one of these two sequences includes the clause e and thus at least one of $v, \neg v$ and one of $v', \neg v'$. Hence, $l \in \mathcal{N}(\Phi,v) \cup \mathcal{N}(\Phi,v')$. Therefore, combining (11.4) and (11.7), we conclude that $\partial a = \{|w|, |w'|\} \subseteq \mathcal{N}(\Phi,e)$. □

Let $\tilde\Phi$ be the formula obtained from $\hat\Phi$ by removing all variables $x \in V(\hat\Phi)$ such that $\{x, \neg x\} \cap \mathcal{N}(\Phi,e) \neq \emptyset$ along with their adjacent clauses.

Claim 11.4. *For any $\tilde\sigma \in S(\tilde\Phi)$ there exists $\sigma \in S(\hat\Psi)$ such that $\sigma_x = \tilde\sigma_x$ for all $x \in V(\tilde\Phi)$.*

Proof. Let $\check\Psi$ be a CNF with variables

$$V(\check\Psi) = V(\hat\Psi) \setminus V(\tilde\Phi) = \{x \in V(\hat\Phi) : \{x, \neg x\} \cap \mathcal{N}(\Phi,e) \neq \emptyset\}.$$

The clauses of $\check\Psi$ include all $a \in F(\hat\Psi)$ such that $\partial a \subseteq V(\check\Psi)$. Additionally, for every clause $a \in F(\hat\Psi)$ that contains exactly one literal l with $|l| \in V(\check\Psi)$ we include the literal l as a unit clause into $\check\Psi$. In light of Claim 11.3, to prove the assertion it suffices to show that $\check\Psi$ is satisfiable. For then we could extend any $\sigma \in S(\tilde\Phi)$ to a satisfying assignment of $\hat\Psi$ by simply setting the variables $x \in V(\hat\Psi) \setminus V(\tilde\Phi)$ in accordance with a satisfying assignment of $\check\Psi$.

As in the proof of Fact 2.2, to construct a satisfying assignment of $\check\Psi$ we fix an order $l_1, \dots, l_k$ of the literals $\mathcal{N}(\Phi,e)$. Let σ_i be the assignment that $\mathtt{PUC}$ outputs on input $\Psi, \{l_i\}$. Further, define a $\{0, \pm 1\}$-valued assignment $(\sigma_x)_{x \in V(\check\Psi)}$ by letting $\sigma_x = \sigma_{i,x}$ for the least index i such that $\{x, \neg x\} \cap \mathcal{L}(\Psi, l_i) \neq \emptyset$.

We claim that

$$\forall a \in F(\check\Psi)\, \exists x \in \partial_{\check\Psi} a : \sigma_x = \mathrm{sign}(x,a) \ ; \tag{11.8}$$

thus, we can turn σ into a satisfying assignment of $\check\Psi$ by assigning those variables y with $\sigma_y = 0$ arbitrarily. To verify (11.8), we consider two cases separately.

Case 1: $|\partial_{\check\Psi} a| = 2$**:** then $a \in F(\Psi)$. Let $\partial a = \{x, x'\}$ and let i be the smallest index such that $\mathcal{L}(\Psi, l_i) \cap \{x, \neg x, x', \neg x'\} \neq \emptyset$. Also let l, l' be the constitutent literals of a such that $|l| = x$ and $|l'| = x'$. Suppose that $l \in \mathcal{L}(\Psi, \{l_i\})$. If $\neg l \notin \mathcal{L}(\Psi, l_i)$, then $\sigma_x = \mathrm{sign}(x, a)$ by construction. Hence, assume that $l, \neg l \in \mathcal{L}(\Psi, l_i)$. Then the construction in Steps 1–2 of PUC ensures that $l' \in \mathcal{L}(\Psi, l_i)$ as well. Moreover, if $\neg l' \notin \mathcal{L}(\Psi, l_i)$, then $\sigma_{x'} = \mathrm{sign}(x', a)$. Finally, the case $l, l', \neg l, \neg l' \in \mathcal{L}(\Psi, l_i)$ cannot occur because otherwise a would have been pruned, i.e., $a \notin F(\hat\Psi)$.

Case 2: $|\partial_{\check\Psi} a| = 1$**:** there exists a clause $b \in F(\hat\Psi)$ and literals l, l' with $|l'| \notin V(\check\Psi)$ such that $b = l \vee l'$ and $a = l$. Let i be the least index such that $\{l, \neg l\} \cap \mathcal{L}(\Psi, \{l_i\}) \neq \emptyset$. If $\neg l \in \mathcal{L}(\Psi, \{l_i\})$, then $\mathrm{PUC}(\Psi, \{l_i\})$ would have added l' to the set $\mathcal{L}(\Psi, \{l_i\})$ as well and thus $|l'| \in V(\check\Psi)$. But $|l'| \notin V(\check\Psi)$. Hence, $\{l, \neg l\}\mathcal{L}(\Psi, \{l_i\}) = \{l\}$ and thus $\sigma_{|l|} = \sigma_{i,|l|} = \mathrm{sign}_{\Psi}(|l|, b) = \mathrm{sign}_{\check\Psi}(|l|, a)$.

Thus, in either case σ satisfies clause a. $\square$

In perfect analogy to the above let $\tilde\Psi$ be the formula obtained from $\hat\Psi$ by removing all variables $x \in V(\hat\Psi)$ such that $\{x, \neg x\} \cap \mathcal{N}(\Phi, e) \neq \emptyset$, along with their adjacent clauses.

Claim 11.5. *For any $\tilde\sigma \in S(\tilde\Psi)$ there exists $\sigma \in S(\hat\Phi)$ such that $\sigma_x = \tilde\sigma_x$ for all $x \in V(\tilde\Psi)$.*

Proof. Let $\check\Phi$ be a CNF with variables $V(\check\Phi) = V(\hat\Phi) \setminus V(\tilde\Psi)$. Include in $\check\Phi$ all $a \in F(\hat\Phi)$ with $\partial a \subseteq V(\check\Phi)$. Moreover, for every $a \in F(\hat\Phi)$ that contains exactly one literal l with $|l| \in V(\check\Psi)$ add l as a clause to $\check\Phi$. As in the proof of Claim 11.4 it suffices to construct a satisfying assignment of $\check\Phi$. Due to (11.1) the same argument as in the proof of Claim 11.4 extends. $\square$

Proof of Lemma 11.1. We use Claim 11.4 to prove that $Z(\hat\Phi) \leq 2^{|\mathcal{N}(\Phi,e)|} Z(\hat\Psi)$; similar reasoning based on Claim 11.5 yields the reverse bound. To show the desired bound split a satisfying assignment $\sigma \in S(\hat\Phi)$ up into two parts $\tilde\sigma = (\sigma_x)_{\{x,\neg x\} \cap \mathcal{N}(\Phi,e) = \emptyset}$, $\check\sigma = (\sigma_x)_{\{x,\neg x\} \cap \mathcal{N}(\Phi,e) \neq \emptyset}$. Claim 11.4 shows that the number of possible first parts $\tilde\sigma$ for $\sigma \in S(\hat\Phi)$ is bounded by $Z(\hat\Psi)$, because every $\tilde\sigma$ extends to a satisfying assignment of $\hat\Psi$. Moreover, the total number of possible second parts is bounded by $2^{|\mathcal{N}(\Phi,e)|}$. $\square$

11.2. **A tail bound.** As a next step we are going to derive a bound on the r.h.s. of (11.3) on random formulas. More specifically, obtain the formula $\mathbf{\Phi}'$ from $\mathbf{\Phi}$ by deleting the last clause $\boldsymbol{a}_m$. Let $\boldsymbol{N}' = |\mathcal{N}(\mathbf{\Phi}', \boldsymbol{a}_m)|$.

Lemma 11.6. *There exists $c = c(d) > 0$ such that for all $t > c$ we have*

$$\mathbb{P}\left[\boldsymbol{N}' > t^2\right] \leq c \exp(-t/c).$$

As a first step we are going to estimate the size of the set $\mathcal{N}(\mathbf{\Phi}', x_1)$ that contains all literals l such that $\mathcal{L}(\mathbf{\Phi}', l) \cap \{x_1, \neg x_1\} \neq \emptyset$.

Claim 11.7. *There exists $c_1 = c_1(d) > 0$ such that for all $t > c_1$ we have* $\mathbb{P}\left[|\mathcal{N}(\mathbf{\Phi}', x_1)| > t\right] \leq c_1 \exp(-t/c_1)$.

Proof. We use a classical branching process argument. Let $\mathcal{R}$ be the set of literals l such that $x_1 \in \mathcal{L}(\mathbf{\Phi}', l)$. By symmetry it suffices to bound $|\mathcal{R}|$.

For every $l \in \mathcal{R}$ there exists an alternating sequence $l = l_0, a_1, l_1, a_2, \ldots, l_k = x_1$ of literals and clauses such that $a_i \equiv \neg l_{i-1} \vee l_i$. Flipping the negations along this sequence yields a reverse sequence $l'_0 = \neg x_1 = \neg l_k, a'_1 = a_k, l'_1 = \neg l_{k-1}, \ldots, l'_k = \neg l$ such that $a'_i \equiv \neg l'_{i-1} \vee l'_i$. Hence, $\mathcal{R}$ is precisely the set of literals l that are reachable from x_1 via such an alternating sequence $l'_0, a'_1, \ldots, l'_k$. Furthermore, for any literal l the expected number of clauses $\boldsymbol{a}_i$ such that $\boldsymbol{a}_i \equiv l \vee l'$ for some other literal l' equals $m/2n \sim d/2$. Therefore, $|\mathcal{R}|$ is stochastically dominated by the progeny of a branching process with offspring $\mathrm{Po}(d/2)$. Standard branching process tail bounds therefore yield the desired bound on $|\mathcal{R}|$. $\square$

Claim 11.8. *There exists $c_2 = c_2(d) > 0$ such that for all $t > c_2$ and for every literal $l \neq x_1$ we have*

$$\mathbb{P}\left[|\mathcal{L}(\mathbf{\Phi}', l)| > t \mid x_1 \in \mathcal{L}(\mathbf{\Phi}', l)\right] \leq c_2 \exp(-t/c_2).$$

Proof. We combine a branching process argument with Bayes' formula. Specifically, because the formula $\mathbf{\Phi}'$ is random, the set $\mathcal{L}(\mathbf{\Phi}', l) \setminus \{l\}$ is random given its size. Hence, for an integer ℓ we have

$$\mathbb{P}\left[x_1 \in \mathcal{L}(\mathbf{\Phi}', l) \mid |\mathcal{L}(\mathbf{\Phi}', l)| = \ell\right] = \frac{\ell - 1}{2n - 1}. \tag{11.9}$$

Furthermore, the size $|\mathscr{L}(\mathbf{\Phi}', l)|$ is stochastically dominated by the progeny of a branching process with offspring $\mathrm{Po}(d/2)$. Therefore, there exists $c_2' = c_2'(d) > 0$ such that for all $t > c_2'$ we have

$$\mathbb{P}\left[|\mathscr{L}(\mathbf{\Phi}', l)| > t\right] \leq c_2' \exp(-t/c_2'). \tag{11.10}$$

Moreover, for any $d > 0$ there exists $c_2'' = c_2''(d) > 0$ such that

$$\mathbb{P}\left[x_1 \in \mathscr{L}(\mathbf{\Phi}', l)\right] \geq c_2''/n. \tag{11.11}$$

Hence, combining (11.9)–(11.11) with Bayes' rule, we obtain for $\ell > c_2'$,

$$\mathbb{P}\left[|\mathscr{L}(\mathbf{\Phi}', l)| = \ell \mid x_1 \in \mathscr{L}(\mathbf{\Phi}', l)\right] \leq \frac{\mathbb{P}\left[x_1 \in \mathscr{L}(\mathbf{\Phi}', \ell) \mid |\mathscr{L}(\mathbf{\Phi}', l)| = \ell\right] \mathbb{P}\left[|\mathscr{L}(\mathbf{\Phi}', l)| = \ell\right]}{\mathbb{P}\left[x_1 \in \mathscr{L}(\mathbf{\Phi}', l)\right]} \leq \frac{c_2'}{c_2''} \ell \exp(-\ell/c_2'),$$

which implies the assertion. □

Proof of Lemma 11.6. Let $\mathfrak{R}(\ell)$ be the event that there exists $l \in \mathscr{N}(\mathbf{\Phi}', x_1) \setminus \{x_1\}$ such that $|\mathscr{L}(\mathbf{\Phi}', l)| > \ell$. Claim 11.7 implies that there exists $c_3 = c_3(d) > 0$ such that

$$\mathbb{P}\left[x_1 \in \mathscr{L}(\mathbf{\Phi}', l)\right] \leq c_3/n. \tag{11.12}$$

Hence, by Claim 11.8, (11.12) and the union bound,

$$\mathbb{P}\left[\mathfrak{R}(\ell)\right] \leq \sum_{l \neq x_1} \mathbb{P}\left[x_1 \in \mathscr{L}(\mathbf{\Phi}', l), |\mathscr{L}(\mathbf{\Phi}', l)| > \ell\right] \leq 2c_2 c_3 \exp(-\ell/c_2). \tag{11.13}$$

Furthermore, Claim 11.7 shows that

$$\mathbb{P}\left[\mathscr{N}(\mathbf{\Phi}', x_1) \setminus \{x_1\} > \ell\right] \leq c_1 \exp(-\ell/c_1). \tag{11.14}$$

Combining (11.13) and (11.14), we obtain

$$\mathbb{P}\left[\sum_{l \in \mathscr{N}(\mathbf{\Phi}', x_1)} |\mathscr{L}(\mathbf{\Phi}', l)| > \ell^2\right] \leq \mathbb{P}\left[\mathfrak{R}(\ell)\right] + \mathbb{P}\left[\mathscr{N}(\mathbf{\Phi}', x_1) \setminus \{x_1\} > \ell\right] \leq c_1 \exp(-\ell/c_1) + 2c_2 c_3 \exp(-\ell/c_2). \tag{11.15}$$

By symmetry the same bound holds with x_1 replaced by $\neg x_1$. Therefore, the assertion follows from (11.15) and the union bound. □

11.3. **The squared martingale difference.** We derive a combinatorial formula for the squared martingale differences $\boldsymbol{X}_i^2$. Let

$$\begin{aligned}
\mathbf{\Delta}(M) &= \log\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M, m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1, m-M))}\right) \cdot \log\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M, m-M))}{Z(\hat{\mathbf{\Phi}}_2(M-1, m-M))}\right), \\
\mathbf{\Delta}'(M) &= \log\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M-1, m-M+1))}{Z(\hat{\mathbf{\Phi}}_1(M-1, m-M))}\right) \cdot \log\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M-1, m-M+1))}{Z(\hat{\mathbf{\Phi}}_2(M-1, m-M))}\right), \\
\mathbf{\Delta}''(M) &= \log\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M, m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1, m-M))}\right) \cdot \log\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M-1, m-M+1))}{Z(\hat{\mathbf{\Phi}}_2(M-1, m-M))}\right).
\end{aligned}$$

Lemma 11.9. *We have* $m\boldsymbol{X}_M^2 = \mathbb{E}\left[\mathbf{\Delta}(M) + \mathbf{\Delta}(M)' - 2\mathbf{\Delta}''(M) \mid \mathfrak{F}_M\right]$.

Proof. This follows from a direct computation. □

Proof of Lemma 2.4. Lemma 2.4 is an immediate consequence of Lemma 11.9. □

11.4. **An L^2-bound.** The following L^2-bound will enable us to deal with error terms.

Lemma 11.10. *Uniformly for all* $1 \leq M \leq m$ *we have* $\mathbb{E}\left[\mathbf{\Delta}(M)^2 + \mathbf{\Delta}'(M)^2 + \mathbf{\Delta}''(M)^2\right] = O(1)$.

Proof. We will bound $\mathbb{E}[\mathbf{\Delta}(M)^2]$; the bounds on the other two terms follow analogously. Invoking the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\mathbb{E}\left[\mathbf{\Delta}(M)^2\right] &= \mathbb{E}\left[\log^2\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M,m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\cdot\log^2\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M,m-M))}{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M))}\right)\right]\\ &\le \mathbb{E}\left[\log^4\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M,m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\right]^{1/2}\mathbb{E}\left[\log^4\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M,m-M))}{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M))}\right)\right]^{1/2}\\ &= \mathbb{E}\left[\log^4\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M,m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\right]. \end{aligned}\tag{11.16}$$

Furthermore, the random formula $\mathbf{\Phi}_1(M,m-M)$ is obtained from $\mathbf{\Phi}_1(M-1,m-M)$ by adding a single random clause $\boldsymbol{a}_M$, which is independent of $\mathbf{\Phi}_1(M-1,m-M)$. Therefore, Lemma 11.1 implies that

$$\log\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M,m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\le|\mathcal{N}(\mathbf{\Phi}_1(M-1,m-M),\boldsymbol{a}_M)|\log 2.\tag{11.17}$$

Moreover, since $|\mathcal{N}(\mathbf{\Phi}_1(M-1,m-M),\boldsymbol{a}_M)|$ has the same distribution as the random variable $\boldsymbol{N}'$ from Lemma 11.6, we obtain

$$\mathbb{E}\left[|\mathcal{N}(\mathbf{\Phi}_1(M-1,m-M),\boldsymbol{a}_M)|^4\right]=O(1)\tag{11.18}$$

uniformly for all M. Finally, the assertion follows from (11.16)–(11.18). □

To facilitate the following steps we introduce truncated versions of $\mathbf{\Delta}(M),\mathbf{\Delta}'(M),\mathbf{\Delta}''(M)$: for $B>0$ and $x>0$ let

$$\Lambda_B(x)=\begin{cases}-B & \text{if }\log(x)<-B,\\ B & \text{if }\log(x)>B,\\ \log(x) & \text{otherwise.}\end{cases}$$

Further, let

$$\begin{aligned}\mathbf{\Delta}_B(M)&=\Lambda_B\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M,m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\cdot\Lambda_B\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M,m-M))}{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M))}\right),\\ \mathbf{\Delta}'_B(M)&=\Lambda_B\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M+1))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\cdot\Lambda_B\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M+1))}{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M))}\right),\\ \mathbf{\Delta}''_B(M)&=\Lambda_B\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M,m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1,m-M))}\right)\cdot\Lambda_B\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M+1))}{Z(\hat{\mathbf{\Phi}}_2(M-1,m-M))}\right).\end{aligned}$$

Combining Lemma 11.10 with the Cauchy-Schwarz inequality, we obtain the following.

Corollary 11.11. *For any $\varepsilon>0$ there exists $B>0$ such that for all $1\le M\le m$ we have*

$$\mathbb{E}|\mathbf{\Delta}(M)-\mathbf{\Delta}_B(M)|+\mathbb{E}\left|\mathbf{\Delta}'(M)-\mathbf{\Delta}'_B(M)\right|+\mathbb{E}\left|\mathbf{\Delta}''(M)-\mathbf{\Delta}''_B(M)\right|<\varepsilon.$$

11.5. **The variance process.** In light of Lemma 11.9, to prove Proposition 2.12 we need to show that

$$\frac{1}{m}\sum_{M=1}^{m}\mathbb{E}\left[\mathbf{\Delta}(M)+\mathbf{\Delta}'(M)-2\mathbf{\Delta}''(M)\mid\mathfrak{F}_M\right]\to\eta(d)^2\qquad\text{in probability.}$$

To this end we divide the above sum up into batches $\bar{\Sigma}(L,L')=\Sigma(L,L')+\Sigma'(L,L')-2\Sigma''(L,L')$, where

$$\begin{aligned}\Sigma(L,L')&=\frac{1}{L'-L}\sum_{M=L}^{L'-1}\mathbb{E}[\mathbf{\Delta}(M)\mid\mathfrak{F}_M],\\ \Sigma'(L,L')&=\frac{1}{L'-L}\sum_{M=L}^{L'-1}\mathbb{E}\left[\mathbf{\Delta}'(M)\mid\mathfrak{F}_M\right],\\ \Sigma''(L,L')&=\frac{1}{L'-L}\sum_{M=L}^{L'-1}\mathbb{E}\left[\mathbf{\Delta}''(M)\mid\mathfrak{F}_M\right].\end{aligned}$$

Then for any sequence $1 = L_0 < \cdots < L_k = m$ we have

$$\frac{1}{n}\sum_{M=1}^{m}\mathbb{E}\left[\boldsymbol{\Delta}(M)+\boldsymbol{\Delta}'(M)-2\boldsymbol{\Delta}''(M)\mid\mathfrak{F}_M\right]=\sum_{i=1}^{k}\frac{L_i-L_{i-1}}{n}\bar{\Sigma}(L_{i-1},L_i).$$

The following lemma is the centerpiece of the proof.

Lemma 11.12. *For any $\varepsilon > 0$ there exists $\omega > 0$ such that uniformly for all $1 \le L < L' \le m$ with $\omega \le L' - L \le 2\omega$ we have*

$$\mathbb{E}\left|\Sigma(L,L')-\mathscr{B}^{\otimes}_{d,t}(\pi_{d,t})\right|+\mathbb{E}\left|\Sigma'(L,L')-\mathscr{B}^{\otimes}_{d,0}(\pi_{d,0})\right|+\mathbb{E}\left|\Sigma''(L,L')-\mathscr{B}^{\otimes}_{d,0}(\pi_{d,0})\right|<\varepsilon+o(1),\qquad \textit{where } t=L/m.$$

We will carry out the details for the first term $\mathbb{E}|\Sigma(L,L')-\mathscr{B}^{\otimes}_{d,t}(\pi_{d,t})|$, which is the most delicate; similar but slightly simpler steps yield the other two estimates. We begin by replacing $\boldsymbol{\Delta}(M)$ by its truncated version $\boldsymbol{\Delta}_B(M)$. Accordingly, let

$$\Sigma_B(L,L')=\frac{1}{L'-L}\sum_{M=L}^{L'-1}\mathbb{E}\left[\boldsymbol{\Delta}_B(M)\mid\mathfrak{F}_M\right],$$

$$\Sigma'_B(L,L')=\frac{1}{L'-L}\sum_{M=L}^{L'-1}\mathbb{E}\left[\boldsymbol{\Delta}'_B(M)\mid\mathfrak{F}_M\right],$$

$$\Sigma''_B(L,L')=\frac{1}{L'-L}\sum_{M=L}^{L'-1}\mathbb{E}\left[\boldsymbol{\Delta}''_B(M)\mid\mathfrak{F}_M\right].$$

Claim 11.13. *For any $\varepsilon > 0$ there exists $B_0 > 0$ such that for all $B > B_0$ and all $L, L' > 0$ we have*

$$\mathbb{E}\left|\Sigma(L,L')-\Sigma_B(L,L')\right|<\varepsilon+o(1).$$

Proof. This is an immediate consequence of Corollary 11.11. □

We proceed to relate the change in the pruned partition function to the marginal distribution of the truth values of the variables of the additional clause $\boldsymbol{a}_M$.

Claim 11.14. *Let $B > 0$. W.h.p. we have*

$$\Lambda_B\left(\frac{Z(\hat{\boldsymbol{\Phi}}_h(M,m-M))}{Z(\hat{\boldsymbol{\Phi}}_h(M-1,m-M))}\right)=\Lambda_B\left(1-\prod_{y\in\partial\boldsymbol{a}_M}\mu_{\hat{\boldsymbol{\Phi}}_h(M-1,m-M)}\left(\boldsymbol{\sigma}_y\neq\mathrm{sign}(y,\boldsymbol{a}_M)\right)\right)+o(1)\qquad(h=1,2).$$

Proof. Since the function Λ_B is bounded and continuous, this follows from Proposition 2.5. □

A combinatorial interpretation of $\Sigma(L,L')$ is that the sum gauges the cumulative effect of adding a total of $L'-L$ 'shared' clauses, one after the other. Claim 11.14 expresses the effect of adding a shared clause in terms of the marginals of the formula $\hat{\boldsymbol{\Phi}}_h(M-1,m-M)$. So long as the total number $L'-L$ of clauses added is not too large, we may expect that this marginal distribution does not shift all to much as we add clauses one by one. This is what the following claim verifies.

Claim 11.15. *Let $t = M/m$. If $L'-L = O(1)$, then w.h.p. we have*

$$\sum_{M=L}^{L'-1}W_1(\pi_{\hat{\boldsymbol{\Phi}}_1(M-1,m-M),\hat{\boldsymbol{\Phi}}_2(M-1,m-M)},\pi_{d,t})=o(1).$$

Proof. This follows from Corollary 2.10. □

As a next step we truncate the functional $\mathscr{B}^{\otimes}_{d,t}$ from (1.4). Hence, for $B > 0$ let

$$\begin{aligned}\mathscr{B}^{\otimes}_{B,d,t}(\pi)=\mathbb{E}\Big[&\Lambda_B\left(1-(\mathbb{1}\{\boldsymbol{r}_{-1,1}=-1\}+\boldsymbol{r}_{-1,1}\boldsymbol{\mu}_{\pi,-1,1,1})(\mathbb{1}\{\boldsymbol{r}_{-1,2}=-1\}+\boldsymbol{r}_{-1,2}\boldsymbol{\mu}_{\pi,-1,2,1})\right)\\&\Lambda_B\left(1-(\mathbb{1}\{\boldsymbol{r}_{-1,1}=-1\}+\boldsymbol{r}_{-1,1}\boldsymbol{\mu}_{\pi,-1,1,2})(\mathbb{1}\{\boldsymbol{r}_{-1,2}=-1\}+\boldsymbol{r}_{-1,2}\boldsymbol{\mu}_{\pi,-1,2,2})\right)\Big].\end{aligned}\qquad(11.19)$$

Claim 11.16. *For any $\varepsilon > 0$ there exists $B_0 > 0$ such that for all $B > B_0$ and all $t \in [0,1]$ we have*

$$\left|\mathscr{B}^{\otimes}_{d,t}(\pi_{d,t})-\mathscr{B}^{\otimes}_{B,d,t}(\pi_{d,t})\right|<\varepsilon.$$

Proof. Since $\mathscr{B}^{\otimes}_{B,d,t}(\pi_{d,t})\uparrow\mathscr{B}^{\otimes}_{d,t}(\pi_{d,t})$ as $B\to\infty$, this follows from Proposition 2.8. □

Proof of Lemma 11.12. Lemma 10.1 ensures that $\mathscr{B}^{\otimes}_{d,t}(\pi_{d,t}) < \infty$ for all t. Moreover, Claims 11.13 and 11.16 imply that we just need to show that for large $B > 0$,

$$\mathbb{E}\left|\Sigma_B(L, L') - \mathscr{B}^{\otimes}_{B,d,t}(\pi_{d,t})\right| < \varepsilon. \tag{11.20}$$

Let $t = L/m$ and let $(\boldsymbol{S}_M)_{L \le M < L'}$ be independent copies of the random variable

$$\prod_{h=1}^{2} \Lambda_B \left(1 - (\mathbb{1}\{\boldsymbol{r}_{-1,1} = -1\} + \boldsymbol{r}_{-1,1}\boldsymbol{\mu}_{\pi_{d,t},-1,1,h})(\mathbb{1}\{\boldsymbol{r}_{-1,2} = -1\} + \boldsymbol{r}_{-1,2}\boldsymbol{\mu}_{\pi_{d,t},-1,2,h})\right).$$

Furthermore, let

$$\boldsymbol{D}_M = \prod_{h=1}^{2} \Lambda_B \left(\frac{Z(\hat{\boldsymbol{\Phi}}_h(M, m-M))}{Z(\hat{\boldsymbol{\Phi}}_h(M-1, m-M))}\right)$$

Then Claims 11.14– 11.15 show that

$$\mathbb{E}\left[W_1\left(\sum_{M=L}^{L-1} \boldsymbol{D}_M, \sum_{M=L}^{L-1} \boldsymbol{S}_M\right)\right] = o(1). \tag{11.21}$$

Finally, since $\sum_{M=L}^{L-1} \boldsymbol{S}_M$ is a sum of bounded independent random variables, (11.20) follows from (11.21) and the strong law of large numbers. □

Proof of Proposition 2.12. The second equality in (2.14) follows from Corollary 11.11, Lemma 11.12 and the triangle inequality. Thus, we are left to verify the first condition. Since Lemma 11.9 shows that

$$\boldsymbol{X}_M^2 = \frac{1}{m}\mathbb{E}\left[\boldsymbol{\Delta}(M) + \boldsymbol{\Delta}(M)' - 2\boldsymbol{\Delta}''(M) \mid \mathfrak{F}_M\right],$$

it suffices to prove that

$$\mathbb{E}\max_{1 \le M \le m} \boldsymbol{\Delta}(M)^2 + \mathbb{E}\max_{1 \le M \le m} \boldsymbol{\Delta}'(M)^2 + \mathbb{E}\max_{1 \le M \le m} \boldsymbol{\Delta}''(M)^2 = o(m). \tag{11.22}$$

We will bound the first expectation; similar arguments apply to the others.

Retracing the steps of the proof of Lemma 11.10, we write

$$\boldsymbol{\Delta}(M)^2 = \log^2\left(\frac{Z(\hat{\boldsymbol{\Phi}}_1(M, m-M))}{Z(\hat{\boldsymbol{\Phi}}_1(M-1, m-M))}\right) \cdot \log^2\left(\frac{Z(\hat{\boldsymbol{\Phi}}_2(M, m-M))}{Z(\hat{\boldsymbol{\Phi}}_2(M-1, m-M))}\right).$$

Since $\boldsymbol{\Phi}_1(M, m-M)$ is obtained from $\boldsymbol{\Phi}_1(M-1, m-M)$ by adding the random clause $\boldsymbol{a}_M$, Lemma 11.1 implies that

$$\log\left(\frac{Z(\hat{\boldsymbol{\Phi}}_h(M, m-M))}{Z(\hat{\boldsymbol{\Phi}}_h(M-1, m-M))}\right) \le |\mathscr{N}(\boldsymbol{\Phi}_h(M-1, m-M), \boldsymbol{a}_M)| \log 2 \qquad (h = 1, 2). \tag{11.23}$$

As $|\mathscr{N}(\boldsymbol{\Phi}_h(M-1, m-M), \boldsymbol{a}_M)|$ has the same distribution as $\boldsymbol{N}'$ from Lemma 11.6, we obtain

$$\mathbb{E}\left[|\mathscr{N}(\boldsymbol{\Phi}_1(M-1, m-M), \boldsymbol{a}_M)|^4\right] = O(1)$$

uniformly for all M. Therefore, Markov's inequality implies

$$\mathbb{P}\left[|\mathscr{N}(\boldsymbol{\Phi}_1(M-1, m-M), \boldsymbol{a}_M)| > m^{1/3}\right] \le O(m^{-4/3}). \tag{11.24}$$

Finally, (11.22) follows from (11.23), (11.24) and the union bound. □

As a final preparation towards the proof of Proposition 2.11 we need a lower bound on $\log Z(\hat{\boldsymbol{\Phi}})$.

Lemma 11.17. *We have* $\mathrm{Var}(\log Z(\hat{\boldsymbol{\Phi}})) = \Omega(n)$.

Proof. Let $\mathfrak{C}$ be the set of isolated sub-formulas of $\hat{\boldsymbol{\Phi}}$ with precisely three clauses and three variables that are acyclic and whose unique variable of degree two appears with the same sign in both its adjacent clauses. Moreover, let $\boldsymbol{C}'$ be the set of isolated sub-formulas of $\hat{\boldsymbol{\Phi}}$ with precisely three clauses and three variables that are acyclic such that the unique variable of degree two appears with two different signs in its adjacent clauses. Then $\mathbb{E}|\mathfrak{C}| = \mathbb{E}|\mathfrak{C}'| = \Omega(n)$ and w.h.p. we have

$$\mathrm{Var}(|\mathfrak{C}| \mid |\mathfrak{C}| + |\mathfrak{C}'|) = \mathrm{Var}(|\mathfrak{C}'| \mid |\mathfrak{C}| + |\mathfrak{C}'|) = \Omega(n). \tag{11.25}$$

Additionally, for each sub-formula $\mathscr{C} \in \mathfrak{C}$ we have $Z(\mathscr{C}) = 5$, while for $\mathscr{C}' \in \mathfrak{C}'$ we have $Z(\boldsymbol{C}') = 4$. Since with the sum ranging over the connected components $\mathscr{C}$ of $\hat{\boldsymbol{\Phi}}$ we have $\log Z(\hat{\boldsymbol{\Phi}}) = \sum_{\mathscr{C}} \log Z(\mathscr{C})$, the assertion follows from (11.25). □

Proof of Corollary 2.11. The corollary is an immediate consequence of Lemma 2.4, Lemma 10.1, Corollary 11.11, Lemma 11.12 and Lemma 11.17. □

12. Proof of Theorem 1.1

We derive Theorem 1.1 from the following general martingale central limit theorem, which is a special case of [33, Theorem 3.2] (see also the subsequent remark there).

Theorem 12.1 ([33, Theorem 3.2])**.** *Let $(\boldsymbol{Z}_{n,i},\mathfrak{F}_{n,i})_{0\le i\le m_n,n\ge1}$ be a zero-mean, square-integrable martingale array with differences $\boldsymbol{X}_{n,i}=\boldsymbol{Z}_{n,i}-\boldsymbol{Z}_{n,i-1}$ for $1\le i\le m_n$. Assume that there exists a constant η^2 such that*

$$\lim_{n\to\infty}\max_{1\le i\le m_n}|\boldsymbol{X}_{n,i}|=0 \qquad \text{in probability,} \tag{12.1}$$

$$\lim_{n\to\infty}\sum_{i=1}^{m_n}\boldsymbol{X}_{n,i}^2=\eta^2 \qquad \text{in probability,} \tag{12.2}$$

$$\mathbb{E}\left[\max_{1\le i\le m_n}\boldsymbol{X}_{n,i}^2\right] \qquad \text{is bounded in } n. \tag{12.3}$$

Then $\boldsymbol{Z}_{n,m_n}$ converges in distribution to a Gaussian distribution with mean zero and variance η^2.

Proof of Theorem 1.1. We apply Theorem 12.1 to the filtration $(\mathfrak{F}_{n,M})_{0\le M\le m_n}$ from Section 2.8 and to the Doob martingale $(\boldsymbol{Z}_{n,M}-\mathbb{E}\left[\boldsymbol{Z}_{n,M}\right])_M$ from (2.13). This is zero-mean by construction and square-integrable, as $\log Z(\hat{\boldsymbol{\Phi}})$ is non-negative and bounded above by n. Let $\boldsymbol{X}_{n,M}=\boldsymbol{Z}_{n,M}-\boldsymbol{Z}_{n,M-1}$ be the martingale differences. Proposition 2.12 immediately implies conditions (12.1)–(12.2) of Theorem 12.1 since L^1-convergence implies convergence in probability. Condition (12.3) also follows from Proposition 2.12 by observing that

$$\mathbb{E}\left[\max_{1\le M\le m_n}\boldsymbol{X}_{n,M}^2\right]\le\mathbb{E}\left[\sum_{M=1}^{m_n}\boldsymbol{X}_{n,M}^2\right]\le\mathbb{E}\left|\sum_{M=1}^{m_n}\boldsymbol{X}_{n,M}^2-\eta(d)^2\right|+\eta(d)^2.$$

Furthermore, Lemma 10.1 guarantees that $\eta(d)<\infty$, while Corollary 2.11 shows that $\eta(d)>0$. Thus, the assertion follows from Theorem 12.1. □

Acknowledgement. Amin Coja-Oghlan's research is supported by DFG CO 646/3, DFG CO 646/5 and DFG CO 646/6. Pavel Zakharov's research is supported by DFG CO 646/6. Haodong Zhu's research is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 945045, and by the NWO Gravitation project NETWORKS under grant no. 024.002.003. Noela Müller's research is supported by the NWO Gravitation project NETWORKS under grant no. 024.002.003. We thank Nicloa Kistler for helpful discussions, and particularly for bringing [19] to our attention.

References

[1] E. Abbe, S. Li, A. Sly: Proof of the contiguity conjecture and lognormal limit for the symmetric perceptron. Proc. 62nd FOCS (2022) 327–338.
[2] E. Abbe, A. Montanari: On the concentration of the number of solutions of random satisfiability formulas. Random Structures and Algorithms **45** (2014) 362–382.
[3] D. Achlioptas, A. Chtcherba, G. Istrate, C. Moore: The phase transition in 1-in-k SAT and NAE 3-SAT. Proc. 12th SODA (2001) 721–722.
[4] D. Achlioptas, A. Coja-Oghlan: Algorithmic barriers from phase transitions. Proc. 49th FOCS (2008) 793–802.
[5] D. Achlioptas, A. Coja-Oghlan, M. Hahn-Klimroth, J. Lee, N. Müller, M. Penschuck, G. Zhou: The number of satisfying assignments of random 2-SAT formulas. Random Structures and Algorithms **58** (2021) 609–647.
[6] D. Achlioptas, C. Moore: Random k-SAT: two moments suffice to cross a sharp threshold. SIAM Journal on Computing **36** (2006) 740–762.
[7] D. Achlioptas, A. Naor, Y. Peres: Rigorous location of phase transitions in hard optimization problems. Nature **435** (2005) 759–764.
[8] D. Achlioptas, Y. Peres: The threshold for random k-SAT is $2^k\ln 2-O(k)$. Journal of the AMS **17** (2004) 947–973.
[9] D. Aldous, J. Steele: The objective method: probabilistic combinatorial optimization and local weak convergence. In: H. Kesten (ed.): Probability on Discrete Structures. Springer (2004).
[10] B. Aspvall, M.F. Plass, R.E. Tarjan: A linear-time algorithm for testing the truth of certain quantified boolean formulas. Information Processing Letters **8** (1979) 121–123.
[11] P. Ayre, A. Coja-Oghlan, P. Gao, N. Müller: The satisfiability threshold for random linear equations. Combinatorica **40** (2020) 179–235
[12] V. Bapst, A. Coja-Oghlan, C. Efthymiou: Planting colourings silently. Combintorics, Probability and Computing **26** (2017) 338-366.
[13] V Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann, D. Vilenchik: The condensation phase transition in random graph coloring. Communications in Mathematical Physics **341** (2016) 543–606.
[14] P. J. Bickel, P. A. Freedman. Some asymptotic theory for the bootstrap. Annals of Statistics **9** (1981) 1196–1217.
[15] B. Bollobás, C. Borgs J. Chayes, J. Kim, D. Wilson: The scaling window of the 2-SAT transition. Random Structures and Algorithms **18** (2001) 201–256.

[16] S. Cao: Central limit theorems for combinatorial optimization problems on sparse Erdős-Rényi graphs. Annals of Applied Probability **31** (2021) 1687–1723.
[17] P. Cheeseman, B. Kanefsky, W. Taylor: Where the *really* hard problems are. Proc. IJCAI (1991) 331–337.
[18] G. Bresler, B. Huang: The algorithmic phase transition of random k-SAT for low degree polynomials. Proc. 62nd FOCS (2021) 298–309.
[19] W.-K. Chen, P. Dey, D. Panchenko: Fluctuations of the free energy in the mixed p-spin models with external field. Probability Theory and Related Fields **168** (2017) 41–53.
[20] V. Chvátal, B. Reed: Mick gets some (the odds are on his side). Proc. 33th FOCS (1992) 620–627.
[21] A. Coja-Oghlan, T. Kapetanopoulos, N. Müller: The replica symmetric phase of random constraint satisfaction problems. Combinatorics, Probability and Computing **29** (2020) 346-422.
[22] A. Coja-Oghlan, F. Krzakala, W. Perkins, L. Zdeborová: Information-theoretic thresholds from the cavity method. Advances in Mathematics **333** (2018) 694–795.
[23] A. Coja-Oghlan, K. Panagiotou: The asymptotic k-SAT threshold. Advances in Mathematics **288** (2016) 985–1068.
[24] A. Coja-Oghlan, N. Wormald: The number of satisfying assignments of random regular k-SAT formulas. Combinatorics, Probability and Computing **27** (2018) 496–530.
[25] J. Ding, A. Sly, N. Sun: Proof of the satisfiability conjecture for large k. **20** Annals of Mathematics **196** (2022) 1–388.
[26] S. Dovgal, É. de Panafieu, V. Ravelomanana: Exact enumeration of satisfiable 2-SAT formulae. arXiv:2108.08067 (2021).
[27] O. Dubois, J. Mandler: The 3-XORSAT threshold. Proc. 43rd FOCS (2002) 769–778.
[28] G. Eagleson: Martingale convergence to mixtures of infinitely divisible laws. Annals of Probability **3** (1975) 557–562.
[29] C. Efthymiou: On sampling symmetric gibbs distributions on sparse random graphs and hypergraphs. Proc. 49th ICALP (2022) #57.
[30] E. Friedgut: Sharp thresholds of graph properties, and the k-SAT problem. Journal of the AMS **12** (1999) 1017–1054.
[31] M. Glasgow, M. Kwan, A. Sah, M. Sawhney: A central limit theorem for the matching number of a sparse random graph. arXiv:2402.05851 (2024).
[32] A. Goerdt: A threshold for unsatisfiability. J. Comput. Syst. Sci. **53** (1996) 469–486
[33] P. Hall, C. Heyde: Martingale limit theory and its applications. Academic Press (1980).
[34] E. Kreačič: Some problems related to the Karp-Sipser algorithm on random graphs. Ph.D. thesis, University of Oxford, 2017.
[35] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborová: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. National Academy of Sciences **104** (2007) 10318–10323.
[36] M. Mézard, A. Montanari: Information, physics and computation. Oxford University Press (2009).
[37] M. Mézard, G. Parisi, R. Zecchina: Analytic and algorithmic solution of random satisfiability problems. Science **297** (2002) 812–815.
[38] R. Monasson, R. Zecchina: The entropy of the k-satisfiability problem. Phys. Rev. Lett. **76** (1996) 3881.
[39] A. Montanari, D. Shah: Counting good truth assignments of random k-SAT formulae. Proc. 18th SODA (2007) 1255–1264.
[40] E. Mossel, J. Neeman, A Sly: Reconstruction and estimation in the planted partition model. Probability Theory and Related Fields (2014) 1–31.
[41] D. Panchenko: On the replica symmetric solution of the K-sat model. Electron. J. Probab. **19** (2014) #67.
[42] D. Panchenko, M. Talagrand: Bounds for diluted mean-fields spin glass models. Probab. Theory Relat. Fields **130** (2004) 319–336.
[43] B. Pittel, G. Sorkin: The satisfiability threshold for k-XORSAT. Combinatorics, Probability and Computing **25** (2016) 236–268.
[44] F. Rassmann: On the number of solutions in random graph k-colouring. Combinatorics, Probability and Computing **28** (2019) 130–158.
[45] R. Robinson, N. Wormald: Almost all regular graphs are Hamiltonian. Random Structures and Algorithms **5** (1994) 363–374.
[46] A. Sly, N. Sun, Y. Zhang: The number of solutions for random regular NAE-SAT. Probability Theory and Related Fields **182** (2022) 1–109.
[47] M. Talagrand: The high temperature case for the random K-sat problem. Probab. Theory Related Fields **119** (2001) 187–212.
[48] L. Valiant: The complexity of enumeration and reliability problems. SIAM Journal on Computing **8** (1979) 410–421.

Arnab Chatterjee, `arnab.chatterjee@tu-dortmund.de`, TU Dortmund, Faculty of Computer Science, 12 Otto-Hahn-St, Dortmund 44227, Germany.

Amin Coja-Oghlan, `amin.coja-oghlan@tu-dortmund.de`, TU Dortmund, Faculty of Computer Science and Faculty of Mathematics, 12 Otto-Hahn-St, Dortmund 44227, Germany.

Noela Müller, `n.s.muller@tue.nl`, Eindhoven University of Technology, Department of Mathematics and Computer Science, MetaForum MF 4.084, 5600 MB Eindhoven, the Netherlands.

Connor Riddlesden, `c.d.riddlesden@tue.nl`, Eindhoven University of Technology, Department of Mathematics and Computer Science, MetaForum MF 4.084, 5600 MB Eindhoven, the Netherlands.

Maurice Rolvien, `maurice.rolvien@tu-dortmund.de`, TU Dortmund, Faculty of Computer Science, 12 Otto-Hahn-St, Dortmund 44227, Germany.

Pavel Zakharov, `pavel.zakharov@tu-dortmund.de`, TU Dortmund, Faculty of Computer Science and Faculty of Mathematics, 12 Otto-Hahn-St, Dortmund 44227, Germany.

Haodong Zhu, `h.zhu1@tue.nl`, Eindhoven University of Technology, Department of Mathematics and Computer Science, 5600 MB Eindhoven, the Netherlands.

BELIEF PROPAGATION GUIDED DECIMATION ON RANDOM k-XORSAT

ARNAB CHATTERJEE, AMIN COJA-OGHLAN, MIHYUN KANG, LENA KRIEG, MAURICE ROLVIEN, GREGORY B. SORKIN

ABSTRACT. We analyse the performance of *Belief Propagation Guided Decimation*, a physics-inspired message passing algorithm, on the random k-XORSAT problem. Specifically, we derive an explicit threshold up to which the algorithm succeeds with a strictly positive probability $\Omega(1)$ that we compute explicitly, but beyond which the algorithm with high probability fails to find a satisfying assignment. In addition, we analyse a thought experiment called the *decimation process* for which we identify a (non-)reconstruction and a condensation phase transition. The main results of the present work confirm physics predictions from [Ricci-Tersenghi and Semerjian: J. Stat. Mech. 2009] that link the phase transitions of the decimation process with the performance of the algorithm, and improve over partial results from a recent article [Yung: Proc. ICALP 2024]. MSc: 60B20, 68W20

1. INTRODUCTION AND RESULTS

1.1. **Background and motivation.** The random k-XORSAT problem shares many characteristics of other intensely studied random constraint satisfaction problems ('CSPs') such as random k-SAT. For instance, random k-XORSAT possesses a sharp satisfiability threshold preceded by a reconstruction or 'shattering' phase transition that affects the geometry of the set of solutions [2, 12, 17, 24]. As in random k-SAT, these transitions appear to significantly impact the performance of certain classes of algorithms [7, 16]. At the same time, random k-XORSAT is more amenable to mathematical analysis than, say, random k-SAT. This is because the XOR operation is equivalent to addition modulo two, which is why a k-XORSAT instance translates into a linear system over $\mathbb{F}_2$. In effect, k-XORSAT can be solved in polynomial time by means of Gaussian elimination. In addition, the algebraic nature of the problem induces strong symmetry properties that simplify its study [3].

Because of its similarities with other random CSPs combined with said relative amenability, random k-XORSAT provides an instructive benchmark. This was noticed not only in combinatorics, but also in the statistical physics community, which has been contributing intriguing 'predictions' on random CSPs since the early 2000s [19, 22]. Among other things, physicists have proposed a message passing algorithm called *Belief Propagation Guided Decimation* ('BPGD') that, according to computer experiments, performs impressively on various random CSPs [21]. Furthermore, Ricci-Tersenghi and Semerjian [25] put forward a heuristic analysis of BPGD on random k-SAT and k-XORSAT. Their heuristic analysis proceeds by way of a thought experiment based on an idealized version of the algorithm. We call this thought experiment the *decimation process*. Based on physics methods Ricci-Tersenghi and Semerjian surmise that the decimation process undergoes two phase transitions, specifically a reconstruction and a condensation transition. A key prediction of Ricci-Tersenghi and Semerjian is that these phase transitions are directly linked to the performance of the BPGD algorithm. Due to the linear algebra-induced symmetry properties, in the case of random k-XORSAT all of these conjectures come as elegant analytical expressions.

The aim of this paper is to verify the predictions from [25] on random k-XORSAT mathematically. Specifically, our aim is to rigorously analyse the BPGD algorithm on random k-XORSAT, and to establish the link between its performance and the phase transitions of the decimation process. A first step towards a rigorous analysis of BPGD on random k-XORSAT was undertaken in a recent contribution by Yung [27]. However, Yung's analysis turns out to be not tight. Specifically, apart from requiring spurious lower bounds on the clause length k, Yung's results do not quite establish the precise connection between the decimation process and the performance of BPGD. One reason for this is that [27] relies on 'annealed' techniques, i.e., essentially moment computations. Here we instead harness 'quenched' arguments that were partly developed in prior work on the rank of random matrices over finite fields [3, 8].

Throughout we let $k \geq 3$ and $n \geq k$ be integers and $d > 0$ a positive real. Let $\boldsymbol{m} \stackrel{\text{dist}}{=} \text{Po}(dn/k)$ and let $\boldsymbol{F} = \boldsymbol{F}(n, d, k)$ be a random k-XORSAT formula with variables $x_1, \ldots, x_n$ and $\boldsymbol{m}$ random clauses of length k. To be precise, every

Amin Coja-Oghlan is supported by DFG CO 646/3, DFG CO 646/5 and DFG CO 646/6.
This research was funded in part by the Austrian Science Fund (FWF) 10.55776/I6502.

clause of $\boldsymbol{F}$ is an XOR of precisely k distinct variables, each of which may or may not come with a negation sign. The $\boldsymbol{m}$ clauses are drawn uniformly and independently out of the set of all $2^k\binom{n}{k}$ possibilities. Thus, d equals the average number of clauses that a given variable x_i appears in. An event $\mathscr{E}$ occurs *with high probability* ('w.h.p.') if $\lim_{n\to\infty}\mathbb{P}[\boldsymbol{F}\in\mathscr{E}]=1$. We always keep d,k fixed as $n\to\infty$.

1.2. **Belief Propagation Guided Decimation.** The first result vindicates the predictions from [25] concerning the success probability of BPGD algorithm. BPGD sets its ambitions higher than merely finding a solution to the k-XORSAT instance $\boldsymbol{F}$: the algorithm attempts to sample a solution uniformly at random. To this end BPGD assigns values to the variables $x_1,\dots,x_n$ of $\boldsymbol{F}$ one after the other. In order to assign the next variable the algorithm attempts to compute the marginal probability that the variable is set to 'true' under a random solution to the k-XORSAT instance, given all previous assignments. More precisely, suppose BPGD has assigned values to the variables $x_1,\dots,x_t$ already. Write $\boldsymbol{\sigma}_{\mathrm{BP}}(x_1),\dots,\boldsymbol{\sigma}_{\mathrm{BP}}(x_t)\in\{0,1\}$ for their values, with 1 representing 'true' and 0 'false'. Further, let $\boldsymbol{F}_{\mathrm{BP},t}$ be the simplified formula obtained by substituting $\boldsymbol{\sigma}_{\mathrm{BP}}(x_1),\dots,\boldsymbol{\sigma}_{\mathrm{BP}}(x_t)$ for $x_1,\dots,x_t$. We drop any clauses from $\boldsymbol{F}_{\mathrm{BP},t}$ that contain variables from $\{x_1,\dots,x_t\}$ only, deeming any such clauses satisfied. Thus, $\boldsymbol{F}_{\mathrm{BP},t}$ is a XORSAT formula with variables $x_{t+1},\dots,x_n$. Its clauses contain at least one and at most k variables, as well as possibly a constant (the XOR of the values substituted in for $x_1,\dots,x_t$).

Let $\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}$ be a uniformly random solution of the XORSAT formula $\boldsymbol{F}_{\mathrm{BP},t}$, assuming that $\boldsymbol{F}_{\mathrm{BP},t}$ remains satisfiable. Then BPGD aims to compute the marginal probability $\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}(x_{t+1})=1\mid\boldsymbol{F}_{\mathrm{BP},t}\right]$ that a random satisfying assignment of $\boldsymbol{F}_{\mathrm{BP},t}$ sets x_{t+1} to true. This is where Belief Propagation ('BP') comes in. An efficient message passing heuristic for computing precisely such marginals, BP returns an 'approximation' $\mu_{\boldsymbol{F}_{\mathrm{BP},t}}$ of $\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}(x_{t+1})=1\mid\boldsymbol{F}_{\mathrm{BP},t}\right]$. We will recap the mechanics of BP in Section 2.2 (the value $\mu_{\boldsymbol{F}_{\mathrm{BP},t}}$ is defined precisely in (2.11)). Having computed the BP 'approximation', BPGD proceeds to assign x_{t+1} the value 'true' with probability $\mu_{\boldsymbol{F}_{\mathrm{BP},t}}$, otherwise sets x_{t+1} to 'false', then moves on to the next variable. The pseudocode is displayed as Algorithm 1.

Data: a random k-XORSAT formula $\boldsymbol{F}$ with variables $x_1,\dots,x_n$ conditioned on being satisfiable
for $t=0,\dots,n-1$ **do**
    compute the BP approximation $\mu_{\boldsymbol{F}_{\mathrm{BP},t}}$;
    set $\boldsymbol{\sigma}_{\mathrm{BP}}(x_{t+1})=\begin{cases}1 & \text{with probability } \mu_{\boldsymbol{F}_{\mathrm{BP},t}}\\ 0 & \text{with probability } 1-\mu_{\boldsymbol{F}_{\mathrm{BP},t}}\end{cases}$;
return $\boldsymbol{\sigma}_{\mathrm{BP}}$;

Algorithm 1: The BPGD algorithm.

Let us pause for a few remarks. First, if the BP approximations are exact, i.e., if $\boldsymbol{F}_{\mathrm{BP},t}$ is satisfiable and $\mu_{\boldsymbol{F}_{\mathrm{BP},t}}=\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}(x_{t+1})=1\mid\boldsymbol{F}_{\mathrm{BP},t}\right]$ for all t, then Bayes' formula shows that BPGD outputs a uniformly random solution of $\boldsymbol{F}$. However, there is no universal guarantee that BP returns the correct marginals. Accordingly, the crux of analysing BPGD is precisely to figure out whether this is the case. Indeed, the heuristic work of [25] ties the accuracy of BP to a phase transition of the decimation process thought experiment, to be reviewed momentarily.

Second, the strategy behind the BPGD algorithm, particularly the message passing heuristic for 'approximating' the marginals, generalizes well beyond k-XORSAT. For instance, the approach applies to k-SAT verbatim. That said, due to the algebraic nature of the XOR operation, BPGD is *far* easier to analyse on k-XORSAT. In fact, in XORSAT the marginal probabilities are guaranteed to be half-integral as seen in Fact 2.3, i.e.,

$$\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{BP},t}}(x_{t+1})=1\mid\boldsymbol{F}_{\mathrm{BP},t}\right]\in\{0,1/2,1\}. \tag{1.1}$$

As a consequence, on XORSAT the BPGD algorithm effectively reduces to a purely combinatorial algorithm called Unit Clause Propagation [19, 25] as per Proposition 6.1, a fact that we will exploit extensively (see Section 6).

1.3. **A tight analysis of** BPGD**.** In order to state the main results we need to introduce a few threshold values. To this end, given d, k and an additional real parameter $\lambda \geq 0$, consider the functions [1]

$$\phi_{d,k,\lambda} : [0,1] \to [0,1], \qquad z \mapsto 1 - \exp\left(-\lambda - d z^{k-1}\right), \tag{1.2}$$

$$\Phi_{d,k,\lambda} : [0,1] \to \mathbb{R}, \qquad z \mapsto \exp\left(-\lambda - d z^{k-1}\right) - \frac{d(k-1)}{k} z^k + d z^{k-1} - \frac{d}{k}. \tag{1.3}$$

Let $\alpha_*(\lambda) = \alpha_*(d,k,\lambda) \in [0,1]$ be the smallest and $\alpha^*(\lambda) = \alpha^*(d,k,\lambda) \geq \alpha_*(d,k,\lambda) \in [0,1]$ the largest fixed point of $\phi_{d,k,\lambda}$. Figure 1 visualizes $\Phi(z)$ for different values of θ. Further, define

$$d_{\min}(k) = \left(\frac{k-1}{k-2}\right)^{k-2}, \quad d_{\mathrm{core}}(k) = \sup\left\{d > 0 : \alpha^*(0) = 0\right\}, \quad d_{\mathrm{sat}}(k) = \sup\left\{d > 0 : \Phi_{d,k,0}(\alpha^*(0)) \leq \Phi_{d,k,0}(0)\right\}. \tag{1.4}$$

The value $d_{\mathrm{sat}}(k)$ is the random k-XORSAT satisfiability threshold [3, 12, 24]. Thus, for $d < d_{\mathrm{sat}}(k)$ the random k-XORSAT formula $\boldsymbol{F}$ possesses satisfying assignments w.h.p., while $\boldsymbol{F}$ is unsatisfiable for $d > d_{\mathrm{sat}}(k)$ w.h.p. Furthermore, $d_{\mathrm{core}}(k)$ equals the threshold for the emergence of a giant 2-core within the k-uniform hypergraph induced by $\boldsymbol{F}$ [3, 23]. This implies that for $d < d_{\mathrm{core}}(k)$ the set of solutions of $\boldsymbol{F}$ is contiguous in a certain well-defined way, while for $d_{\mathrm{core}}(k) < d < d_{\mathrm{sat}}(k)$ the set of solutions shatters into an exponential number of well-separated clusters [16, 19]. Moreover, a simple linear time algorithm is known to find a solution w.h.p. for $d < d_{\mathrm{core}}(k)$ [16]. The relevance of $d_{\min}(k)$ will emerge momentarily. A bit of calculus reveals that

$$0 < d_{\min}(k) < d_{\mathrm{core}}(k) < d_{\mathrm{sat}}(k) < k. \tag{1.5}$$

The following theorem determines the precise clause-to-variable densities where BPGD succeeds/fails. To be precise, in the 'successful' regime BPGD does not actually succeed with *high* probability, but with an explicit probability strictly between zero and one, which is displayed in Figure 2 for $k = 3, 4, 5$.

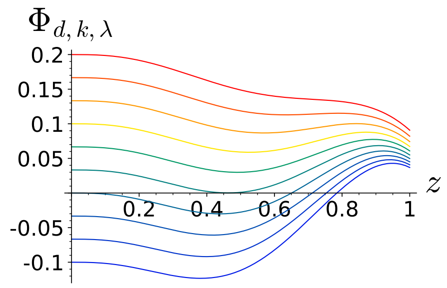


FIGURE 1. $\Phi_{d,k,\lambda}$ for $k = 3$ and $d = 2.4$, for λ from 0 to 0.3 (maximum at $z = 0$) and from 0.4 to 0.9

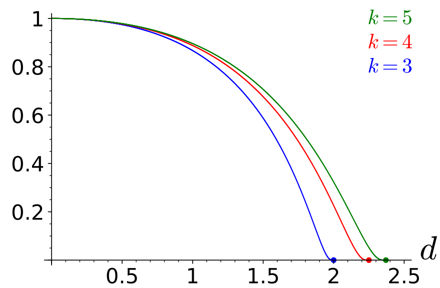


FIGURE 2. Success probability of BPGD for $0 < d < d_{\min}(k)$ and various k.

Theorem 1.1. *Let $k \geq 3$.*

(i) If $d < d_{\min}(k)$, then

$$\lim_{n\to\infty} \mathbb{P}\left[\mathtt{BPGD}(\boldsymbol{F}) \textit{ finds a satisfying assignment}\right] = \exp\left(-\frac{d^2(k-1)^2}{4}\int_0^1 \frac{z^{2k-4}(1-z)}{1-d(k-1)z^{k-2}(1-z)}\,\mathrm{d}z\right). \tag{1.6}$$

(ii) If $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$, then

$$\mathbb{P}\left[\mathtt{BPGD}(\boldsymbol{F}) \textit{ finds a satisfying assignment}\right] = o(1).$$

Theorem 1.1 vindicates the predictions from Ricci-Tersenghi and Semerjian [25, Section 4] as to the performance of BPGD, and improves over the results from Yung [27]. Specifically, Theorem 1.1 (i) verifies the formula for the success probability from [25, Eq. (38)]. Combinatorially, the formula (1.6) results from the possible presence of bounded length cycles (so called toxic cycles) that may cause the algorithm to run into contradictions. By contrast, Yung has no positive result on the performance of BPGD. Moreover, Yung's negative results [27, Theorems 2–3]

[1] The function $\Phi_{d,k,\lambda}$ is known in physics parlance as the "Bethe free entropy" [8, 19]. The stationary points of $\Phi_{d,k,\lambda}$ coincide with the fixed points of $\phi_{d,k,\lambda}$, as we will verify in Section 2.1.

only apply to $k \geq 9$ and to $d > d_{\mathrm{core}}(k)$, while Theorem 1.1 (ii) covers all $k \geq 3$ and kicks in at the correct threshold $d_{\min}(k) < d_{\mathrm{core}}(k)$ predicted in [25].

1.4. **The decimation process.** In addition to the BPGD algorithm itself, the heuristic work [25] considers an idealised version of the algorithm, the *decimation process*. This thought experiment highlights the conceptual reasons behind the success/failure of BPGD. Just like BPGD, the decimation process assigns values to variables one after the other for good. But instead of the BP 'approximations' the decimation process uses the *actual* marginals given its previous decisions. To be precise, suppose that the input formula $\boldsymbol{F}$ is satisfiable and that variables $x_1, \ldots, x_t$ have already been assigned values $\boldsymbol{\sigma}_{\mathrm{DC}}(x_1), \ldots, \boldsymbol{\sigma}_{\mathrm{DC}}(x_t)$ in the previous iterations. Obtain $\boldsymbol{F}_{\mathrm{DC},t}$ by substituting the values $\boldsymbol{\sigma}_{\mathrm{DC}}(x_1), \ldots, \boldsymbol{\sigma}_{\mathrm{DC}}(x_t)$ for $x_1, \ldots, x_t$ and dropping any clauses that do not contain any of $x_{t+1}, \ldots, x_n$. Thus, $\boldsymbol{F}_{\mathrm{DC},t}$ is a XORSAT formula with variables $x_{t+1}, \ldots, x_n$. Let $\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}$ be a random satisfying assignment of $\boldsymbol{F}_{\mathrm{DC},t}$. Then the decimation process sets x_{t+1} according to the true marginal $\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]$, thus ultimately returning a uniformly random satisfying assignment of $\boldsymbol{F}$.

Data: a random k-XORSAT formula $\boldsymbol{F}$, conditioned on being satisfiable
for $t = 0, \ldots, n-1$ **do**
  compute $\pi_{\boldsymbol{F}_{\mathrm{DC},t}} = \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]$;
  set $\boldsymbol{\sigma}_{\mathrm{DC}}(x_t) = \begin{cases} 1 & \text{with probability } \pi_{\boldsymbol{F}_{\mathrm{DC},t}} \\ 0 & \text{with probability } 1 - \pi_{\boldsymbol{F}_{\mathrm{DC},t}} \end{cases}$;
return $\boldsymbol{\sigma}_{\mathrm{DC}}$;

Algorithm 2: The decimation process.

Clearly, if indeed the BP 'approximations' are correct, then the decimation process and BPGD are identical. Thus, a key question is for what parameter regimes the two process coincide or diverge, respectively. As it turns out, this question is best answered by parametrize not only in terms of the average variable degree d, but also in terms of the 'time' parameter t of the decimation process.

1.5. **Phase transitions of the decimation process.** Ricci-Tersenghi and Semerjian heuristically identify several phase transitions in terms of d and t that the decimation process undergoes. We will confirm these predictions mathematically and investigate how they relate to the performance of BPGD.

The first set of relevant phase transitions concerns the so-called non-reconstruction property. Roughly speaking, non-reconstruction means that the marginal $\pi_{\boldsymbol{F}_{\mathrm{DC},t}} = \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]$ is determined by short-range rather than long-range effects. Since Belief Propagation is essentially a local algorithm, one might expect that the (non-)reconstruction phase transition coincides with the threshold up to which BPGD succeeds; cf. the discussions in [5, 17].

To define (non-)reconstruction precisely, we associate a bipartite graph $G(\boldsymbol{F}_{\mathrm{DC},t})$ with the formula $\boldsymbol{F}_{\mathrm{DC},t}$. The vertices of this graph are the variables and clauses of $\boldsymbol{F}_{\mathrm{DC},t}$. Each variable is adjacent to the clauses in which it appears. For a (variable or clause) vertex v of $G(\boldsymbol{F}_{\mathrm{DC},t})$ let ∂v be the set of vs neighbours. More generally, for an integer $\ell \geq 1$ let $\partial^{\ell} v$ be the set of vertices of $G(\boldsymbol{F}_{\mathrm{DC},t})$ at shortest path distance precisely ℓ from v. Following [17], we say that $\boldsymbol{F}_{\mathrm{DC},t}$ has the *non-reconstruction property* if

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{E}\left[\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \,\middle|\, \boldsymbol{F}_{\mathrm{DC},t}, \left\{\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(y)\right\}_{y\in\partial^{2\ell}x_{t+1}}\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]\right| \,\middle|\, \boldsymbol{F} \text{ satisfiable}\right] = 0. \tag{1.7}$$

Conversely, $\boldsymbol{F}_{\mathrm{DC},t}$ has the *reconstruction property* if

$$\liminf_{\ell\to\infty} \liminf_{n\to\infty} \mathbb{E}\left[\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \,\middle|\, \boldsymbol{F}_{\mathrm{DC},t}, \left\{\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(y)\right\}_{y\in\partial^{2\ell}x_{t+1}}\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_{t+1}) = 1 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]\right| \,\middle|\, \boldsymbol{F} \text{ sat.}\right] > 0. \tag{1.8}$$

To parse (1.7), notice that in the left probability term we condition on both the outcome $\boldsymbol{F}_{\mathrm{DC},t}$ of the first t steps of the decimation process and on the values $\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(y)$ that the random solution $\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}$ assigns to the variables y at distance exactly 2ℓ from x_{t+1}. By contrast, in the right probability term we only condition on $\boldsymbol{F}_{\mathrm{DC},t}$. Thus, the second probability term matches the probability $\pi_{\boldsymbol{F}_{\mathrm{DC},t}}$ from the decimation process. Hence, (1.7) compares the probability that a random solution sets x_{t+1} to one given the values $\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(y)$ of *all* variables y at distance 2ℓ from x_{t+1} with plain marginal probability that x_{t+1} is set to one. What (1.7) asks is that these two probabilities be asymptotically equal in the limit of large ℓ, with high probability over the choice of $\boldsymbol{F}$ and the prior steps of the decimation process. Thus, so long as non-reconstruction holds 'long-range effects', meaning anything beyond distance 2ℓ for large enough but fixed ℓ, are negligible.

Confirming the predictions from [25], the following theorem identifies the precise regimes of d, t where (non-)reconstruction holds. To state the theorem, we need to know that for $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$ the polynomial $d(k-1)z^{k-2}(1-z)-1$ has precisely two roots $0 < z_* = z_*(d,k) < z^* = z^*(d,k) < 1$; we are going to prove this as part of Proposition 2.2 below. Let

$$\lambda_* = \lambda_*(d,k) = -\log(1-z_*) - \frac{z_*}{(k-1)(1-z_*)} > \lambda^* = \lambda^*(d,k) = \max\left\{0, -\log(1-z^*) - \frac{z^*}{(k-1)(1-z^*)}\right\} \geq 0, \quad (1.9)$$

$$\theta_* = \theta_*(d,k) = 1-\exp(-\lambda_*) > \theta^* = \theta^*(d,k) = 1-\exp(-\lambda^*). \quad (1.10)$$

Additionally, let $\lambda_{\mathrm{cond}}(d,k)$ be the solution to the ODE

$$\frac{\partial \lambda_{\mathrm{cond}}(d,k)}{\partial d} = -\frac{\alpha^*(\lambda_{\mathrm{cond}}(d,k))^k - \alpha_*(\lambda_{\mathrm{cond}}(d,k))^k}{k(\alpha^*(\lambda_{\mathrm{cond}}(d,k)) - \alpha_*(\lambda_{\mathrm{cond}}(d,k)))}, \qquad \lambda_{\mathrm{cond}}(d_{\mathrm{sat}}(k),k) = 0 \quad (1.11)$$

on the interval $(d_{\min}, d_{\mathrm{sat}}]$ and set $\theta_{\mathrm{cond}} = \theta_{\mathrm{cond}}(d,k) = 1-\exp(-\lambda_{\mathrm{cond}}(d,k))$. Note that

$$\theta^* < \theta_{\mathrm{cond}} < \theta_*.$$

Theorem 1.2. *Let $k \geq 3$ and let $0 \leq t = t(n) \leq n$ be a sequence such that $\lim_{n\to\infty} t/n = \theta \in (0,1)$.*

(i) If $d < d_{\min}(k)$, then $\boldsymbol{F}_{\mathrm{DC},t}$ has the non-reconstruction property w.h.p.
(ii) If $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$ and $\theta < \theta^*$ or $\theta > \theta_{\mathrm{cond}}$, then $\boldsymbol{F}_{\mathrm{DC},t}$ has the non-reconstruction property w.h.p.
(iii) If $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$ and $\theta^* < \theta < \theta_{\mathrm{cond}}$, then $\boldsymbol{F}_{\mathrm{DC},t}$ has the reconstruction property w.h.p.

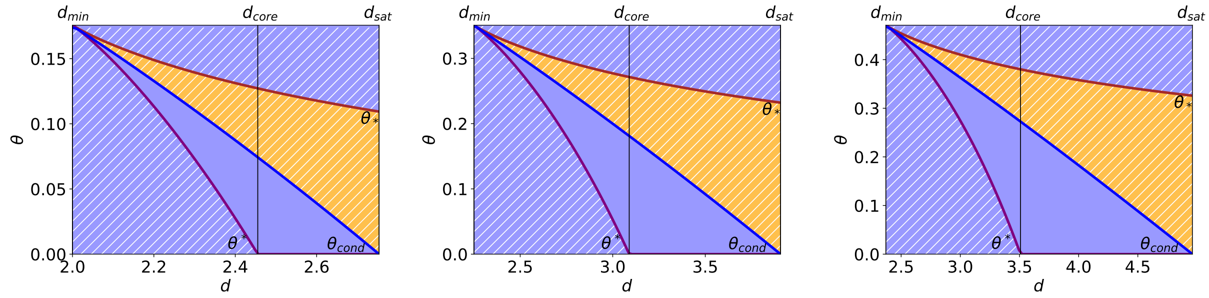
Theorem 1.2 shows that $d_{\min}(k)$ marks the precise threshold of d up to which the decimation process $\boldsymbol{F}_{\mathrm{DC},t}$ exhibits non-reconstruction for *all* $0 \leq t \leq n$ w.h.p. By contrast, for $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$ there is a regime of t where reconstruction occurs. In fact, as Proposition 2.2 shows, for $d > d_{\mathrm{core}}(k)$ we have $\theta^* = 0$ and thus reconstruction holds even at $t = 0$, i.e., for the original, undecimated random formula $\boldsymbol{F}$. Prior to the contribution [25], it had been suggested that this precise scenario (reconstruction on the original problem instance) is the stone on which BPGD stumbles [5]. In fact, Yung's negative result kicks in at this precise threshold $d_{\mathrm{core}}(k)$. However, Theorems 1.1 and 1.2 show that matters are more subtle. Specifically, for $d_{\min}(k) < d < d_{\mathrm{core}}(k)$ reconstruction, even though absent in the initial formula $\boldsymbol{F}$, occurs at a later 'time' $t > 0$ as decimation proceeds, which suffices to trip BPGD up. Also, remarkably, Theorem 1.2 shows that non-reconstruction is not 'monotone'. The property holds for $\theta < \theta^*$ and then again for $\theta > \theta_{\mathrm{cond}}$, but not on the interval $(\theta^*, \theta_{\mathrm{cond}})$ as visualised in Figure 3.

But there is one more surprise. Namely, Theorem 1.2 (ii) might suggest that for $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$ Belief Propagation manages to compute the correct marginals for $t/n \sim \theta > \theta_{\mathrm{cond}}$, as non-reconstruction kicks back in. But remarkably, this is not quite true. Despite the fact that non-reconstruction holds, BPGD goes astray because the algorithm starts its message passing process from a mistaken, oblivious initialisation. As a consequence, for $t/n \sim \theta \in (\theta_{\mathrm{cond}}, \theta^*)$ the BP 'approximations' remain prone to error. To be precise, the following result identifies the precise 'times' where BP succeeds/fails. To state the result let $\mu_{\boldsymbol{F}_{\mathrm{DC},t}}$ denote the BP 'approximation' of the true marginal $\pi_{\boldsymbol{F}_{\mathrm{DC},t}}$ of variable x_{t+1} in the formula $\boldsymbol{F}_{\mathrm{DC},t}$ created by the decimation process (see Section 2.2 for a reminder of the definition). Also recall that $\pi_{\boldsymbol{F}_{\mathrm{DC},t}}$ denotes the correct marginal as used by the decimation process.

Theorem 1.3. *Let $k \geq 3$ and let $0 \leq t = t(n) \leq n$ be a sequence such that $\lim_{n\to\infty} t/n = \theta \in (0,1)$.*

(i) If $0 < d < d_{\min}(k)$ then $\mu_{\boldsymbol{F}_{\mathrm{DC},t}} = \pi_{\boldsymbol{F}_{\mathrm{DC},t}}$ w.h.p.
(ii) If $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$ and $\theta < \theta_{\mathrm{cond}}$ or $\theta > \theta_*$, then $\mu_{\boldsymbol{F}_{\mathrm{DC},t}} = \pi_{\boldsymbol{F}_{\mathrm{DC},t}}$ w.h.p.
(iii) If $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$ and $\theta_{\mathrm{cond}} < \theta < \theta_*$, then $\mathbb{E}\left|\mu_{\boldsymbol{F}_{\mathrm{DC},t}} - \pi_{\boldsymbol{F}_{\mathrm{DC},t}}\right| = \Omega(1)$.

The upshot of Theorems 1.2–1.3 is that the relation between the accuracy of BP and reconstruction is subtle. Everything goes well so long as $d < d_{\min}$ as non-reconstruction holds throughout and the BP approximations are correct. But if $d_{\min} < d < d_{\mathrm{sat}}$ and $\theta^* < \theta < \theta_{\mathrm{cond}}$, then Theorem 1.2 (iii) shows that reconstruction occurs. Nonetheless, Theorem 1.3 (ii) demonstrates that the BP approximations remain valid in this regime. By contrast, for $\theta_{\mathrm{cond}} < \theta < \theta_*$ we have non-reconstruction by Theorem 1.2 (iii), but Theorem 1.3 (iii) shows that BP misses its mark with a non-vanishing probability. Finally, for $\theta > \theta_*$ everything is in order once again as BP regains its footing and non-reconstruction holds. Unfortunately BPGD is unlikely to reach this happy state because the algorithm is bound to make numerous mistakes at times $t/n \in (\theta_{\mathrm{cond}}, \theta_*)$.



(A) $k = 3$ (B) $k = 4$ (C) $k = 5$

FIGURE 3. The phase diagrams for $k = 3,4,5$ with $d \in (d_{\min}, d_{\text{sat}})$ on the horizontal and θ on the vertical axis. The hatched area displays the regime $\theta < \theta^*$ and $\theta_{\text{cond}} < \theta$ where non reconstruction holds. In the non hatched area, where $\theta^* < \theta < \theta_{\text{cond}}$, we have reconstruction. Similarly, the blue area displays $\theta < \theta_{\text{cond}}$ and $\theta > \theta_*$ where BP is correct whereas in the orange area, BP is inaccurate.

Theorems 1.2 and 1.3 confirm the predictions from [25, Section 4]. To be precise, while θ_{cond} matches the predictions of Ricci-Tersenghi and Semerjian, the ODE formula (1.11) for the threshold, which is easy to evaluate numerically, does not appear in [25]. Instead of the ODE formulation, Ricci-Tersenghi and Semerjian define λ_{cond} as the (unique) $\lambda \geq 0$ such that $\Phi_{d,k,\lambda}(\alpha_*) = \Phi_{d,k,\lambda}(\alpha^*)$; Proposition 2.2 below shows that both are equivalent. Illustrating Theorems 1.2–1.3, Figure 3 displays the phase diagram in terms of d and $\theta \sim t/n$ for $k = 3,4,5$.

2. OVERVIEW

This section provides an overview of the proofs of Theorems 1.1–1.3. In the final paragraph we conclude with a discussion of further related work. *We assume throughout that $k \geq 3$ is an integer and that $0 < d < d_{\text{sat}}(k)$. Moreover, $t = t(n)$ denotes an integer sequence $0 \leq t(n) \leq n$ such that $\lim_{n\to\infty} t(n)/n = \theta \in (0,1)$.*

2.1. Fixed points and thresholds.

2.1. Fixed points and thresholds. The first item on our agenda is to study the functions $\phi_{d,k,\lambda}, \Phi_{d,k,\lambda}$ from (1.2)–(1.3). Specifically, we are concerned with the maxima of $\Phi_{d,k,\lambda}$ and the fixed points of $\phi_{d,k,\lambda}$, the combinatorial relevance of which will emerge as we the analyse `BPGD` and the decimation process. We begin by observing that the fixed points of $\phi_{d,k,\lambda}$ are precisely the stationary points of $\Phi_{d,k,\lambda}$.

Fact 2.1. *For any $d > 0, \lambda \geq 0$ the stationary points $z \in (0,1)$ of $\Phi_{d,k,\lambda}$ coincide with the fixed points of $\phi_{d,k,\lambda}$ in $(0,1)$. Furthermore, for a fixed point $z \in (0,1)$ of $\phi_{d,k,\lambda}$ we have*

$$\Phi''_{d,k,\lambda}(z) \begin{cases} < 0 & \text{if } \phi'_{d,k,\lambda}(z) < 1, \\ = 0 & \text{if } \phi'_{d,k,\lambda}(z) = 1, \\ > 0 & \text{if } \phi'_{d,k,\lambda}(z) > 1. \end{cases} \tag{2.1}$$

Proof. Differentiating $\Phi_{d,k,\lambda}$, we obtain

$$\Phi'_{d,k,\lambda}(z) = d(k-1)z^{k-2}\left(\phi_{d,k,\lambda}(z) - z\right). \tag{2.2}$$

Hence, a point $z \in (0,1)$ is a fixed point of $\phi_{d,k,\lambda}$ iff $\Phi'_{d,k,\lambda}(z) = 0$. Differentiating (2.2) once more, we obtain

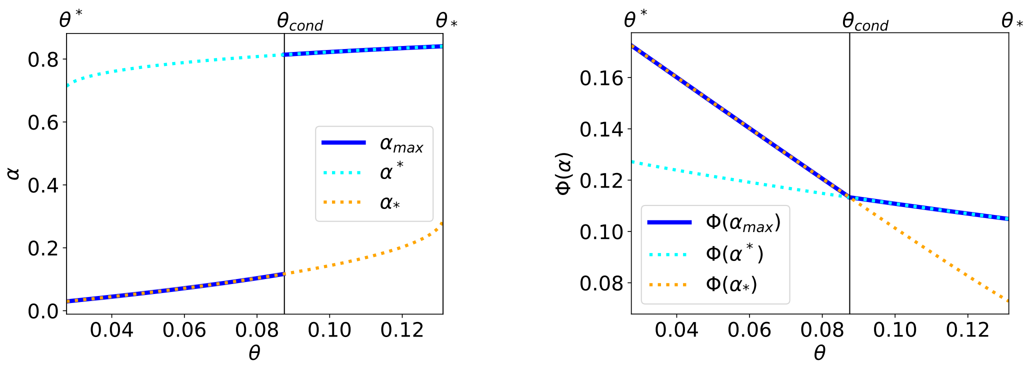
$$\Phi''_{d,k,\lambda}(z) = d(k-1)z^{k-3}\left[(k-2)\left(\phi_{d,k,\lambda}(z) - z\right) + z\left(\phi'_{d,k,\lambda}(z) - 1\right)\right]. \tag{2.3}$$

Clearly, if $\phi_{d,k,\lambda}(z) = z$, then (2.3) simplifies to $\Phi''_{d,k,\lambda}(z) = d(k-1)z^{k-2}(\phi'_{d,k,\lambda}(z) - 1)$, whence (2.1) follows. $\square$

We recall that $0 \leq \alpha_* = \alpha_*(d,k,\lambda) \leq \alpha^* = \alpha^*(d,k,\lambda) \leq 1$ are the smallest and the largest fixed point of $\phi_{d,k,\lambda}$ in $[0,1]$, respectively. Fact 2.1 shows that $\Phi_{d,k,\lambda}$ attains its global maximum in $[0,1]$ at α_* or α^*. Let

$$\alpha_{\max} = \alpha_{\max}(d,k,\lambda) \in \{\alpha_*, \alpha^*\}$$

be the maximiser of $\Phi_{d,k,\lambda}$; if $\Phi_{d,k,\lambda}(\alpha_*) = \Phi_{d,k,\lambda}(\alpha^*)$, set $\alpha_{\max} = \alpha_*$. The following proposition characterises the fixed points of $\phi_{d,k,\lambda}$ and the maximiser $\alpha_{\max}$.



(A) $\alpha_{\max}$ (B) $\Phi(\alpha_{\max})$

FIGURE 4. $\alpha_{\max}$ and $\Phi(\alpha_{\max})$ for $d = 2.4$ and $k = 3$ from θ^* to θ_*.

Proposition 2.2.

(i) If $d < d_{\min}(k)$, then for all $\lambda > 0$ we have $\alpha_*(d,k,\lambda) = \alpha^*(d,k,\lambda)$, the function $\lambda \in (0,\infty) \mapsto \alpha_*(d,k,\lambda) \in (0,1)$ is analytic, and $\alpha_*(d,k,\lambda)$ is the unique stable fixed point of $\phi_{d,k,\lambda}$.

(ii) If $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$, then the polynomial $d(k-1)z^{k-2}(1-z)-1$ has precisely two roots $0 < z_* < z^* < 1$, the numbers λ_*, λ^* from (1.9) satisfy $0 \le \lambda^* < \lambda_*$ and the following is true.

(a) If $\lambda < \lambda^*$ or $\lambda > \lambda_*$, then $\alpha_*(d,k,\lambda) = \alpha^*(d,k,\lambda) \in (0,1)$ is the unique stable fixed point of $\phi_{d,k,\lambda}$.

(b) If $\lambda^* < \lambda < \lambda_*$, then $0 < \alpha_*(d,k,\lambda) < \alpha^*(d,k,\lambda) < 1$ are the only stable fixed points of $\phi_{d,k,\lambda}$.

(c) The functions $\lambda \in (0,\lambda_*) \mapsto \alpha_*(d,k,\lambda)$ and $\lambda \in (\lambda^*,\infty) \mapsto \alpha^*(d,k,\lambda)$ are analytic.

(d) If $d_{\min}(k) < d < d_{\mathrm{sat}}(k)$, then the solution λ_{cond} of (1.11) satisfies $\lambda^* < \lambda_{\mathrm{cond}} = \lambda_{\mathrm{cond}}(d) < \lambda_*$ and

$$\alpha_{\max}(d,k,\lambda) = \begin{cases} \alpha_*(d,k,\lambda) & \text{if } \lambda < \lambda_{\mathrm{cond}}, \\ \alpha^*(d,k,\lambda) & \text{if } \lambda > \lambda_{\mathrm{cond}}. \end{cases}$$

Furthermore, $\Phi_{d,k,\lambda}(\alpha^*(d,k,\lambda)) \neq \Phi_{d,k,\lambda}(\alpha_*(d,k,\lambda))$ unless $\lambda = \lambda_{\mathrm{cond}}$. Thus, the function $\lambda \mapsto \alpha_{\max}(d,k,\lambda)$ is analytic on $(0,\lambda_{\mathrm{cond}})$ and on $(\lambda_{\mathrm{cond}},\infty)$, but discontinuous at $\lambda = \lambda_{\mathrm{cond}}$.

2.2. **Belief Propagation.** Having done our analytic homework, we proceed to recall how Belief Propagation computes the 'approximations' $\mu_{\boldsymbol{F}_{\mathrm{BP},t}}$ that the BPGD algorithm relies upon. We will see that due to the inherent symmetries of XORSAT the Belief Propagation computations simplify and boil down to a simpler message passing process called Warning Propagation. Subsequently we will explain the connection between Warning Propagation and the fixed points α_*, α^* of $\phi_{d,k,\lambda}$.

It is probably easiest to explain BP on a general XORSAT instance F with a set $V(F)$ of variables and a set $C(F)$ of clauses of lengths between one and k. As in Section 1.5 we consider the graph $G(F)$ induced by F, with vertex set $V(F) \cup C(F)$ and an edge xa between $x \in V(F)$ and $a \in C(F)$ iff a contains x. Let $\partial v = \partial_F v$ be the set of neighbours of $v \in V(F) \cup C(F)$. Additionally, given an assignment $\tau \in \{0,1\}^{\partial a}$ of the variables that appear in a, we write $\tau \models a$ iff τ satisfies a.

With each clause/variable pair x, a such that $x \in \partial a$ Belief Propagation associates two sequences of 'messages' $(\mu_{F,x\to a,\ell})_{\ell\ge0}$, $(\mu_{F,a\to x,\ell})_{\ell\ge0}$ directed from x to a and from a to x, respectively. These messages are probability distributions on $\{0,1\}$, i.e.,

$$\mu_{F,x\to a,\ell} = (\mu_{F,x\to a,\ell}(0), \mu_{F,x\to a,\ell}(1)),\ \mu_{F,x\to a,\ell} = (\mu_{F,a\to x,\ell}(0), \mu_{F,a\to x,\ell}(1)) \in [0,1]^2 \qquad \text{and} \tag{2.4}$$

$$\mu_{F,x\to a,\ell}(0) + \mu_{F,x\to a,\ell}(1) = \mu_{F,a\to x,\ell}(0) + \mu_{F,a\to x,\ell}(1) = 1. \tag{2.5}$$

The initial messages are uniform, i.e.,

$$\mu_{F,x\to a,0}(s) = \mu_{F,a\to x,0}(s) = 1/2 \qquad (s \in \{0,1\}). \tag{2.6}$$

Further, the messages at step $\ell+1$ are obtained from the messages at step ℓ via the *Belief Propagation equations*

$$\mu_{F,a\to x,\ell+1}(s) \propto \sum_{\tau\in\{0,1\}^{\partial a}} \mathbb{1}\{\tau_x = s,\, \tau \models a\} \prod_{y\in\partial a\setminus\{x\}} \mu_{F,y\to a,\ell}(\tau_y), \tag{2.7}$$

$$\mu_{F,x\to a,\ell+1}(s) \propto \prod_{b\in\partial x\setminus\{a\}} \mu_{F,b\to x,\ell}(s). \tag{2.8}$$

In (2.7)–(2.8) the $\propto$ -symbol represents the normalisation required to ensure that the updated messages satisfy (2.5). In the case of (2.8) such a normalization may be impossible because the expressions on the r.h.s. could vanish for both $s=0$ and $s=1$. In this event we agree that

$$\mu_{F,x\to a,\ell+1}(s) = \begin{cases} \mu_{F,x\to a,\ell}(s) & \text{if } \mu_{F,x\to a,\ell}(s) \neq 1/2 \\ \mathbb{1}\{s=0\} & \text{otherwise} \end{cases} \qquad (s\in\{0,1\});$$

in other words, we retain the messages from the previous iteration unless its value was $1/2$, in which case we set $\mu_{F,x\to a,\ell+1}(0)=1$. The same convention applies to $\mu_{F,a\to x,\ell+1}(s)$. Further, at any time t the BP messages render a heuristic 'approximation' of the marginal probability that a random solution to the formula F sets a variable x to $s\in\{0,1\}$:

$$\mu_{F,x,\ell}(s) \propto \prod_{b\in\partial x} \mu_{F,b\to x,\ell}(s). \tag{2.9}$$

We set $\mu_{F,x,\ell}(0) = 1-\mu_{F,x,\ell}(1) = 1$ if the normalization in (2.9) fails, i.e., if $\sum_{s\in\{0,1\}}\prod_{b\in\partial x}\mu_{F,b\to x,\ell}(s)=0$.

Fact 2.3. *The BP messages and marginals are half-integral for all t, i.e., for all $t\geq 0$ and $s\in\{0,1\}$ we have*

$$\mu_{F,x\to a,\ell}(s), \mu_{F,a\to x,\ell}(s), \mu_{F,x,\ell}(s) \in \{0,1/2,1\}. \tag{2.10}$$

Furthermore, for all $\ell > 2\sum_{a\in C(F)}|\partial a|$ we have $\mu_{F,x,\ell}(s)=\mu_{F,x,\ell+1}(s)$.

Proof. The half-integrality (2.10) follows from a straightforward induction on ℓ. Furthermore, another induction on ℓ and inspection of (2.7)–(2.8) shows that for any x,a,ℓ such that $\mu_{F,x\to a,\ell}(1)\neq 1/2$ we have $\mu_{F,x\to a,\ell+1}(s) = \mu_{F,x\to a,\ell}(s)$ $(s\in\{0,1\})$. A similar statement holds for $\mu_{F,a\to x,\ell+1}(s)$. In particular, the number of messages that take the value $1/2$ is monotonically decreasing in ℓ. Since the total number of messages is bounded by $2\sum_{a\in C(F)}|\partial a|$, we conclude that the messages will have converged pointwise after this number of iterations. □

Finally, in light of Fact 2.3 it makes sense to define the approximations for `BPGD` by letting

$$\mu_{\boldsymbol{F}_{\mathrm{BP},t}} = \lim_{\ell\to\infty}\mu_{\boldsymbol{F}_{\mathrm{BP},t},x_{t+1},\ell}(1), \qquad\qquad \mu_{\boldsymbol{F}_{\mathrm{DC},t}} = \lim_{\ell\to\infty}\mu_{\boldsymbol{F}_{\mathrm{DC},t},x_{t+1},\ell}(1). \tag{2.11}$$

2.3. **Warning Propagation.** Thanks to the half-integrality (2.10) of the messages, Belief Propagation is equivalent to a purely combinatorial message passing procedure called *Warning Propagation* ('WP') [19]. Similar as BP, WP also associates two message sequences $(\omega_{F,x\to a,\ell},\omega_{F,a\to x,\ell})_{\ell\geq 0}$ with every adjacent clause/variable pair. The messages take one of three possible discrete values $\{\mathtt{f},\mathtt{u},\mathtt{n}\}$ ('frozen', 'uniform', 'null'). To trace the BP messages from Section 2.2 actually only the two values $\{\mathtt{n},\mathtt{u}\}$ would be necessary. However, the third value $\mathtt{f}$ will prove useful in order to compare the BP approximations with the actual marginals. Perhaps unexpectedly given the all-uniform initialisation (2.6), we launch WP from all-frozen start values:

$$\omega_{F,x\to a,0} = \omega_{F,a\to x,0} = \mathtt{f} \qquad\qquad \text{for all } a,x. \tag{2.12}$$

Subsequently the messages get updated according to the rules

$$\omega_{F,a\to x,\ell+1} = \begin{cases} \mathtt{n} & \text{if } \omega_{F,y\to a,\ell}=\mathtt{n} \text{ for all } y\in\partial a\setminus\{x\}, \\ \mathtt{f} & \text{if } \omega_{F,y\to a,\ell}\neq\mathtt{u} \text{ for all } y\in\partial a\setminus\{x\} \text{ and } \omega_{F,y\to a,\ell}\neq\mathtt{n} \text{ for at least one } y\in\partial a\setminus\{x\}, \\ \mathtt{u} & \text{otherwise,} \end{cases} \tag{2.13}$$

$$\omega_{F,x\to a,\ell+1} = \begin{cases} \mathtt{n} & \text{if } \omega_{F,b\to x,\ell}=\mathtt{n} \text{ for at least one } b\in\partial x\setminus\{a\}, \\ \mathtt{f} & \text{if } \omega_{F,b\to x,\ell}\neq\mathtt{n} \text{ for all } b\in\partial x\setminus\{a\} \text{ and } \omega_{F,b\to x,\ell}=\mathtt{f} \text{ for at least one } b\in\partial x\setminus\{a\}, \\ \mathtt{u} & \text{otherwise.} \end{cases} \tag{2.14}$$

In addition to the messages we also define the *mark* of variable node x by letting

$$\omega_{F,x,\ell} = \begin{cases} \mathtt{n} & \text{if } \omega_{F,b\to x,\ell} = \mathtt{n} \text{ for at least one } b \in \partial x, \\ \mathtt{f} & \text{if } \omega_{F,b\to x,\ell} \neq \mathtt{n} \text{ for all } b \in \partial x \text{ and } \omega_{F,b\to x,\ell} = \mathtt{f} \text{ for at least one } b \in \partial x, \\ \mathtt{u} & \text{otherwise.} \end{cases} \tag{2.15}$$

The following statement summarises the relationship between BP and WP.

Fact 2.4. *For all $t \geq 0$ and all x, a we have*

$$\mu_{x\to a,\ell}(1) = 1/2 \qquad \Leftrightarrow \qquad \omega_{F,x\to a,\ell} \neq \mathtt{n}, \tag{2.16}$$

$$\mu_{a\to x,\ell}(1) = 1/2 \qquad \Leftrightarrow \qquad \omega_{F,a\to x,\ell} \neq \mathtt{n}, \tag{2.17}$$

$$\mu_{x,\ell}(1) = 1/2 \qquad \Leftrightarrow \qquad \omega_{F,x,\ell} \neq \mathtt{n}. \tag{2.18}$$

Moreover, for all $\ell > 2|C(F)|$ we have $\omega_{F,x\to a,\ell} = \omega_{F,x\to a,\ell+1}$ and $\omega_{F,a\to x,\ell} = \omega_{F,a\to x,\ell+1}$.

Proof. The fact that $\omega_{F,x\to a,\ell} = \omega_{F,x\to a,\ell+1}$ and $\omega_{F,a\to x,\ell} = \omega_{F,a\to x,\ell+1}$ for all $\ell > 2|C(F)|$ follows from the observation that the number of $\mathtt{f}$-messages is monotonically decreasing, while the number of $\mathtt{n}$-messages is monotonically increasing. The equations (2.16)–(2.18) follow by induction on ℓ. Initially all the messages are uniform in BP, i.e., $\mu_{x\to a,0}(1) = \mu_{a\to x,0}(1) = 1/2$. By contrast, in WP, we start with all frozen values to both variables and clauses as given by (2.12).Then from (2.13),(2.14) and (2.15), for $\ell = 0$,(2.16)–(2.18) holds true. For $\ell = 1$, we get the messages and marginals in BP obtained from the messages at initial step. From (2.7) it follows that if the marginals are uniform then from WP arguments (2.13), it is sure that $\omega_{F,a\to x,1} \neq \mathtt{n}$ because $\omega_{F,y\to a,0} = \mathtt{f}$. The same argument is valid for the other way round. If the WP message at step $\ell = 1$ is not null, then the BP message from (2.7) after normalization become $1/2$. So for $\ell = 1$, (2.16) holds true.
Let us assume the (2.16) is true for any step ℓ.Then for step $\ell + 1$ the messages in BP is obtained from step ℓ as in (2.7) is $\frac{1}{2}$ implies in WP message $\omega_{F,a\to x,\ell+1} \neq \mathtt{n}$ because $\omega_{F,y\to a,\ell} = \mathtt{u}$ for at least one $y \in \partial a \setminus \{x\}$. Similarly, if the WP message $\omega_{F,a\to x,\ell+1} \neq \mathtt{n}$ implies this can be either "uniform" or "frozen". Now, if there will be at least one uniform incoming message then $\mu_{a\to x,\ell+1}(1) = 1/2$ and for all frozen incoming messages it is straightforward from the initialization of WP (2.12) which corresponds to $\mu_{a\to x,\ell+1}(1) = 1/2$. So at step $\ell + 1$, (2.16) holds true. We conclude that (2.16) holds true for every ℓ. Similarly, by induction on ℓ we can conclude that (2.17)–(2.18) also hold true for every ℓ. □

Fact 2.4 implies that the WP messages and marks 'converge' in the limit of large ℓ, in the sense that eventually they do not change any more. Let $\omega_{F,x\to a}, \omega_{F,a\to x}, \omega_{F,x} \in \{\mathtt{f}, \mathtt{u}, \mathtt{n}\}$ be these limits. Furthermore, let $V_{\mathtt{f},\ell}(F)$, $V_{\mathtt{u},\ell}(F)$, $V_{\mathtt{n},\ell}(F)$ be the sets of variables with the respective mark after $\ell \geq 0$ iterations. Also let $V_{\mathtt{f}}(F), V_{\mathtt{u}}(F), V_{\mathtt{n}}(F)$ be the sets of variables where the limit $\omega_{F,x}$ takes the respective value. The following statement traces WP on the random formula $\boldsymbol{F}_{\mathrm{DC},t}$ produced by the decimation process.

Proposition 2.5. *Let $\varepsilon > 0$ and assume that $d > 0$, $t = t(n) \sim \theta n$ satisfy one of the following conditions:*

(i) $d < d_{\min}$, *or*
(ii) $d > d_{\min}$ *and* $\theta \notin \{\theta_*, \theta^*\}$.

Then there exists $\ell_0 = \ell_0(d, \theta, \varepsilon) > 0$ such that for any fixed $\ell \geq \ell_0$ with $\lambda = -\log(1-\theta)$ w.h.p. we have

$$\left|t + |V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| - \alpha_* n\right| < \varepsilon n, \qquad \left|t + |V_{\mathtt{f},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| - (\alpha^* - \alpha_*)n\right| < \varepsilon n, \qquad \left|V_{\mathtt{n}}(\boldsymbol{F}_{\mathrm{DC},t}) \triangle V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})\right| < \varepsilon n. \tag{2.19}$$

2.4. **The check matrix.** Since the XOR operation is equivalent to addition modulo two, a XORSAT formula F with variables $x_1, \ldots, x_n$ and clauses $a_1, \ldots, a_m$ translates into a linear system over $\mathbb{F}_2$, as follows. Let A_F be the $m \times n$-matrix over $\mathbb{F}_2$ whose (i, j)-entry equals one iff variable x_j appears in clause a_i. Adopting coding parlance, we refer to A_F as the *check matrix* of F. Furthermore, let $y_F \in \mathbb{F}_2^m$ be the vector whose ith entry is one plus the sum of any constant term and the number of negation signs of clause a_i mod two. Then the solutions $\sigma \in \mathbb{F}_n^n$ of the linear system $A_F\sigma = y_F$ are precisely the satisfying assignments of F.

The algebraic properties of A_F therefore have a direct impact on the satisfiability of F. For example, if A_F has rank m, we may conclude immediately that F is satisfiable. Furthermore, the set of solutions of F is an affine subspace of $\mathbb{F}_2^n$ (if non-empty). In effect, if F is satisfiable, then the number of satisfying assignments equals the size of the kernel of A_F. Hence the nullity $\mathrm{nul}\, A_F = \dim\ker A_F$ of the check matrix is a key quantity.

Indeed, the single most significant ingredient towards turning the heuristic arguments from [25] into rigorous proofs is a formula for the nullity of the check matrix of the XORSAT instance $\boldsymbol{F}_{\mathrm{DC},t}$ from the decimation process.

To unclutter the notation set $\boldsymbol{A}_t = A_{\boldsymbol{F}_{\mathrm{DC},t}}$. We derive the following proposition from a recent general result about the nullity of random matrices over finite fields [8, Theorem 1.1]. The proposition clarifies the semantics of the function $\Phi_{d,k,\lambda}$ and its maximiser $\alpha_{\max}$. In physics jargon $\Phi_{d,k,\lambda}$ is known as the Bethe free entropy.

Proposition 2.6. *Let $d > 0$ and $\lambda = -\log(1-\theta)$. Then*

$$\lim_{n\to\infty} \mathrm{nul}\, \boldsymbol{A}_t = \Phi_{d,k,\lambda}(\alpha_{\max}) \qquad \textit{in probability.}$$

2.5. **Null variables.** Proposition 2.6 enables us to derive crucial information about the set of satisfying assignments of $\boldsymbol{F}_{\mathrm{DC},t}$. Specifically, for any XORSAT instance F with variables $x_1,\ldots,x_n$ let $V_0(F)$ be the set of variables x_i such that $\sigma_i = 0$ for all $\sigma \in \ker A_F$. We call the variables $x_i \in V_0(F)$ *null variables.* Since the set of solutions of F, if non-empty, is a translation of $\ker A_F$, any two solutions σ, σ' of F set the variables in $V_0(F)$ to exactly the same values. The following proposition shows that WP identifies certain variables as null.

Proposition 2.7. *W.h.p. the following two statements are true for any fixed integer $\ell > 0$.*

(i) We have $V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \subseteq V_0(\boldsymbol{F}_{\mathrm{DC},t})$.
(ii) We have $|V_{\mathtt{u},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cap V_0(\boldsymbol{F}_{\mathrm{DC},t})| = o(n)$.

Propositions 2.6 and 2.7 enable us to calculate the number of null variables of $\boldsymbol{F}_{\mathrm{DC},t}$, so long as we remain clear of the point θ_{cond} where $\alpha_{\max}$ is discontinuous.

Proposition 2.8. *If $\theta \neq \theta_{\mathrm{cond}}$ then $|V_0(\boldsymbol{F}_{\mathrm{DC},t})| = \alpha_{\max} n + o(n)$ w.h.p.*

Let us briefly summarise what we have learned thus far. First, because all Belief Propagation messages are half-integral, BP reduces to WP. Second, Proposition 2.5 shows that the fixed points α_*, α^* of $\phi_{d,k,\lambda}$ determine the number of variables marked `n` or `f` by WP. Third, the function $\Phi_{d,k,\lambda}$ and its maximiser $\alpha_{\max}$ govern the nullity of the check matrix and thereby the number of null variables of $\boldsymbol{F}_{\mathrm{DC},t}$. Clearly, the null variables x_i are precisely the ones whose actual marginals $\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{F}_{\mathrm{DC},t}}(x_i) = s \mid \boldsymbol{F}_{\mathrm{DC},t}\right]$ are *not* uniform. As a next step, we investigate whether BP/WP identify these variables correctly.

In light of Proposition 2.5, in order to investigate the accuracy of BP it suffices to compare the *numbers* of variables marked `n` by WP with the true marginals. The following corollary summarises the result.

Corollary 2.9. *For any d, θ the following statements are true.*

(i) If $d < d_{\min}$, or *$d > d_{\min}$ and $\theta < \theta_{\mathrm{cond}}$,* or *$d > d_{\min}$ and $\theta > \theta_*$, then $|V_0(\boldsymbol{F}_{\mathrm{DC},t}) \triangle V_{\mathtt{n}}(\boldsymbol{F}_{\mathrm{DC},t})| = o(n)$ w.h.p.*
(ii) If $d > d_{\min}$ and $\theta_{\mathrm{cond}} < \theta < \theta_*$, then $|V_0(\boldsymbol{F}_{\mathrm{DC},t}) \triangle V_{\mathtt{n}}(\boldsymbol{F}_{\mathrm{DC},t})| = \Omega(n)$ w.h.p.

Thus, so long as $d < d_{\min}$ or $d > d_{\min}$ and $\theta < \theta_{\mathrm{cond}}$ or $\theta > \theta_*$, the BP/WP approximations are mostly correct. By contrast, if $d > d_{\min}$ and $\theta_{\mathrm{cond}} < \theta < \theta_*$, the BP/WP approximations are significantly at variance with the true marginals w.h.p. Specifically, w.h.p. BP deems $\Omega(n)$ frozen variables unfrozen, thereby setting itself up for failure. Indeed, Corollary 2.9 easily implies Theorem 1.3, which in turn implies Theorem 1.1 (ii) without much ado.

In addition, to settle the (non-)reconstruction thresholds set out in Theorem 1.2 we need to investigate the *conditional* marginals given the values of variables at a certain distances from x_{t+1} as in (1.7). This is where the extra value `f` from the construction of WP enters. Indeed, for a XORSAT instance F with variables $x_1,\ldots,x_n$ and an integer ℓ let $V_{0,\ell}(F)$ be the set of variables x_i such that $\sigma_i = 0$ for all $\sigma \in \ker A_F$ for which $\sigma_h = 0$ for all variables $x_h \in \partial^\ell x_i$.

Corollary 2.10. *Assume that $d > d_{\min}$ and let $\varepsilon > 0$.*

(i) If $\theta < \theta_{\mathrm{cond}}$, then for any fixed ℓ we have $|V_{\mathtt{f},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cap V_{0,\ell}(\boldsymbol{F}_{\mathrm{DC},t})| < \varepsilon n$ w.h.p.
(ii) If $\theta > \theta_{\mathrm{cond}}$, then there exists $\ell_0 = \ell_0(d,\theta,\varepsilon)$ such that for any fixed $\ell > \ell_0$ we have

$$|(V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cup V_{\mathtt{f},\ell}(\boldsymbol{F}_{\mathrm{DC},t})) \triangle V_{0,\ell}(\boldsymbol{F}_{\mathrm{DC},t})| < \varepsilon n \qquad \textit{w.h.p.}$$

Comparing the number of actually frozen variables with the ones marked `f` by WP, we obtain Theorem 1.2.

2.6. **Proving `BPGD` successful.** We are left to prove Theorem 1.1. First, we need to compute the (strictly positive) success probability of `BPGD` for $d < d_{\min}$. At this point, the fact that `BPGD` has a fair chance of succeeding for $d < d_{\min}$ should not come as a surprise. Indeed, Corollary 2.9 implies that the BP approximations of the marginals are mostly correct for $d < d_{\min}$, at least on the formula $\boldsymbol{F}_{\mathrm{DC},t}$ created by the decimation process. Furthermore, so long as the marginals are correct, the decimation process $\boldsymbol{F}_{\mathrm{DC},t}$ and the execution of the `BPGD` algorithm $\boldsymbol{F}_{\mathrm{BP},t}$

move in lockstep. The sole difficulty in analysing BPGD lies in proving that the estimates of the algorithm are not just mostly correct, but correct up to only a *bounded* expected number of discrepancies over the entire execution of the algorithm. To prove this fact we combine the method of differential equations with a subtle analysis of the sources of the remaining bounded number of discrepancies. These discrepancies result from the presence of short (i.e., bounded-length) cycles in the graph $G(\boldsymbol{F})$. Finally, the proof of the second (negative) part of Theorem 1.1 follows by coupling the execution of BPGD with the decimation process, and invoking Theorem 1.3. The details of both arguments can be found in Section 6.

2.7. **Discussion.** The thrust of the present work is to verify the predictions from [25] on the BPGD algorithm and the decimation process rigorously. Concerning the decimation process, the main gap in the deliberations of Ricci-Tersenghi and Semerjian [25] that we needed to plug is the proof of Proposition 2.8 on the actual number of null variables in the decimation process. The proof of Proposition 2.8, in turn, hinges on the formula for the nullity from Proposition 2.6, whereas Ricci-Tersenghi and Semerjian state the (as it turns out, correct) formulas for the nullity and the number of null variables based on purely heuristic arguments.

Regarding the analysis of the BPGD algorithm, Ricci-Tersenghi and Semerjian state that they rely on the heuristic techniques from the insightful article [11] to predict the formula (1.6), but do not provide any further details; the article [11] principally employs heuristic arguments involving generating functions. By contrast, the method that we use to prove (1.6) is a bit more similar to that of Frieze and Suen [13] for the analysis of a variant of the unit clause algorithm on random k-SAT instances, for which they also obtain the asymptotic success probability. Yet by comparison to the argument of Frieze and Suen, we pursue a more combinatorially explicit approach that demonstrates that certain small sub-formulas that we call 'toxic cycles' are responsible for the failure of BPGD. Specifically, the proof of (1.6) combines the method of differential equations with Poissonisation. Finally, the proof of Theorem 1.1 (ii) is an easy afterthought of the analysis of the decimation process.

Yung's work [27] on random k-XORSAT is motivated by the 'overlap gap paradigm' [14], the basic idea behind which is to show that a peculiar clustered geometry of the set of solutions is an obstacle to certain types of algorithms. Specifically, Yung only considers the Unit Clause Propagation algorithm and (a truncated version of) BPGD. Following the path beaten in [20], Yung performs moment computations to establish the overlap gap property. However, moment computations (also called 'annealed computations' in physics jargon) only provide one-sided bounds. As a consequence, Yung's results require spurious lower bounds on the clause length k ($k \geq 9$ for Unit Clause and $k \geq 13$ for BPGD). By contrast, the present proof strategy pivots on the number of null variables rather than overlaps, and Proposition 2.8 provides the precise 'quenched' count of null variables. A further improvement over [27] is that the present analysis pinpoints the *precise* threshold up to which BPGD (as well as Unit Clause) succeeds for any $k \geq 3$. Specifically, Yung proves that these algorithms fail for $d > d_{\mathrm{core}}$, while Theorem 1.1 shows that failure occurs already for $d > d_{\min}$ with $d_{\min} < d_{\mathrm{core}}$. Conversely, Theorem 1.1 shows that the algorithms succeed with a non-vanishing probability for $d < d_{\min}$. Thus, Theorem 1.1 identifies the correct threshold for the success of BPGD, as well as the correct combinatorial phenomenon that determines this threshold, namely the onset of reconstruction in the decimation process (Theorems 1.2 and 1.3).

The BPGD algorithm as detailed in Section 2.2 applies to a wide variety of problems beyond random k-XORSAT. Of course, the single most prominent example is random k-SAT. Lacking the symmetries of XORSAT, random k-SAT does not allow for the simplification to discrete messages; in particular, the BP messages are not generally half-integral. In effect, BP and WP are no longer equivalent. In addition to random k-XORSAT, the article [25] also provides a heuristic study of BPGD on random k-SAT. But once again due to the lack of half-integrality, the formulas for the phase transitions no longer come as elegant finite-dimensional expressions. Instead, they now come as infinite-dimensional variational problems. Furthermore, the absence of half-integrality also entails that the present proof strategy does not extend to k-SAT.

The lack of inherent symmetry in random k-SAT can partly be compensated by assuming that the clause length k is sufficiently large (viz. larger than some usually unspecified constant k_0). Under this assumption the random k-SAT version of both the decimation process and the BPGD algorithm have been analysed rigorously [6, 10]. The results are in qualitative agreement with the predictions from [25]. In particular, the BPGD algorithm provably fails to find satisfying assignments on random k-SAT instances even below the threshold where the set of satisfying assignments shatters into well-separated clusters [1, 17]. Furthermore, on random k-SAT a more sophisticated message passing algorithm called Survey Propagation Guided Decimation has been suggested [21, 25]. While on random XORSAT Survey Propagation and Belief Propagation are equivalent, the two algorithms are substantially different on random k-SAT. One might therefore hope that Survey Propagation Guided Decimation outperforms BPGD on random k-SAT and finds satisfying assignments up to the aforementioned shattering transition. A negative result to the effect that Survey Propagation Guided Decimation fails asymptotically beyond the shattering transition point for large enough k exists [15]. Yet a complete analysis of Belief/Survey Propagation Guided Decimation on random k-SAT for any $k \geq 3$ in analogy to the results obtained here for random k-XORSAT remains an outstanding challenge.

Finally, returning to random k-XORSAT, a question for future work may be to investigate the performance of various types of algorithms such as greedy, message passing or local search that aim to find an assignment that violates the least possible number of clauses. Of course, this question is relevant even for $d > d_{\mathrm{sat}}(k)$. A first step based on the heuristic 'dynamical cavity method' was recently undertaken by Maier, Behrens and Zdeborová [18].

2.8. **Preliminaries and notation.** Throughout we assume that $k \geq 3$ and $0 < d < d_{\min}$ and $\theta \in (0,1)$ are fixed independently of n. We always let $t = t(n) \in \{0,1,\dots,n\}$ be an integer sequence such that $\lim_{n\to\infty} t/n = \theta$. Unless specified otherwise we tacitly assume that n is sufficiently large for our various estimates to hold. Asymptotic notation such as $O(\cdot)$ refers to the limit of large n by default, with k,d,θ fixed. We continue to denote by $\alpha_* = \alpha_*(\lambda) = \alpha_*(d,k,\lambda)$ and $\alpha^* = \alpha^*(\lambda) = \alpha^*(d,k,\lambda)$ the smallest/largest fixed points of $\phi_{d,k,\lambda}$ in $[0,1]$ and by $\lambda_* = \lambda_*(d,k)$, $\lambda^* = \lambda^*(d,k)$, $\theta_* = \theta_*(d,k)$, $\theta^* = \theta^*(d,k)$ the quantities defined in (1.9)–(1.10).

For a formula F and a partial assignment $\sigma : U \to \{0,1\}$ with $U \subseteq V(F)$ let $F[\sigma]$ be the simplified formula obtained by substituting constants for the variables in U. The *length* of a clause of $F[\sigma]$ is defined as the number of variables from $V(F) \setminus U$ that the clause contains.

The following fact provides the correctness of BP on formulas represented by acyclic graphs $G(F)$.

Fact 2.11 ([19, Chapter 14])**.** *For a XORSAT Formula F with an acyclic bipartite graph $G(F)$ the BP marginals as defined in* (2.9) *are exact, i.e.*

$$\lim_{\ell\to\infty} \mu_{F,x,\ell}(1) = \mathbb{P}\left[\boldsymbol{\sigma}_F(x) = 1\right].$$

2.9. **Organisation.** The rest of the paper is organised as follows. Section 3 contains the proof of Proposition 2.2. Subsequently in Section 4 we investigate Warning Propagation to prove Propositions 2.5 and 2.7. Furthermore, Section 5 deals with the study of the check matrix; here we prove Propositions 2.6 and 2.8 as well as Corollaries 2.9 and 2.10. Additionally, with all these preparations completed we put all the pieces together to complete the proofs of Theorems 1.2 and 1.3 in Section 5.5. Finally, Section 6 contains the proof of Theorem 1.1.

3. Proof of Proposition 2.2

Even though a few steps are mildly intricate, the proof of Proposition 2.2 mostly consists of 'routine calculus'. As a convenient shorthand we introduce

$$\zeta_\lambda(z) = \zeta_{d,k,\lambda}(z) = \phi_{d,k,\lambda}(z) - z = 1 - \exp\left(-\lambda - dz^{k-1}\right) - z.$$

Its derivatives read

$$\zeta'_\lambda(z) = d(k-1)\, z^{k-2}\, \exp(-\lambda - dz^{k-1}) - 1 \quad \text{and} \tag{3.1}$$

$$\zeta''_\lambda(z) = d(k-1)\, z^{k-3}\, \exp(-\lambda - dz^{k-1})\, \left[(k-2) - d(k-1)z^{k-1}\right]. \tag{3.2}$$

Also let

$$z_0 = z_0(d,k) = \left(\frac{k-2}{d(k-1)}\right)^{\frac{1}{k-1}}. \tag{3.3}$$

We begin by investigating the zeros of ζ_λ, obviously identical with fixed points of $\phi_{d,k,\lambda}$.

Lemma 3.1. *Assume that $\lambda > 0$.*

(i) The function ζ_λ has either one or three zeros in $z \in [0,1]$, possibly including multiple zeros. If ζ_λ has three zeros, then at least one lies in the interval $[0, z_0]$ and at least one lies in the interval $[z_0, 1]$.
(ii) Also, ζ_λ has at most two stationary points, a minimum and a maximum, and if it has both, the minimum occurs left of the maximum.
(iii) If ζ_λ has a unique zero, then α_* is a stable fixed point of $\phi_{d,k,\lambda}$ and $\sup_{z\in[0,1]} \phi'_{d,k,\lambda}(z) < 1$.

(iv) *If ζ_λ has three zeros but no double zero, then α_*, α^* are stable fixed points of $\phi_{d,k,\lambda}$. Additionally, $\phi_{d,k,\lambda}$ possesses an unstable fixed point $\alpha_{\mathrm{u}} \in (\alpha_*, \alpha^*)$. Furthermore, there exists $\varepsilon = \varepsilon(d,\lambda) > 0$ such that*

$$\sup_{z\in[0,\alpha_*+\varepsilon]} \phi'_{d,k,\lambda}(z) < 1, \qquad \sup_{z\in[\alpha^*-\varepsilon,1]} \phi'_{d,k,\lambda}(z) < 1.$$

Proof. Since $\zeta_\lambda(0) > 0$ and $\zeta_\lambda(1) < 0$, the number of zeros must be odd, so towards (i) it suffices to show that there cannot be more than three zeros. Indeed, by Rolle's theorem, between any two zeros of ζ_λ there is a zero of ζ'_λ. So, if ζ_λ had four or more zeros then ζ'_λ would have at least three zeros in $(0,1]$, and in turn ζ''_λ would have at least two. From (3.2) it is clear that ζ''_λ has only two zeros, at $z = 0$ (outside the relevant range) and at the inflection point where $k-2 = d(k-1)z^{k-1}$, namely for $z = z_0$. So, ζ''_λ has at most two zeros, thus ζ_λ has at most three zeros, therefore either one or three.

The second assertion follows from $\zeta''_\lambda(z_0) = 0$ and that by inspection of (3.2), $\zeta''_\lambda(z)$ is decreasing in z, so a local minimum of ζ_λ at z_1 implies $\zeta''_\lambda(z_1) > 0$ thus $z_1 < z_0$, and symmetrically a local maximum at z_2 implies that $z_2 > z_0$.

Moving on to (iii), we observe that $\zeta_\lambda(\alpha_*) = 0$. Furthermore, since $\zeta_\lambda(0) > 0$ while $\zeta_\lambda(1) < 0$, we conclude that $\zeta'_\lambda(\alpha_*) < 0$, which implies that $0 < \phi'_{d,k,\lambda}(\alpha_*) < 1$. Hence, α_* is a stable fixed point.

With respect to (iv), if ζ_λ has three zeros, then $\alpha_* < \alpha^*$ are the smallest and the largest zero, respectively. Since we assume that ζ_λ does not have a double zero, the same reasoning as under (iii) shows that $\zeta'_\lambda(\alpha_*) < 0$ and thus $0 < \phi'_{d,k,\lambda}(\alpha_*) < 1$. Further, if ζ_λ has three zeros, then by Rolle's theorem and (ii) the function has a local minimum followed by a local maximum, which is followed by the zero α^*. Hence, $\zeta'_\lambda(\alpha^*) < 0$, and thus $0 < \phi'_{d,k,\lambda}(\alpha^*) < 1$. □

The following statement implies that $\phi_{d,k,\lambda}$ has only a single fixed point if $d < d_{\min}$.

Lemma 3.2. *Let $\lambda > 0$. If $d < d_{\min}$, then ζ_λ has a unique zero and is strictly decreasing.*

Proof. Suppose that z is a zero of ζ_λ. Then $\exp(-\lambda - dz^{k-1}) = 1 - z$ and thus

$$\phi'_{d,k,\lambda}(z) = d(k-1)z^{k-2}\exp(-\lambda - dz^{k-1}) = d(k-1)(z^{k-2} - z^{k-1}). \tag{3.4}$$

The expression on the r.h.s. is positive for $z \in (0,1)$ and zero at $z \in \{0,1\}$. Moreover, its derivative works out to be

$$\frac{\partial}{\partial z} d(k-1)(z^{k-2} - z^{k-1}) = d(k-1)z^{k-3}(k-2-(k-1)z).$$

Thus, the expression on the r.h.s. of (3.4) takes its maximum value of $d((k-2)/(k-1))^{k-2}$ at $z^\dagger = (k-2)/(k-1)$. Hence, (3.4) implies that $\phi'_{d,k,\lambda}(z) < 1$ and thus $\zeta'_\lambda(z) < 0$. Consequently, the function $\phi_{d,k,\lambda}$ only has stable fixed points and thus has only a single fixed point by Lemma 3.1. □

Proceeding to average degrees $d > d_{\min}$, we verify that the values λ_*, λ^* from Section 1.5 are well defined and satisfy the inequality (1.9).

Lemma 3.3. *If $d > d_{\min}$, then the polynomial $d(k-1)z^{k-2}(1-z) - 1$ has precisely two roots $0 < z_* < z^* < 1$ and the values λ_*, λ^* defined in (1.9) satisfy $\lambda_* > \lambda^*$. Furthermore, $d_{\mathrm{core}} > d_{\min}$ and $\lambda^* = 0$ iff $d \geq d_{\mathrm{core}}$.*

Proof. Let $z^\dagger = (k-2)/(k-1)$. The polynomial $z^{k-2}(1-z)$ is non-negative on $[0,1]$, strictly increasing on $[0, z^\dagger]$ and strictly decreasing on $[z^\dagger, 1]$. Hence, at $z^\dagger$ the polynomial attains its maximum value of

$$\max_{0\leq z\leq 1} z^{k-2}(1-z) = \frac{(k-2)^{k-2}}{(k-1)^{k-1}}. \tag{3.5}$$

If $d > d_{\min}$, the equation

$$z^{k-2}(1-z) = \frac{1}{d(k-1)}. \tag{3.6}$$

therefore has two distinct solutions $0 < z_* < z^\dagger < z^* < 1$. Letting

$$\mathfrak{l}(z) = -\log(1-z) - \frac{z}{(1-z)(k-1)},$$

we obtain $\lambda_* = \mathfrak{l}(z_*)$ and $\lambda^* = \max\{\mathfrak{l}(z^*), 0\}$.

The function $\mathfrak{l}(z)$ is positive and monotonically increasing on $(0, z^\dagger)$, and monotonically decreasing on $(z^\dagger, 1)$. Indeed, the derivative works out to be

$$\mathfrak{l}'(z) = \frac{k-2-(k-1)z}{(k-1)(1-z)^2}, \tag{3.7}$$

which is positive for small $z > 0$ and has its unique zero at $z^\dagger$. Since $z_* < z^\dagger$, we conclude that $\lambda_* > 0$.

Further, [8, Theorem 1.2] shows that at $d = d_{\mathrm{core}}$ we have $\mathfrak{l}(z^*) = 0$. Since z^* is an increasing function of d while $\mathfrak{l}(z)$ is strictly decreasing in $z > z^\dagger$, we conclude that $\mathfrak{l}(z^*) < 0$ for $d > d_{\mathrm{core}}$, $\mathfrak{l}(z^*) = 0$ for $d = d_{\mathrm{core}}$ and $\mathfrak{l}(z^*) = \lambda^* > 0$ for $d_{\min} < d < d_{\mathrm{core}}$.

Thus, we are left to verify that $\lambda_* > \lambda^*$, which amounts to showing that $\mathfrak{l}(z^*) < \mathfrak{l}(z_*)$. Rearranging (3.6) into $d = 1/((k-1)(1-z_*)z_*^{k-2})$ and $d = 1/((k-1)(1-z^*)z^{*\,k-2})$ and applying the inverse function theorem, we obtain

$$\frac{\partial z_*}{\partial d} = -\frac{(k-1)(1-z_*)^2 z_*^{k-1}}{k-2-(k-1)z_*}, \qquad \frac{\partial z^*}{\partial d} = -\frac{(k-1)(1-z^*)^2 z^{*\,k-1}}{k-2-(k-1)z^*}. \tag{3.8}$$

Combining (3.7) and (3.8) with the chain rule, we arrive at

$$\frac{\partial}{\partial d}\mathfrak{l}(z_*) = -z_*^{k-1}, \qquad \frac{\partial}{\partial d}\mathfrak{l}(z^*) = -z^{*\,k-1}. \tag{3.9}$$

Since $z^* > z_*$ for all $d > d_{\min}$, integrating (3.9) on d shows that $\lambda_* > \lambda^*$, thereby completing the proof. □

We are ready to identify the zeros of ζ_λ for $d > d_{\min}$, depending on the regime of λ.

Lemma 3.4. *Let $\lambda > 0$ and assume that $d > d_{\min}$.*

(i) If $\lambda < \lambda^*$, then ζ_λ has a unique zero.
(ii) If $\lambda^* < \lambda < \lambda_*$, then ζ_λ has three distinct zeros.
(iii) If $\lambda > \lambda_*$, then ζ_λ has a unique zero.

Proof. Assume that $d > d_{\min}$. For fixed k and d, the function ζ_λ varies continuously with λ, so there are contiguous regimes of λ where it has one zero, regimes where it has three zeros, and these regimes are divided by critical values of λ where ζ_λ has three zeros two of which consist of a double zero. In this case, the slope at the double zero is also 0. (By Rolle's theorem, the slope is 0 somewhere between the two zeros, and this is the limiting case.)

Thus, the separation between the regimes with one and three zeros occurs at values of λ such that $\zeta_\lambda(z) = \zeta'_\lambda(z) = 0$. Recalling the definition of ζ_λ and the derivative ζ'_λ from (3.1), we obtain

$$1 - z = \exp(-\lambda - dz^{k-1}) \qquad \text{and} \qquad d(k-1)z^{k-2} = \frac{1}{\exp(-\lambda - dz^{k-1})}. \tag{3.10}$$

Substituting the left equation for the exponential in the right equation, we conclude that (3.10) holds only if z is a solution to (3.6). Further, substituting the two solutions $0 < z_* < z^\dagger = (k-2)/(k-1) < z^*$ into either one of the equations from (3.10) and solving for λ, we obtain

$$\lambda_* = -\log(1-z_*) - \frac{z_*}{(1-z_*)(k-1)}, \qquad \lambda^\star = -\log(1-z^*) - \frac{z^*}{(1-z^*)(k-1)}.$$

Observe that $\lambda^* = \max\{\lambda^\star, 0\}$.

Suppose $0 < \lambda < \lambda_*$. Since $\zeta_{\lambda_*}(z_*) = 0$, the function $\lambda \mapsto \zeta_\lambda(z_*)$ is strictly increasing and $\zeta_\lambda(0) > 0$, we conclude that ζ_λ has a zero in the interval $(0, z_*)$. Similarly, if $\lambda > \lambda^*$, then the function ζ_λ has a zero in the interval $(z^*, 1)$. Hence, (ii) is an immediate consequence of Lemma 3.1.

Now assume that $0 < \lambda < \lambda^*$. Since $\lambda_* > \lambda^*$ by Lemma 3.3, Lemma 3.1 implies that ζ_{λ^*} has precisely three zeros. The largest one is $\alpha^* = z^*$, satisfies $\alpha^* > z^\dagger > z_0$, is a double zero and simultaneously a local maximum of ζ_{λ^*}. Since α^* is a double zero and a local maximum, the smallest zero α_* satisfies $\alpha_* < z_0$ by Rolle's theorem. Hence, $\zeta'_{\lambda^*}(z) < 0$ for all $0 < z < \alpha_*$. Since the function $\lambda \mapsto \zeta_\lambda(z)$ is strictly increasing for all $z \in (0,1)$, Lemma 3.1(i) implies that for $\lambda < \lambda^*$ only a single zero remains, which is smaller than z_0.

Finally, suppose that $\lambda > \lambda_* > \lambda^*$. Lemma 3.1(i) implies that ζ_{λ_*} has precisely three zeros, with a double zero occurring at z_* and another zero at $\alpha^*(\lambda_*) > z^\dagger > z_0$. By Lemma 3.1 and the choice of z_*, λ_*, the double zero at z_* is a local minimum. Therefore, $\zeta'_{\lambda_*}(z) < 0$ for all $z > \alpha^*$. Since the function $\lambda \mapsto \zeta_\lambda(z)$ is strictly increasing for all $z \in (0,1)$, we conclude that $\zeta_\lambda(z) > 0$ for all $\lambda > \lambda_*$ and $z \in [0, z_0]$. Hence, by Lemma 3.1(i) for $\lambda > \lambda_*$ only a single zero remains, which lies in the interval $[z_0, 1]$. □

Combining Lemmas 3.3 and 3.4, we can now verify the analytic properties of the function $\lambda \mapsto \alpha_*$ and $\lambda \mapsto \alpha^*$.

Lemma 3.5. *Let $0 < d < d_{\mathrm{sat}}$ and $\lambda > 0$.*

(i) If $d < d_{\min}$, then the function $\lambda \in (0,\infty) \mapsto \alpha_* = \alpha^*$ is analytic with derivative

$$\frac{\partial \alpha_*}{\partial \lambda} = \frac{1-\alpha_*}{1-d(k-1)\alpha_*^{k-2}(1-\alpha_*)} < 1. \tag{3.11}$$

(ii) If $d > d_{\min}$, then $\lambda \in (0,\lambda_*) \mapsto \alpha_*$ is analytic with derivative (3.11).

(iii) If $d > d_{\min}$, then $\lambda \in (\lambda^*,\infty) \mapsto \alpha^*$ is analytic differentiable with derivative

$$\frac{\partial \alpha^*}{\partial \lambda} = \frac{1-\alpha^*}{1-d(k-1)\alpha^{*\,k-2}(1-\alpha^*)}.$$

Proof. Assume that $d > d_{\min}$ and $\lambda \in (0,\lambda_*)$. We know from the proof of Lemma 3.4 that z_* is a double root and local minimum of ζ_{λ_*}. Furthermore, $z_* < z_0$ and the function $\lambda \mapsto \zeta_\lambda(z)$ is strictly increasing in λ. Hence, Lemma 3.1 implies that for any $0 < \lambda < \lambda_*$ the function ζ_λ has a unique zero in $(0, z_*)$. Similarly, if $d < d_{\min}$ then Lemma 3.2 shows that ζ_λ has a unique zero at α_*. Therefore, the implicit function theorem implies that in cases (i) and (ii) the function $\lambda \mapsto \alpha_*$ is continuously differentiable.

Thus, we are left to work out $\partial\alpha_*(d,k,\lambda)/\partial\lambda$. Consider the function $\mathscr{A} : \binom{z}{\lambda} \mapsto \binom{\zeta_\lambda(z)}{\lambda}$, which is one-to-one in an open interval around α_*. The Jacobi matrix reads

$$D\mathscr{A} = \begin{pmatrix} \partial(\phi_{d,k,\lambda})/\partial\alpha - 1 & \partial\phi_{d,k,\lambda}/\partial\lambda \\ 0 & 1 \end{pmatrix}.$$

Furthermore,

$$\begin{aligned} \frac{\partial\phi_{d,k,\lambda}}{\partial\alpha}\Big|_{\alpha=\alpha_*} &= d(k-1)\alpha^{k-2}\exp(-\lambda - d\alpha^{k-1})\big|_{\alpha=\alpha_*} = d(k-1)\alpha_*^{k-2}(1-\alpha_*), \\ \frac{\partial\phi_{d,k,\lambda}}{\partial t}\Big|_{\alpha=\alpha_*} &= \exp(-\lambda - d\alpha^{k-1})\big|_{\alpha=\alpha_*} = 1-\alpha_*. \end{aligned}$$

Hence, by the inverse function theorem the derivative of $\mathscr{A}^{-1}$ reads

$$(D\mathscr{A})^{-1} = \begin{pmatrix} \left[\partial\phi_{d,k,\lambda}/\partial\alpha - 1\right]^{-1} & -\left[\partial\phi_{d,k,\lambda}/\partial\lambda\right] / \left[\partial\phi_{d,k,\lambda}/\partial\alpha - 1\right] \\ 0 & 1 \end{pmatrix}, \qquad \text{and thus}$$

$$\frac{\partial\alpha_*}{\partial\lambda} = -\frac{\partial\phi_{d,k,\lambda}/\partial\lambda}{\partial\phi_{d,k,\lambda}/\partial\alpha - 1} = \frac{1-\alpha_*}{1-d(k-1)\alpha_*^{k-2}(1-\alpha_*)}.$$

Thus, we obtain (i) and (ii). A similar argument applies to $\lambda \in (\lambda^*,\infty) \mapsto \alpha^*$ in the case $d > d_{\min}$ and yields (iii). $\square$

As a final preparation towards the proof of Proposition 2.2 we investigate the solution λ_{cond} to the differential equation (1.11); notice that Lemma 3.5 shows that this ODE does indeed possess a unique solution on $(d_{\min}, d_{\text{sat}}]$.

Lemma 3.6. *For any $0 < d < d_{\text{sat}}$ we have $0 < \lambda_{\text{cond}} < \lambda_*$. Furthermore, for all $0 < \lambda < \lambda_{\text{cond}}$ we have $\Phi_{d,k,\lambda}(\alpha_*) > \Phi_{d,k,\lambda}(\alpha^*)$, while $\Phi_{d,k,\lambda}(\alpha_*) < \Phi_{d,k,\lambda}(\alpha^*)$ for $\lambda_* > \lambda > \lambda_{\text{cond}}$.*

Proof. For $d < d_{\text{sat}}$ define

$$\lambda^*_{\text{cond}} = \inf\{\lambda \ge 0 : \Phi_{d,k,\lambda}(\alpha^*) > \Phi_{d,k,\lambda}(\alpha_*)\} \tag{3.12}$$

For any $d < d_{\text{sat}}$ we have $\Phi_{d,k,0}(0) > \Phi_{d,k,0}(z)$ for all $0 < z \le 1$; this follows from the characterisation of the k-XORSAT threshold from [3, Theorem 1.1]. Hence, $\lambda^*_{\text{cond}} > 0$ for all $d < d_{\text{sat}}$.

Further, the function ζ_{λ_*} has a double zero and a local minimum at $\alpha_* = z_*$. Since the sign of $\zeta_{\lambda_*}(z)$ matches the sign of $\Phi'_{d,k,\lambda_*}(z)$, this means that $\Phi_{d,k,\lambda_*}(\alpha^*) > \Phi_{d,k,\lambda_*}(\alpha_*)$. Hence, there exists $\varepsilon > 0$ such that for $0 < \lambda_* - \varepsilon < \lambda < \lambda_*$ we have $\Phi_{d,k,\lambda}(\alpha^*) > \Phi_{d,k,\lambda}(\alpha_*)$. Therefore,

$$0 < \lambda^*_{\text{cond}} < \lambda_*. \tag{3.13}$$

As a next step we show that

$$\Phi_{d,k,\lambda}(\alpha_*) < \Phi_{d,k,\lambda}(\alpha^*) \qquad \text{for } \lambda_* > \lambda > \lambda^*_{\text{cond}}. \tag{3.14}$$

To this end, we compute the derivatives of $\Phi_{d,k,\lambda}(\alpha_*)$, $\Phi_{d,k,\lambda}(\alpha^*)$ with respect to $0<\lambda<\lambda_*$. Since α_*,α^* are stationary points of $\Phi_{d,k,\lambda}$, the chain rule yields

$$\frac{\partial}{\partial\lambda}\Phi_{d,k,\lambda}(\alpha_*)=\frac{\partial\Phi_{d,k,\lambda}}{\partial\lambda}\Big|_{\alpha_*}+\frac{\partial\Phi_{d,k,\lambda}}{\partial\alpha}\Big|_{\alpha_*}\frac{\partial\alpha_*}{\partial\lambda}=\frac{\partial\Phi_{d,k,\lambda}}{\partial\lambda}\Big|_{\alpha_*}=-\exp(-\lambda-d\alpha_*^{k-1})=\alpha_*-1, \tag{3.15}$$

$$\frac{\partial}{\partial\lambda}\Phi_{d,k,\lambda}(\alpha^*)=\alpha^*-1. \tag{3.16}$$

Since $\alpha_*<\alpha^*$ for all $\lambda^*<\lambda<\lambda_*$, (3.14) follows from (3.15)–(3.16).

Finally, we verify that $\lambda^*_{\mathrm{cond}}$ equals the solution λ_{cond} to the differential equation (1.11). Recalling the definition (3.12), we see that it suffices to check that $\Phi_{d,k,\lambda_{\mathrm{cond}}}(\alpha_*)=\Phi_{d,k,\lambda_{\mathrm{cond}}}(\alpha^*)$ for all $d_{\min}<d<d_{\mathrm{sat}}$. To this end, we notice that by definition of d_{sat} we have $\Phi_{d_{\mathrm{sat}},k,0}(0)=\Phi_{d_{\mathrm{sat}},k,0}(\alpha^*)$, in line with the initial condition $\lambda_{\mathrm{cond}}(d_{\mathrm{sat}})=0$. Additionally, we claim that $\lambda_{\mathrm{cond}}(d_{\mathrm{sat}})$ satisfies

$$\frac{\partial\Phi_{d,k,\lambda_{\mathrm{cond}}}}{\partial d}\Big|_{\alpha^*}=\frac{\partial\Phi_{d,k,\lambda_{\mathrm{cond}}}}{\partial d}\Big|_{\alpha_*}.$$

Indeed, using the chain rule and the fact that α_*,α^* are stationary points, with $\lambda=\lambda(d)$ we obtain

$$\begin{aligned}\frac{\partial\Phi_{d,k,\lambda_{\mathrm{cond}}}(\alpha_*)}{\partial d}&=\frac{\partial\Phi_{d,k,\lambda_{\mathrm{cond}}}}{\partial d}\Big|_{\alpha_*,\lambda_{\mathrm{cond}}}+\frac{\partial\Phi_{d,k,\lambda_{\mathrm{cond}}}}{\partial\alpha}\Big|_{\alpha_*,\lambda_{\mathrm{cond}}}\frac{\partial\alpha_*}{\partial d}+\frac{\partial\Phi_{d,k,\lambda}}{\partial\lambda}\Big|_{\alpha_*,\lambda}\\&=\frac{\partial\Phi_{d,k,\lambda_{\mathrm{cond}}}}{\partial d}\Big|_{\alpha_*,\lambda_{\mathrm{cond}}}+\frac{\partial\Phi_{d,k,\lambda}}{\partial\lambda}\Big|_{\alpha_*,\lambda}=\alpha_*^{k-1}+(\alpha_*-1)\frac{\partial\lambda_{\mathrm{cond}}}{\partial d}.\end{aligned}$$

Analogously,

$$\frac{\partial\Phi_{d,k,\lambda_{\mathrm{cond}}}(\alpha^*)}{\partial d}=\alpha^{*\,k-1}+(\alpha^*-1)\frac{\partial\lambda_{\mathrm{cond}}}{\partial d}.$$

Hence, the solution λ_{cond} to (1.11) satisfies $\Phi_{d,k,\lambda_{\mathrm{cond}}}(\alpha_*)=\Phi_{d,k,\lambda_{\mathrm{cond}}}(\alpha^*)$, and thus $\lambda_{\mathrm{cond}}=\lambda^*_{\mathrm{cond}}$. Therefore, the assertion follows from (3.15) and (3.14). □

Proof of Proposition 2.2. The first assertion is an immediate consequence of Lemmas 3.2 and 3.5. Moreover, the second assertion follows from Lemmas 3.3, 3.4 and 3.5. Finally, the last assertion follows from Lemma 3.6. □

4. WARNING PROPAGATION AND LOCAL WEAK CONVERGENCE

In this section we prove Propositions 2.5 and 2.7. The proofs rely on the concept of local weak convergence. Specifically, we are going to set up a Galton-Watson process that mimics the local topology of the graph $G(\boldsymbol{F}_{\mathrm{DC},t})$ up to any fixed depth ℓ. Subsequently we will analyse WP on the Galton-Watson tree and argue that the result extends to $G(\boldsymbol{F}_{\mathrm{DC},t})$.

4.1. **Local weak convergence.** The construction of the Galton-Watson process $\mathbb{T}=\mathbb{T}(d,k,t)$ is pretty straightforward. The process has two types called *variable nodes* and *check nodes*. The process starts with a single variable node v_0. Furthermore, each variable node begets a $\mathrm{Po}(d)$ number of check nodes as offspring, while the offspring of a check node is a $\mathrm{Bin}(k-1,1-t/n)$ number of variable nodes.

Let $\mathbb{T}$ be the Galton-Watson tree rooted at v_0 that this process generates; $\mathbb{T}$ may be infinite. Hence, for an integer ℓ obtain $\mathbb{T}^{(\ell)}$ from $\mathbb{T}$ by deleting all variable/check nodes at distance greater than 2ℓ from v_0. Thus, $\mathbb{T}^{(\ell)}$ is a finite random tree rooted at v_0. For any graphs T,T' rooted at v,v', respectively, we write $T\cong T'$ if there is a graph isomorphism $\iota:T\to T'$ such that $\iota(v)=v'$. Furthermore, for a vertex v of $G(\boldsymbol{F}_{\mathrm{DC},t})$ and an integer ℓ we let $\partial^{\le\ell}_{\boldsymbol{F}_{\mathrm{DC},t}}v$ be the subgraph obtained from $G(\boldsymbol{F}_{\mathrm{DC},t})$ by deleting all vertices at distance greater than 2ℓ from v, rooted at v. Finally, for a rooted graph g and an integer ℓ we let $\boldsymbol{N}_t^{(\ell)}(g)$ be the number of vertices v of $G(\boldsymbol{F}_{\mathrm{DC},t})$ such that $\partial^{\le\ell}_{\boldsymbol{F}_{\mathrm{DC},t}}v\cong g$.

Lemma 4.1. *For any rooted tree g we have*

$$\mathbb{E}\left|\boldsymbol{N}_t^{(\ell)}(g)-(n-t)\mathbb{P}\left[\mathbb{T}^{(\ell)}\cong g\right]\right|=o(n). \tag{4.1}$$

Proof. The proof is based on a routine second moment argument; that is, we claim that

$$\mathbb{E}\left[\boldsymbol{N}_t^{(\ell)}(g)\right]=(n-t)\mathbb{P}\left[\mathbb{T}^{(\ell)}\cong g\right]+o(n),\qquad\mathbb{E}\left[\boldsymbol{N}_t^{(\ell)}(g)^2\right]=(n-t)^2\mathbb{P}\left[\mathbb{T}^{(\ell)}\cong g\right]^2+o(n^2). \tag{4.2}$$

Combining (4.2) with the Markov and Chebyshev inequalities then yields the assertion.

We prove (4.2) and thereby (4.1) by induction on ℓ. Recall that $\boldsymbol{F}_{\mathrm{DC},t}$ is a XORSAT instance with variables $x_{t+1},\dots,x_n$. Let us begin with the estimate of the first moment. Due to the linearity of expectation, it suffices to show that

$$\mathbb{P}\left[\partial^{\le\ell}(\boldsymbol{F}_{\mathrm{DC},t},x_{t+1})\cong g\right]=\mathbb{P}\left[\mathbb{T}^{(\ell)}\cong g\right]+o(1). \tag{4.3}$$

For $\ell=0$ there is nothing to show. Hence, suppose that (4.1) is true with ℓ replaced by $\ell-1$. Furthermore, let Δ be the degree of the root r of g and let $1\le\kappa_1\le\ldots\le\kappa_\Delta\le k$ be the degrees of the children of the root; thus, we order the children of r so that their degrees are increasing. For an integer $1\le i\le k$ let K_i be the number $j\in[\Delta]$ such that $\kappa_j=i$. Further, let $(g_{i,j})_{1\le i\le\Delta,1\le j\le\kappa_i}$ be the trees pending on the grandchildren of the root. In addition, let $\boldsymbol{\Delta}$ be the degree of x_{t+1} in $G(\boldsymbol{F}_{\mathrm{DC},t})$ and let $1\le\boldsymbol{\kappa}_1\le\ldots\le\boldsymbol{\kappa}_{\boldsymbol{\Delta}}\le k$ be the degrees of the neighbours of x_{t+1}. Then $\partial^{\le\ell}(\boldsymbol{F}_{\mathrm{DC},t},x_{t+1})\cong g$ is possible only if $\Delta=\boldsymbol{\Delta}$ and $\boldsymbol{\kappa}_i=\kappa_i$ for all $1\le i\le\Delta$. Since the clauses of the random formula $\boldsymbol{F}$ are drawn uniformly and independently and $G(\boldsymbol{F}_{\mathrm{DC},t})$ is obtained from $G(\boldsymbol{F})$ by deleting the variable nodes $x_1,\dots,x_t$ along with any ensuing isolated check nodes, we conclude that the event $\mathscr{D}=\{\boldsymbol{\Delta}=\Delta,\bigwedge_{1\le i\le\Delta}\boldsymbol{\kappa}_i=\kappa_i\}$ has probability

$$\mathbb{P}[\mathscr{D}]=\mathbb{P}[\mathrm{Po}(d)=\Delta]\binom{\Delta}{K_1,\dots,K_k}\prod_{i=1}^{k}\mathbb{P}[\mathrm{Bin}(k-1,1-t/n)=i]^{K_i}. \tag{4.4}$$

Further, let $\mathscr{G}=\{g_{i,j}:1\le i\le\Delta,1\le j\le\kappa_i\}$ and let $\mathscr{E}$ be the event that $\boldsymbol{N}_t^{(\ell-1)}(\gamma)=(n-t)\mathbb{P}\left[\mathbb{T}^{(\ell-1)}\cong\gamma\right]+o(n)$ for all $\gamma\in\mathscr{G}$. Then by induction we have

$$\mathbb{P}[\mathscr{E}\mid\mathscr{D}]=1-o(1). \tag{4.5}$$

Now, obtain $G^-(\boldsymbol{F}_{\mathrm{DC},t})$ from $G(\boldsymbol{F}_{\mathrm{DC},t})$ by deleting x_{t+1} along with its adjacent check nodes. Let $\boldsymbol{N}_t^{(\ell),-}(g_{i,j})$ be the number of vertices v of $G^-(\boldsymbol{F}_{\mathrm{DC},t})$ such that $\partial^{\le\ell}(G^-(\boldsymbol{F}_{\mathrm{DC},t}),v)\cong g_{i,j}$. Moreover, let $\mathscr{E}^-$ be the event that $\boldsymbol{N}_t^{(\ell-1),-}(g_{i,j})=(n-t)\mathbb{P}\left[\mathbb{T}^{(\ell-1)}\cong g_{i,j}\right]+o(n)$ for all i,j. Since x_{t+1} has degree $\Delta=O(1)$ given $\mathscr{D}$ and all adjacent check nodes have degree at most k, (4.5) implies that

$$\mathbb{P}[\mathscr{E}^-\mid\mathscr{D}]=1-o(1). \tag{4.6}$$

Finally, since $\boldsymbol{F}_{\mathrm{DC},t}$ is uniformly random, given $\mathscr{D}$ the checks a of $\boldsymbol{F}_{\mathrm{DC},t}$ adjacent to x_{t+1} simply choose their other neighbours uniformly at random from the variable nodes $x_{t+2},\dots,x_n$ of $G^-(\boldsymbol{F}_{\mathrm{DC},t})$. Therefore, (4.4) implies that

$$\mathbb{P}\left[\partial^{\le\ell}(\boldsymbol{F}_{\mathrm{DC},t},x_{t+1})\cong g\right]=\mathbb{P}\left[\mathbb{T}^{(\ell)}\cong g\right]+o(1),$$

thereby proving (4.3) and thus the first part of (4.2).

The proof of the second part of (4.2) (the estimate of the second moment) proceeds along similar lines, except that we need to explore the depth-2ℓ neighbourhoods of two variable nodes of $\boldsymbol{F}_{\mathrm{DC},t}$ simultaneously. Specifically, the proof of the second moment bound comes down to showing that

$$\mathbb{P}\left[\partial^{\le\ell}(\boldsymbol{F}_{\mathrm{DC},t},x_{t+1})\cong g,\partial^{\le\ell}(\boldsymbol{F}_{\mathrm{DC},t},x_{t+2})\cong g\right]=\mathbb{P}\left[\mathbb{T}^{(\ell)}\cong g\right]^2+o(1). \tag{4.7}$$

Exploiting that the variable nodes x_{t+1},x_{t+2} are at distance greater than 4ℓ w.h.p., we conduct a similar induction as above to verify (4.7) and thus (4.2). □

4.2. **Proof of Proposition 2.5.** To prove Proposition 2.5 we estimate the sizes $|V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})|,|V_{\mathtt{f},\ell}(\boldsymbol{F}_{\mathrm{DC},t})|$ separately. Recall that $\theta\sim t/n$.

Lemma 4.2. *Let $\varepsilon>0$ and assume that one of the following conditions is satisfied:*

(i) $d<d_{\min}$, or

(ii) $d>d_{\min}$ and $|\theta_*-\theta|>\varepsilon$.

Then there exists $\ell_0=\ell_0(d,\varepsilon)>0$ such that for any fixed $\ell\ge\ell_0$ with $\lambda=-\log(1-\theta)$ w.h.p. we have

$$\left|t+|V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})|-\alpha_*n\right|<\varepsilon n.$$

Proof. In light of Lemma 4.1 it suffices to investigate WP on the random tree $\mathbb{T}^{(\ell)}$ for large enough ℓ. Specifically, let $p^{(\ell)}$ be the probability that WP marks the root of $\mathbb{T}^{(\ell)}$ as $\mathtt{n}$. In formulas, recalling (2.15), this means that

$$p^{(\ell)}=\mathbb{P}\left[\omega_{\mathbb{T}^{(\ell)},r,\ell}=\mathtt{n}\right]\text{ for }\ell\ge1,\qquad\text{and }p^{(0)}=0. \tag{4.8}$$

Let $\boldsymbol{\Delta}$ be the degree of the root r of $\mathbb{T}^{(\ell)}$ and let $\boldsymbol{\kappa}_1,\ldots,\boldsymbol{\kappa}_{\boldsymbol{\Delta}}$ be the degrees of the children of r. Since the sub-trees of $\mathbb{T}^{(\ell)}$ pending on the grandchildren of r are independent copies of $\mathbb{T}^{(\ell-1)}$, the WP update rules (2.13)–(2.14) yield the recurrence

$$p^{(\ell)} = 1-\mathbb{E}\left[\prod_{i=1}^{\boldsymbol{\Delta}}\left(1-\prod_{j=0}^{\boldsymbol{\kappa}_i-1} p^{(\ell-1)}\right)\right] \qquad (\ell>0). \tag{4.9}$$

By the construction of $\mathbb{T}$ the degree $\boldsymbol{\Delta}$ of r has distribution $\mathrm{Po}(d)$. Furthermore, each child of r has $\mathrm{Bin}(k-1,1-t/n)$ children; thus, $\boldsymbol{\kappa}_i-1\overset{\mathrm{dist}}{=}\mathrm{Bin}(k-1,1-t/n)$. Consequently, (4.9) yields

$$\begin{aligned} p^{(\ell)} &= 1-\exp(-d)\sum_{\Delta=0}^{\infty}\frac{d^{\Delta}}{\Delta!}\left(1-\sum_{\kappa=0}^{k-1}\binom{k-1}{\kappa}\exp(-\lambda\kappa)(1-\exp(-\lambda))^{k-1-\kappa}p^{(\ell-1)\kappa}\right)^{\Delta} \\ &= 1-\exp\left(-d\left(1-\exp(-\lambda)(1-p^{(\ell-1)})\right)^{k-1}\right). \end{aligned} \tag{4.10}$$

Letting $z^{(\ell)} = 1-\exp(-\lambda)(1-p^{(\ell)})$ and recalling the definition (1.2) of $\phi_{d,k,\lambda}$, we see that (4.10) amounts to

$$z^{(\ell)} = \phi_{d,k,\lambda}(z^{(\ell-1)}). \tag{4.11}$$

Moreover, Lemma 3.1 (iii)–(iv), Lemma 3.2 and Lemma 3.4 show that if (i) or (ii) above hold, then $\phi_{d,k,\lambda}$ is a contraction on $[0,\alpha_*]$. Therefore, (4.11) shows that $\lim_{\ell\to\infty}p^{(\ell)} = \frac{\alpha_*-\theta}{1-\theta}$. Thus, the assertion follows from Lemma 4.1. □

Lemma 4.3. *Let $\varepsilon>0$ and assume that $d>0$, $t=t(n)$ are such that one of the following conditions is satisfied:*

(i) $d<d_{\min}$, or
(ii) $d>d_{\min}$, $|\theta_*-\theta|>\varepsilon$ and $|\theta^*-\theta|>\varepsilon$.

Then there exists $\ell_0=\ell_0(d,\varepsilon)>0$ such that for any fixed $\ell\geq\ell_0$ with $\lambda=-\log(1-\theta)$ w.h.p. we have

$$\left||V_{\mathtt{f},\ell}(\boldsymbol{F}_{\mathrm{DC},t})|-(\alpha^*-\alpha_*)n\right|<\varepsilon n.$$

Proof. Once again it suffices to trace WP on $\mathbb{T}^{(\ell)}$ for large ℓ. As in the proof of Lemma 4.2, let

$$p^{(\ell)} = \mathbb{P}\left[\omega_{\mathbb{T}^{(\ell)},r,\ell}\neq \mathtt{u}\right] \text{ for } \ell\geq 1, \qquad \text{and } p^{(0)}=1. \tag{4.12}$$

Then with $\boldsymbol{\Delta}$ the degree of r and $\boldsymbol{\kappa}_1,\ldots,\boldsymbol{\kappa}_{\boldsymbol{\Delta}}$ the degrees of the children of r, the WP update rules (2.13)–(2.14) translate into

$$p^{(\ell)} = 1-\mathbb{E}\left[\prod_{i=1}^{\boldsymbol{\Delta}}\left(1-\prod_{j=0}^{\boldsymbol{\kappa}_i-1} p^{(\ell-1)}\right)\right] \qquad (\ell>0), \tag{4.13}$$

Thus, the recurrence is identical to (4.8), but this time with the initial condition $p^{(0)}=1$. Hence, letting $z^{(\ell)} = 1-\exp(-\lambda)(1-p^{(\ell)})$ and $z^{(0)}=1$ and retracing the steps towards (4.11), we obtain

$$z^{(\ell)} = 1-\exp(-\lambda)(1-p^{(\ell)}). \tag{4.14}$$

Invoking Lemmas 3.1, 3.2 and 3.4, we conclude that (i) or (ii) ensure that $\phi_{d,k,\lambda}$ contracts on $[0,\alpha^*]$. Consequently, (4.14) implies that $\lim_{\ell\to\infty}p^{(\ell)} = \frac{\alpha^*-\theta}{1-\theta}$. Thus, the assertion follows from Lemmas 4.1 and 4.2. □

Finally, we compare the set $V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})$ obtained after a (large but) bounded number of iterations with the ultimate sets $V_{\mathtt{n}}(\boldsymbol{F}_{\mathrm{DC},t})$ obtained upon convergence of WP. The proof of the following lemma is an adaptation of the argument from [23] for cores of random hypergraphs.

Lemma 4.4. *Assume that $\theta\in(0,1)\setminus\{\theta_*,\theta^*\}$. Then for any $\varepsilon>0$ there exists $\ell_0=\ell_0(d,\varepsilon,\theta)$ such that for all $\ell>\ell_0$ we have $|V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})\triangle V_{\mathtt{n}}(\boldsymbol{F}_{\mathrm{DC},t})|<\varepsilon n$ w.h.p.*

Proof. In place of the WP message passing process from Section 2.3 we consider the following simpler peeling process, which reproduces the same set $V_{\mathtt{n}}(\boldsymbol{F}_{\mathrm{DC},t})$. Let $\boldsymbol{G}_0=G(\boldsymbol{F})$ be the bipartite graph induced by $\boldsymbol{F}_{\mathrm{DC},t}$. For $h\geq 0$ obtain $\boldsymbol{G}_{h+1}$ from $\boldsymbol{G}_h$ by performing the following peeling operation.

Remove all check nodes of degree one along with their variable node neighbours. (4.15)

Clearly, this process will reach a fixed point (i.e., $\boldsymbol{G}_{h+1} = \boldsymbol{G}_h$) after at most $\boldsymbol{m}$ iterations. Moreover, a straightforward induction on ℓ shows that $V(\boldsymbol{G}_0) \setminus V(\boldsymbol{G}_\ell) = V_{\mathtt{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})$ and thus $V(\boldsymbol{G}_0) \setminus V(\boldsymbol{G}_{\boldsymbol{m}}) = V_{\mathtt{n}}(\boldsymbol{F}_{\mathrm{DC},t})$. Hence, it suffices to prove that for large enough $\ell = \ell(d,\varepsilon,\theta)$ we have

$$|V(\boldsymbol{G}_\ell) \triangle V(\boldsymbol{G}_{\boldsymbol{m}})| < \varepsilon n \qquad \text{w.h.p.} \tag{4.16}$$

Towards the proof of (4.16) let $\boldsymbol{d}_h = (\boldsymbol{d}_h(u))_{u \in V(\boldsymbol{G}_h) \cup C(\boldsymbol{G}_h)}$ be the degree sequence of $\boldsymbol{G}_h$. By the principle of deferred decisions $\boldsymbol{G}_h$ is uniformly random given $\boldsymbol{d}_h$. Further, let

$$\boldsymbol{\Delta}_h(j) = \left|\left\{x \in V(\boldsymbol{G}_h) : \boldsymbol{d}_h(x) = j\right\}\right|, \qquad \boldsymbol{\Delta}'_h(j) = \left|\left\{a \in C(\boldsymbol{G}_h) : \boldsymbol{d}_h(a) = j\right\}\right|$$

be the number of variable/check nodes of degree $j \geq 0$. Pick $\delta = \delta(d,\varepsilon,\theta)$, $\delta' = \delta'(d,\delta,\theta)$, $\delta''(d,\delta',\theta)$, small enough and $\ell \geq \ell_0(d,\delta'',\theta)$ large enough. Then Lemma 4.2 implies that w.h.p.

$$|V(\boldsymbol{G}_\ell) \setminus V(\boldsymbol{G}_{\ell+1})| < \delta'' n. \tag{4.17}$$

Furthermore, we claim that

$$\sum_{j\geq 0}\left|\frac{\boldsymbol{\Delta}_\ell(j)}{|V(\boldsymbol{G}_\ell)|} - \mathbb{P}\left[\mathrm{Po}(d(1-\alpha_*^{k-1})) = j\right]\right| < \delta', \qquad \sum_{j\geq 2}\left|\frac{\boldsymbol{\Delta}'_\ell(j)}{|C(\boldsymbol{G}_\ell)|} - \frac{\mathbb{P}\left[\mathrm{Bin}(k,1-\alpha_*) = j\right]}{\mathbb{P}\left[\mathrm{Bin}(k,1-\alpha_*) \geq 2\right]}\right| < \delta'. \tag{4.18}$$

Indeed, Lemma 4.1 shows that we just need to study WP on the random tree $\mathbb{T}^{(\ell)}$, as in the proof of Lemma 4.2. Thus, let $\boldsymbol{\Delta}$ be the degree of the root variable and let $\boldsymbol{\kappa}_1,\ldots,\boldsymbol{\kappa}_{\boldsymbol{\Delta}}$ be the degrees of the children of the root. Since the sub-trees pending on the children of the root are independent copies of $\mathbb{T}^{(\ell-1)}$, Lemma 4.2 shows that the probability that any one of the $\boldsymbol{\Delta}$ children sends a $\mathtt{n}$-message to r falls into the interval $(1-\alpha_*^{k-1}-\delta'', 1-\alpha_*^{k-1}+\delta'')$, provided that ℓ is large enough. Since $\boldsymbol{\Delta} \overset{\mathrm{dist}}{=} \mathrm{Po}(d)$, the first part of (4.18) follows from Poisson thinning.

Similarly, to obtain the second part of (4.18) consider a clause a that is a child of the root r of $\mathbb{T}^{(\ell)}$. Then by the same token as in the previous paragraph the number of children of a that do not send a $\mathtt{n}$-message after ℓ iterations of WP lies in the interval $(1-\alpha_*^{k-1}-\delta'', 1-\alpha_*^{k-1}+\delta'')$. Furthermore, the number of children $a' \neq a$ of r has distribution $\mathrm{Po}(\boldsymbol{\Delta})$. Hence, the probability that the WP-message from r to a equals $\mathtt{n}$ comes to $\alpha_* \pm \delta''$, and this event is independent of the messages that the children of a send to a. Finally, the probability that one of the messages that a receives after ℓ iterations of WP differs from the message received after $\ell-1$ iterations is smaller than δ'' for large enough ℓ. Since the peeling process removes any checks a with at least $k-1$ incoming $\mathtt{n}$-messages, we obtain (4.18).

To complete the proof we are going to deduce from (4.17)–(4.18) that the peeling process (4.15) will remove no more than $\varepsilon n/2$ further nodes from $\boldsymbol{G}_\ell$ before it stops. Following [23], we consider a slowed-down version of the process where no longer all checks of degree one get removed simultaneously, but rather one-at-a-time. Let $(\boldsymbol{G}_\ell[v])_{v\geq 0}$ be the sequences of graphs produced by this modified process, with $\boldsymbol{G}_\ell[0] = \boldsymbol{G}_\ell$ and $\boldsymbol{G}_\ell[v+1] = \boldsymbol{G}_\ell[v]$ if all checks of $\boldsymbol{G}_\ell[v]$ have degree at least two. Further, let $\boldsymbol{U}_\ell[v]$ be the number of unary checks of $\boldsymbol{G}_\ell[v]$. Let $\mathscr{D}$ be the event that the bounds (4.17)–(4.18) hold. Then it suffices to prove that on the event $\{\boldsymbol{U}_\ell[v] > 0\} \cap \mathscr{D}$ we have

$$\mathbb{E}\left[\boldsymbol{U}_\ell[v+1] - \boldsymbol{U}_\ell[v] \mid \boldsymbol{U}_\ell[v]\right] < 0 \qquad \text{for all } 0 \leq v \leq \varepsilon n/2. \tag{4.19}$$

Invoking the principle of deferred decisions, in order to verify (4.19) we compute the expected number of new degree one checks produced by the removal of a single random variable node $\boldsymbol{x}$. Due to (4.18), for $v \leq \varepsilon n/2$ the expected number of neighbours a of $\boldsymbol{x}$ of degree precisely two is bounded by

$$d\mathbb{P}\left[\mathrm{Bin}(k-1,1-\alpha_*) = 1\right] + \delta = d(k-1)(1-\alpha_*)\alpha_*^{k-2} + \delta = \phi'_{d,k,\lambda}(\alpha_*) + \delta < 1,$$

provided that $\delta > 0$ is chosen sufficiently small. Hence, we obtain (4.19). $\square$

Proof of Proposition 2.5. The proposition is an immediate consequence of Lemmas 4.2–4.4. $\square$

4.3. **Proof of Proposition 2.7.** We deal with the two claims separately. Towards the first claim we establish the following stronger, deterministic statement.

Lemma 4.5. *For any XORSAT instance F with variables $V_n = \{x_1,\ldots,x_n\}$ and any integer $\ell \geq 0$ we have $V_{\mathtt{n},\ell}(F) \subseteq V_0(F)$.*

Proof. We proceed by induction on ℓ. For $\ell \le 1$ there is nothing to show because $V_{\mathtt{n},\ell}(F) = \emptyset$ by construction. Hence, assume that $\ell > 1$ and that $V_{\mathtt{n},\ell-1}(F) \subseteq V_0(F)$. If $x \in V_{\mathtt{n},\ell-1}$, then (2.15) shows that there exists a check node $b \in \partial x$ such that $\omega_{F,b\to x,\ell} = \mathtt{n}$. Furthermore, (2.13) shows that if $\omega_{F,b\to x,\ell} = \mathtt{n}$, then for all $y \in \partial b \setminus \{x\}$ we have $\omega_{F,y\to b,\ell-1} = \mathtt{n}$. Additionally, (2.14) shows that if $\omega_{F,y\to b,\ell-1} = \mathtt{n}$, then there exists $a \in \partial y \setminus \{b\}$ such that $\omega_{F,a\to y,\ell-2} = \mathtt{n}$. Hence, (2.15) ensures that $\omega_{F,y,\ell-2} = \mathtt{n}$ and thus

$$y \in V_0(F) \qquad \text{for all } y \in \partial b \setminus \{x\} \tag{4.20}$$

by induction. Now suppose that $\partial b = \{x_{j_1}, \dots, x_{j_h}\}$ with pairwise distinct indices $1 \le j_1, \dots, j_h \le n$ such that $x = x_{j_1}$. Consider $\sigma \in \ker A(F)$. Then (4.20) implies that $\sigma_{j_2} = \dots = \sigma_{j_h} = 0$. Consequently, $\sigma_{j_1} = 0$ and thus $x \in V_0(F)$. □

The following lemma deals with the variables that WP marks $\mathtt{u}$.

Lemma 4.6. *For any fixed $\ell \ge 0$ we have $|V_{\mathtt{u},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cap V_0(\boldsymbol{F}_{\mathrm{DC},t})| = o(n)$ w.h.p.*

Proof. We are going to show by induction on ℓ that $\mathbb{E}|V_{\mathtt{u},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cap V_0(\boldsymbol{F}_{\mathrm{DC},t})| = o(n)$. To this end, because the distribution of $\boldsymbol{F}_{\mathrm{DC},t}$ is invariant under permutations of the variables $x_{t+1}, \dots, x_n$, it suffices to show that

$$\mathbb{P}\left[x_n \in V_{\mathtt{u},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cap V_0(\boldsymbol{F}_{\mathrm{DC},t})\right] = o(1). \tag{4.21}$$

Indeed, let $\mathscr{A}$ be the event that the depth-2ℓ neighbourhood $\partial^{\le \ell} x_n$ of x_n in $\boldsymbol{F}_{\mathrm{DC},t}$ is acyclic. Since Lemma 4.1 shows that $\mathbb{P}[\mathscr{A}] = 1 - o(1)$, towards (4.21) it suffices to prove that on the event $\mathscr{A}$ we have

$$x_n \notin V_{\mathtt{u},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cap V_0(\boldsymbol{F}_{\mathrm{DC},t}). \tag{4.22}$$

But (4.22) follows from the well known fact that BP is exact on acyclic factor graphs (see Fact 2.11). □

Proof of Proposition 2.7. The proposition is an immediate consequence of Lemmas 4.5 and 4.6. □

5. Analysis of the check matrix

In this section we prove Propositions 2.6 and 2.8. Proposition 2.6 is an easy consequence of [8, Theorem 1.1]. Furthermore, Proposition 2.8 follows from Proposition 2.6 by interpolating on the parameter λ; a related argument was recently used in [9] to show that certain random combinatorial matrices have full rank w.h.p. In addition, we prove Corollaries 2.9 and 2.10 and subsequently complete the proofs of Theorems 1.2–1.3.

5.1. **Proof of Proposition 2.6.** We use a general result [8, Theorem 1.1] about the rank of sparse random matrices from a fairly universal class of distributions. The definition of this general random matrix goes as follows. Let $\mathfrak{d}, \mathfrak{k} \ge 0$ be integer-values random variables such that $0 < \mathbb{E}[\mathfrak{d}^3] + \mathbb{E}[\mathfrak{k}^3] < \infty$. Moreover, let $(\mathfrak{d}_i, \mathfrak{k}_i)_{i\ge 0}$ be families of mutually independent random variables such that $\mathfrak{d}_i \overset{\mathrm{dist}}{=} \mathfrak{d}$ and $\mathfrak{k}_i \overset{\mathrm{dist}}{=} \mathfrak{k}$. Let $\bar{\mathfrak{d}} = \mathbb{E}[\mathfrak{d}]$ and $\bar{\mathfrak{k}} = \mathbb{E}[\mathfrak{k}]$ and for an integer $n > 0$ let $\mathfrak{m} = \mathfrak{m}_n \overset{\mathrm{dist}}{=} \mathrm{Po}(\bar{\mathfrak{d}} n / \bar{\mathfrak{k}})$. The sequence $(\mathfrak{m}_n)_n$ is independent of $(\mathfrak{d}_i, \mathfrak{k}_i)_{i\ge 0}$. Further, let $\mathfrak{S}_n$ be the event that

$$\sum_{i=1}^{n} \mathfrak{d}_i = \sum_{i=1}^{\mathfrak{m}_n} \mathfrak{k}_i. \tag{5.1}$$

It is a known fact that $\mathbb{P}[\boldsymbol{S}_n] = \Omega(n^{-1/2})$ [8, Proposition 1.10]. Given that $\mathfrak{S}_n$ occurs, create a simple random bipartite graph $\mathfrak{G}_n$ with a set $\mathfrak{V}_n = \{\mathfrak{x}_1, \dots, \mathfrak{x}_n\}$ of *variable nodes* and a set $\mathfrak{C}_n = \{\mathfrak{c}_1, \dots, \mathfrak{c}_{\mathfrak{m}_n}\}$ of *check nodes* uniformly at random subject to the condition that $\mathfrak{x}_j$ has degree $\mathfrak{d}_j$ and $\mathfrak{c}_i$ has degree $\mathfrak{k}_i$ for all $1 \le j \le n$ and $1 \le i \le \mathfrak{m}_n$. Finally, let $\mathfrak{A}_n$ be the biadjacency matrix of $\mathfrak{G}_n$. Thus, $\mathfrak{A}_n$ has size $\mathfrak{m}_n \times n$ and its (i,j)-entry equals 1 iff $\mathfrak{x}_j$ and $\mathfrak{c}_i$ are adjacent in $\mathfrak{G}_n$.

Theorem 5.1 (special case of [8, Theorem 1.1])**.** *Let $\mathfrak{D}(z) = \sum_{h=0}^{\infty} \mathbb{P}[\mathfrak{d} = h] z^h$ and $\mathfrak{K}(z) = \sum_{h=0}^{\infty} \mathbb{P}[\mathfrak{k} = h] z^h$ be the probability generating functions of $\mathfrak{d}, \mathfrak{k}$, respectively. Furthermore, let*

$$\mathfrak{F} : [0,1] \to \mathbb{R}, \qquad z \mapsto \mathfrak{D}(1 - \mathfrak{K}'(z)/\mathfrak{K}'(1)) - \frac{\mathfrak{D}'(1)}{\mathfrak{K}'(1)}(1 - \mathfrak{K}(z) - (1-z)\mathfrak{K}'(z)). \tag{5.2}$$

Then

$$\lim_{n\to\infty} \frac{1}{n} \operatorname{nul} \mathfrak{A}_n = \max_{z \in [0,1]} \mathfrak{F}(z) \qquad \textit{in probability.}$$

We now derive Proposition 2.6 from Theorem 5.1 by identifying suitable distributions $\mathfrak{d}, \mathfrak{k}$ such that $\mathfrak{A}_n$ resembles $\boldsymbol{A}_t$.

Proof of Proposition 2.6. Recall that $0 \le t = t(n) \le n$ satisfies $t = \theta n + o(n)$ or a fixed $0 \le \theta \le 1$. We continue to set $\lambda = -\log(1-\theta)$. We are going to construct several random matrices that can be coupled such that their nullities differ by no more than $o(n)$ w.h.p. The first of these random matrices is the matrix $\boldsymbol{A}_t$ from Proposition 2.6, and the last is the matrix $\mathfrak{A}_n$ from Theorem 5.1, with suitably chosen $\mathfrak{d}, \mathfrak{k}$.

For a start, consider the check matrix $\boldsymbol{A}' = \boldsymbol{A}_0$ of the original, 'undecimated' k-XORSAT formula $\boldsymbol{F} = \boldsymbol{F}_{\mathrm{DC},0}$. Obtain $\boldsymbol{A}'_t$ from $\boldsymbol{A}'$ by adding t new rows to $\boldsymbol{A}'$. Each of these rows contains precisely a single non-zero entry. The positions of the non-zero entries are chosen uniformly without replacement. Thus, the extra t rows have the effect of fixing t uniformly random coordinates to zero. Since the distribution of the random matrix $\boldsymbol{A}'$ is invariant under column permutations, we conclude that

$$\operatorname{nul} \boldsymbol{A}_t \stackrel{\text{dist}}{=} \operatorname{nul} \boldsymbol{A}'_t. \tag{5.3}$$

Further, let $\boldsymbol{A}[\lambda]$ be the matrix obtained from $\boldsymbol{A}'$ by adding a random number of $\boldsymbol{l} = \mathrm{Po}(\lambda n)$ of rows. Each of these rows contains a single non-zero entry, which is placed in a uniformly random position. The extra rows are chosen mutually independently (thus, 'with replacement') and independently of $\boldsymbol{A}'$. By Poisson thinning, for any column index $j \in [n]$ the probability that one of the new $\boldsymbol{l}$ rows has a non-zero entry in the jth column equals $1-\exp(-\lambda) = \theta$. Since $t \sim \theta n$, the total number of such indices j has distribution $\mathrm{Bin}(n,\theta)$. Since $\mathbb{P}\left[|\mathrm{Bin}(n,\theta) - t| \le \sqrt{n}\log n\right] \ge 1 - 1/n$ by the Chernoff bound, we can couple $\boldsymbol{A}'_t$ and $\boldsymbol{A}[\lambda]$ such that

$$\operatorname{nul} \boldsymbol{A}'_t = \operatorname{nul} \boldsymbol{A}[\lambda] + o(n) \text{ w.h.p.} \tag{5.4}$$

Finally, let $\boldsymbol{A}'[\lambda]$ be the matrix obtained as follows. Let $\mathfrak{d}, \mathfrak{k}$ have probability generating functions

$$\mathfrak{D}(z) = \exp((\lambda + d)(z-1)), \qquad \mathfrak{K}(z) = \frac{dz^k + k\lambda z}{d + k\lambda}. \tag{5.5}$$

In other words, $\mathfrak{d}$ has distribution $\mathrm{Po}(d+\lambda)$ while $\mathfrak{k}$ equals one with probability $k\lambda/(d+k\lambda)$ and equals k with probability $d/(d+k\lambda)$. The definition (5.5) readily yields

$$\bar{\mathfrak{d}} = \mathfrak{D}'(1) = \lambda + d, \qquad \bar{\mathfrak{k}} = \mathfrak{K}'(1) = \frac{k(d+\lambda)}{d+k\lambda}. \tag{5.6}$$

Hence, the number $\mathfrak{m} = \mathfrak{m}_n \stackrel{\text{dist}}{=} \mathrm{Po}(n\bar{\mathfrak{d}}/\bar{\mathfrak{k}})$ of rows of $\mathfrak{A} = \mathfrak{A}_n$ can be written as a sum of independent random variables $\mathfrak{m} = \mathfrak{m}' + \mathfrak{m}''$ with distributions

$$\mathfrak{m}' = \mathrm{Po}(dn/k), \qquad \mathfrak{m}'' = \mathrm{Po}(\lambda n). \tag{5.7}$$

The first summand $\mathfrak{m}'$ prescribes the number of rows of $\mathfrak{A}$ with k non-zero entries, while $\mathfrak{m}''$ details the number of rows with a single non-zero entry. Consequently, (5.7) shows that the numbers of rows with k or with just a single non-zero entry have the same distributions in both $\mathfrak{A}$ and $\boldsymbol{A}[\lambda]$.

We are left to argue that in $\mathfrak{A}$ the positions of the non-zero entries in the different rows are nearly independent and uniform. To see this, let $(\boldsymbol{h}_{i,j})_{1 \le i \ge \mathfrak{m}, 1 \le j \le k}$ be a family of mutually independent and uniform random variables with values in $[n] = \{1, \ldots, n\}$. Moreover, let $\boldsymbol{X}$ be the number of indices $1 \le i \le \mathfrak{m}'$ such that there exist $1 \le j_1 < j_2 \le k$ such that $\boldsymbol{h}_{i,j_1} = \boldsymbol{h}_{i,j_2}$; in other words, $\boldsymbol{h}_{i,1}, \ldots, \boldsymbol{h}_{i,k}$ fail to be pairwise distinct. A routine calculation shows that

$$\mathbb{E}[\boldsymbol{X}] = O(1). \tag{5.8}$$

Now, let us think of $(\boldsymbol{h}_{i,j})_{1 \le i \le \mathfrak{m}', 1 \le j \le k}$ and $(\boldsymbol{h}_{i,1})_{\mathfrak{m}' < i \le \mathfrak{m}}$ as the 'bins' where $k\mathfrak{m}' + \mathfrak{m}''$ randomly tossed 'balls' land. Then the standard Poissonisation of the balls-into-bins experiment shows that given the event (5.1) the loads of the bins are distributed precisely as the vector $(\mathfrak{d}_1, \ldots, \mathfrak{d}_n)$. Therefore, (5.8) shows that $\boldsymbol{A}[\lambda], \mathfrak{A}$ can be coupled such that

$$\operatorname{nul} \boldsymbol{A}[\lambda] = \operatorname{nul} \mathfrak{A} + o(n) \text{ w.h.p.} \tag{5.9}$$

Combining (5.3), (5.4) and (5.9), we see that $\boldsymbol{A}_t$ and $\mathfrak{A}$ can be coupled such that

$$\operatorname{nul} \boldsymbol{A}_t = \operatorname{nul} \mathfrak{A} + o(n) \qquad \text{w.h.p.} \tag{5.10}$$

Hence, Theorem 5.1 implies that

$$\lim_{n \to \infty} \frac{1}{n} \operatorname{nul} \boldsymbol{A}_t = \max_{z \in [0,1]} \mathfrak{F}(z) \qquad \text{in probability.} \tag{5.11}$$

Further, recalling the definitions (5.2), (5.5) of $\mathfrak{F}, \mathfrak{D}, \mathfrak{K}$ and performing a bit of calculus, we verify that $\mathfrak{F}(z)$ coincides with the function $\Phi_{d,k,\lambda}(z)$ from (1.3). Finally, the assertion follows from (5.11) and the fact that $\Phi_{d,k,\lambda}(\alpha_{\max}) = \max_{z\in[0,1]} \Phi_{d,k,\lambda}(z)$. $\square$

5.2. **Proof of Proposition 2.8.** We continue to work with the random matrix $\boldsymbol{A}[\lambda]$ from the above proof of Proposition 2.6. As we recall, this matrix is obtained by adding $\boldsymbol{l} = \mathrm{Po}(\lambda n)$ stochastically independent new rows to the matrix $A(\boldsymbol{F})$ that each contain a single non-zero entry in a uniformly random position. Combining (5.3)–(5.4), we see that

$$|\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]] - \mathbb{E}[\mathrm{nul}\,\boldsymbol{A}_t]| = o(n) \qquad \text{for } \lambda = \log(1-\theta). \tag{5.12}$$

Towards the proof of Proposition 2.8 we observe that $\mathrm{nul}\,\boldsymbol{A}[\lambda], \mathrm{nul}\,\boldsymbol{A}_t$ concentrate about their expectations.

Lemma 5.2. *We have*

$$\mathbb{P}\left[|\,\mathrm{nul}\,\boldsymbol{A}_t - \mathbb{E}[\mathrm{nul}\,\boldsymbol{A}_t]| > \sqrt{n}\log n\right] = o(n^{-10}), \qquad \mathbb{P}\left[|\,\mathrm{nul}\,\boldsymbol{A}[\lambda] - \mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]]| > \sqrt{n}\log n\right] = o(n^{-10}). \tag{5.13}$$

Proof. We combine the Azuma–Hoeffding inequality with the simple observation that the nullity satisfies a Lipschitz condition. Specifically, adding or removing a single row to a matrix changes the nullity by at most one. We apply this observation to the matrix $\boldsymbol{A}'_t$ from the proof of Proposition 2.6, which consists of $\boldsymbol{m}+t$ independent random rows. Indeed, Azuma-Hoeffding implies together with the Lipschitz property that

$$\mathbb{P}\left[|\boldsymbol{A}'_t - \mathbb{E}[\boldsymbol{A}'_t \mid \boldsymbol{m}]| > u \mid \boldsymbol{m}\right] \le 2\exp\left(-\frac{u^2}{2(\boldsymbol{m}+t)}\right) \qquad \text{for any } u>0. \tag{5.14}$$

Furthermore, Bennett's concentration inequality for Poisson variables shows that

$$\mathbb{P}\left[|\boldsymbol{m} - dn/k| > \sqrt{n}\log^{2/3} n\right] = o(n^{-10}). \tag{5.15}$$

Combining (5.14)–(5.15) with the Lipschitz property and setting $u = \sqrt{n}\log^{2/3} n$, we obtain the first part of (5.13).

Similar reasoning applies to the second matrix $\boldsymbol{A}[\lambda]$; for given $\boldsymbol{l}$ and $\boldsymbol{m}$ the Lipschitz property yields

$$\mathbb{P}\left[|\boldsymbol{A}'_t - \mathbb{E}[\boldsymbol{A}'_t \mid \boldsymbol{l},\boldsymbol{m}]| > u \mid \boldsymbol{l},\boldsymbol{m}\right] \le 2\exp\left(-\frac{u^2}{2(\boldsymbol{l}+\boldsymbol{m})}\right) \qquad \text{for any } u>0. \tag{5.16}$$

Moreover, in analogy to (5.15) we have

$$\mathbb{P}\left[|\boldsymbol{l} - \lambda n| > \sqrt{n}\log^{2/3} n\right] = o(n^{-10}). \tag{5.17}$$

Thus, (5.15)–(5.17) and Azuma-Hoeffding imply the second part of (5.13). $\square$

We are going to estimate $|V_0(\boldsymbol{F}_{\mathrm{DC},t})|$ by way of estimating changes of $\mathrm{nul}\,\boldsymbol{A}[\lambda]$ as λ varies. Since $\mathrm{nul}\,\boldsymbol{A}[\lambda]/n$ converges to $\Phi_{d,k,\lambda}(\alpha_{\max})$ by Proposition 2.6, we thus need to estimate the derivative $\frac{\partial}{\partial\lambda}\Phi_{d,k,\lambda}(\alpha_{\max})$.

Lemma 5.3. *Let $d>0$ and assume that*

(i) $d < d_{\min}$, or

(ii) $d > d_{\min}$ and $\lambda \in (0,\infty)\setminus\{\lambda_{\mathrm{cond}}\}$.

Then

$$\frac{\partial}{\partial\lambda}\Phi_{d,k,\lambda}(\alpha_{\max}) = \alpha_{\max} - 1. \tag{5.18}$$

Proof. The seeming difficulty is that $\alpha_{\max} = \alpha_{\max}(\lambda)$ varies with λ. Yet Proposition 2.2 (iii) ensures that the function $\lambda \mapsto \alpha_{\max}$ is continuously differentiable for $\lambda \neq \lambda_{\mathrm{cond}}$. Moreover, Fact 2.1 shows that $\alpha_{\max}$ is a local maximum of $\Phi_{d,k,\lambda}$. Hence, applying the chain rule we obtain

$$\frac{\partial}{\partial\lambda}\Phi_{d,k,\lambda}(\alpha_{\max}) = \frac{\partial\Phi_{d,k,\lambda}}{\partial\lambda}\Big|_{\lambda,\alpha_{\max}} + \frac{\partial\Phi_{d,k,\lambda}}{\partial\alpha}\Big|_{\lambda,\alpha_{\max}}\frac{\partial\alpha_{\max}}{\partial\lambda} = \frac{\partial\Phi_{d,k,\lambda}}{\partial\lambda}\Big|_{\lambda,\alpha_{\max}} = -\exp\left(-\lambda - d\alpha_{\max}^{k-1}\right). \tag{5.19}$$

In fact, since Fact 2.1 shows that $\alpha_{\max}$ is a fixed point of $\phi_{d,k,\lambda}$, the r.h.s. of (5.19) simplifies to (5.18). $\square$

Complementing the analytic formula (5.18), we now derive a combinatorial interpretation of the derivative of the nullity. For a matrix A of size $m\times n$ let $V_0(A)$ be the set of all indices $i\in[n]$ such that $\sigma_i = 0$ for all $\sigma\in\ker A$.

Lemma 5.4. *For any $d,\lambda>0$ we have*

$$\frac{\partial}{\partial\lambda}\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]] = \frac{\mathbb{E}|V_0(\boldsymbol{A}[\lambda])|}{n} - 1.$$

Proof. Recall that $\boldsymbol{A}[\lambda]$ is obtained from $A(\boldsymbol{F})$ by adding $\boldsymbol{m}''\overset{\text{dist}}{=}\mathrm{Po}(\lambda n)$ stochastically independent rows with a single non-zero entry in a uniformly random position. Consequently,

$$\begin{aligned}\frac{\partial}{\partial\lambda}\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]] &= \frac{\partial}{\partial\lambda}\sum_{\ell=0}^{\infty}\mathbb{P}\left[\boldsymbol{m}''=\ell\right]\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]\mid \boldsymbol{m}''=\ell] = \sum_{\ell=0}^{\infty}\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]\mid \boldsymbol{m}''=\ell]\frac{\partial}{\partial\lambda}\frac{(\lambda n)^{\ell}}{\ell!}\exp(-\lambda n)\\ &= \sum_{\ell=0}^{\infty}\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]\mid \boldsymbol{m}''=\ell]\left(\mathbb{1}\{\ell\geq 1\}\frac{(\lambda n)^{\ell-1}}{(\ell-1)!}-\frac{(\lambda n)^{\ell}}{\ell!}\right)\exp(-\lambda n)\\ &= \sum_{\ell=0}^{\infty}\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]\mid \boldsymbol{m}''=\ell]\left(\mathbb{P}\left[\boldsymbol{m}''=\ell\right]-\mathbb{P}\left[\boldsymbol{m}''=\ell+1\right]\right).\end{aligned}\tag{5.20}$$

Hence, obtain $\boldsymbol{A}[\lambda]^+$ from $\boldsymbol{A}[\lambda]$ by adding one more row with a single non-zero entry in a uniformly random position $\boldsymbol{j}^+\in[n]$. Then $\boldsymbol{A}[\lambda]^+-\boldsymbol{A}[\lambda]=-\mathbb{1}\{\boldsymbol{j}^+\in V_0(A[\lambda])\}$. Hence, (5.20) yields

$$\frac{\partial}{\partial\lambda}\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]] = -\mathbb{E}\left[\mathrm{nul}(\boldsymbol{A}[\lambda]^+)-\mathrm{nul}(\boldsymbol{A}[\lambda])\right]=\mathbb{P}\left[\boldsymbol{j}^+\in V_0(\boldsymbol{A}[\lambda])\right]-1=\frac{\mathbb{E}|V_0(\boldsymbol{A}[\lambda])|}{n}-1,$$

as claimed. □

With these preparations in place we can now derive the desired formulas for $|V_0(\boldsymbol{A}_t)|$. We treat the cases $\alpha_{\max}=\alpha_*$ and $\alpha_{\max}=\alpha^*$ separately.

Lemma 5.5. *Assume that $d,\lambda>0$ satisfy*

$$\Phi_{d,k,\lambda}(\alpha_*)>\Phi_{d,k,\lambda}(\alpha)\quad\textit{for all }\alpha\in[0,1]\setminus\{\alpha_*\}.\tag{5.21}$$

Then $|V_0(\boldsymbol{A}[\lambda])|=\alpha_*n+o(n)$ w.h.p.

Proof. Proposition 2.5 and Lemma 4.5 yield the lower bound

$$|V_0(\boldsymbol{A}[\lambda])|\geq\alpha_*n+o(n)\qquad\text{w.h.p.}\tag{5.22}$$

To derive the matching upper bound, fix $\varepsilon>0$ and assume that the event $\mathscr{E}=\{|V_0(\boldsymbol{A}[\lambda])|>(\alpha_*+\varepsilon)n\}$ has probability $\mathbb{P}[\mathscr{E}]>\varepsilon$. Then by Proposition 2.2 (iii) there exists $\lambda'>\lambda$ such that $\alpha_{\max}(\lambda'')=\alpha_*(\lambda'')$ and $\alpha_*(\lambda'')<\alpha_*(\lambda)+\varepsilon^2/2$ for all $\lambda''\in[\lambda,\lambda']$. Hence, Lemmas 5.3 yields

$$\Phi_{d,k,\lambda'}(\alpha_{\max}(\lambda'))-\Phi_{d,k,\lambda}(\alpha_{\max}(\lambda))\leq\int_{\lambda}^{\lambda'}(\alpha_*(\lambda'')-1)\,\mathrm{d}\lambda''\leq(\lambda'-\lambda)(\alpha_*(\lambda)-\varepsilon^2/2/-1).\tag{5.23}$$

Combining (5.23) with Proposition 2.6 and Lemma 5.2, we obtain

$$n^{-1}\left[\mathbb{E}\left[\mathrm{nul}\,\boldsymbol{A}[\lambda']\right]-\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]]\right]\leq(\lambda'-\lambda)(\alpha_*(\lambda)-\varepsilon^2/2/-1+o(1)).\tag{5.24}$$

On the other hand, since adding checks can only increase the number of frozen variables, Lemma 5.4 shows that

$$n^{-1}\left[\mathbb{E}\left[\mathrm{nul}\,\boldsymbol{A}[\lambda']\right]-\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]]\right]\geq(\lambda'-\lambda)(\alpha_*(\lambda)+\mathbb{P}[\mathscr{E}]\,\varepsilon-1+o(1))\geq(\lambda'-\lambda)(\alpha_*(\lambda)+\varepsilon^2-1+o(1)).\tag{5.25}$$

Finally, since (5.24) and (5.25) contradict each other, we have refuted the assumption $\mathbb{P}[\mathscr{E}]>\varepsilon$. □

Lemma 5.6. *Assume that $d,\lambda>0$ are such that*

$$\Phi_{d,k,\lambda}(\alpha^*)>\Phi_{d,k,\lambda}(\alpha)\quad\textit{for all }\alpha\in[0,1]\setminus\{\alpha^*\}.\tag{5.26}$$

Then $|V_0(\boldsymbol{A}[\lambda])|=\alpha^*n+o(n)$ w.h.p.

Proof. We use a similar strategy as in the proof of Lemma 5.5. Hence, assume that $d,\lambda>0$ satisfy (5.26). Combining Proposition 2.5 and Lemma 4.6, we see that $|V_0(\boldsymbol{A}[\lambda])|\leq\alpha^*n+o(n)$ w.h.p. Now choose a small enough $\varepsilon>0$ and assume that $\mathscr{E}=\{|V_0(\boldsymbol{A}[\lambda])|<(\alpha^*-\varepsilon)n\}$ occurs with probability $\mathbb{P}[\mathscr{E}]>\varepsilon$. Then Proposition 2.2 shows that there exists $\lambda'<\lambda$ such that $\alpha_{\max}(\lambda'')=\alpha^*(\lambda'')$ and $\alpha^*(\lambda'')>\alpha^*(\lambda)-\varepsilon^2/2$ for all $\lambda''\in[\lambda,\lambda']$. Hence, Lemmas 5.3 yields

$$\Phi_{d,k,\lambda}(\alpha_{\max}(\lambda))-\Phi_{d,k,\lambda'}(\alpha_{\max}(\lambda'))=\int_{\lambda}^{\lambda'}(\alpha^*(\lambda'')-1)\,\mathrm{d}\lambda''\geq(\lambda'-\lambda)(\alpha^*(\lambda)-\varepsilon^2/2/-1).\tag{5.27}$$

But once again because adding checks can only increase the number of frozen variables, Lemma 5.4 yields

$$n^{-1}\left[\mathbb{E}[\mathrm{nul}\,\boldsymbol{A}[\lambda]]-\mathbb{E}\left[\mathrm{nul}\,\boldsymbol{A}\left[\lambda'\right]\right]\right]\leq(\lambda'-\lambda)(\alpha^*(\lambda)-\mathbb{P}[\mathscr{E}]\,\varepsilon-1+o(1))\leq(\lambda'-\lambda)(\alpha_*(\lambda)-\varepsilon^2-1+o(1)).\tag{5.28}$$

However, Proposition 2.6 and Lemma 5.3 show that (5.27)–(5.28) are in contradiction. □

Proof of Proposition 2.8. Since $\alpha_{\max}\in\{\alpha_*,\alpha^*\}$, the assertion is an immediate consequence of Lemmas 5.5–5.6. □

5.3. **Proof of Corollary 2.9.** There are four cases to consider separately. Let $\varepsilon > 0$.

Case 1: $d < d_{\min}$**:** As Proposition 2.2 (i) shows, in this case we have $\alpha_* = \alpha^*$ for all $\lambda > 0$; thus, the function $\phi_{d,k,\lambda}$ has only the single fixed point α_*, which is stable. Furthermore, Proposition 2.5 shows that $||V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| - \alpha_* n| < \varepsilon n/2$ for large enough ℓ w.h.p. Moreover, Proposition 2.8 yields $|V_0(\boldsymbol{F}_{\mathrm{DC},t})| = \alpha_* n + o(n)$ w.h.p. Therefore, Proposition 2.7 implies that $|V_0(\boldsymbol{F}_{\mathrm{DC},t}) \triangle V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| < \varepsilon n$ w.h.p. for large enough ℓ. Since $|V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| \subseteq V_0(\boldsymbol{F}_{\mathrm{DC},t})$ w.h.p. and $|V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \triangle V_{\mathrm{n}}(\boldsymbol{F}_{\mathrm{DC},t})| < \varepsilon n$ by (2.19), the assertion follows.

Case 2: $d_{\min} < d < d_{\mathrm{sat}}$ and $\theta > \theta_*$: A similar argument as under Case 1 applies. Indeed, Proposition 2.2 (ii) shows that $\alpha_* = \alpha^*$ is the unique and stable fixed point of $\phi_{d,k,\lambda}$. Since $||V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| - \alpha_* n| < \varepsilon n/2$ for large ℓ w.h.p. by Proposition 2.5 and $|V_0(\boldsymbol{F}_{\mathrm{DC},t})| = \alpha_* n + o(n)$ w.h.p. by Proposition 2.8, Proposition 2.7 yields $|V_0(\boldsymbol{F}_{\mathrm{DC},t}) \triangle V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| < \varepsilon n$ w.h.p. Therefore, (2.19) implies the assertion.

Case 3: $d_{\min} < d < d_{\mathrm{sat}}$ and $\theta < \theta_{\mathrm{cond}}$: Proposition 2.2 (ii) shows that $\alpha_* < \alpha^*$ in this case. Moreover, Proposition 2.5 yields $||V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| - \alpha_* n| < \varepsilon n/2$ for large ℓ w.h.p., while Proposition 2.8 and Proposition 2.2 (iii) imply that $|V_0(\boldsymbol{F}_{\mathrm{DC},t})| = \alpha_* n + o(n)$ w.h.p. Thus, the same steps as in Cases 1–2 complete the proof.

Case 4: $d_{\min} < d < d_{\mathrm{sat}}$ and $\theta_{\mathrm{cond}} < \theta < \theta_*$: Once again Proposition 2.2 (ii) shows that $\alpha_* < \alpha^*$, Proposition 2.5 yields $||V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| - \alpha_* n| < \varepsilon n/2$ for large ℓ w.h.p., and Proposition 2.8 and Proposition 2.2 (iii) show that $|V_0(\boldsymbol{F}_{\mathrm{DC},t})| = \alpha^* n + o(n)$ w.h.p. Since $|V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| \subseteq V_0(\boldsymbol{F}_{\mathrm{DC},t})$ w.h.p., the assertion follows from (2.19) and the fact that $\alpha_* < \alpha^*$.

5.4. **Proof of Corollary 2.10.** Assume first that $\theta < \theta_{\mathrm{cond}}$. Then Corollary 2.9 shows that $|V_0(\boldsymbol{F}_{\mathrm{DC},t}) \triangle V_{\mathrm{n}}(\boldsymbol{F}_{\mathrm{DC},t}) = o(n)$ for large enough ℓ. Since $V_{\mathrm{n}}(\boldsymbol{F}_{\mathrm{DC},t}) \cap V_{\mathtt{f}}(\boldsymbol{F}_{\mathrm{DC},t}) = \emptyset$ by construction, the first assertion follows.

Now suppose $\theta > \theta_{\mathrm{cond}}$. Then Proposition 2.5 yields $||V_{\mathtt{f},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| - \alpha^* n| < \varepsilon n/2$ for large ℓ w.h.p., while Proposition 2.8 and Proposition 2.2 (iii) show that $|V_0(\boldsymbol{F}_{\mathrm{DC},t})| = \alpha^* n + o(n)$ w.h.p. Additionally, Proposition 2.5 shows that $|V_{\mathrm{u},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cap V_0(\boldsymbol{F}_{\mathrm{DC},t})| < \varepsilon n$ for large ℓ, which implies the assertion.

5.5. **Proofs of Theorems 1.2 and 1.3.** We begin with the following observation.

Lemma 5.7. *Let $\boldsymbol{\sigma} \in \ker(\boldsymbol{F}_{\mathrm{DC},t})$ be uniformly random. For any $\ell > 0$ w.h.p. we have*

$$\mathbb{P}\left[\boldsymbol{\sigma}_{x_{t+1}} = 0 \mid \boldsymbol{F}_{\mathrm{DC},t}, \boldsymbol{\sigma}_{\partial^{2\ell} x_{t+1}}\right] = \frac{1}{2}\left(1 + \mathbb{1}\{x_{t+1} \in V_{\mathtt{f},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cup V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})\}\right), \tag{5.29}$$

$$\pi_{\boldsymbol{F}_{\mathrm{DC},t}} = \mathbb{P}\left[\boldsymbol{\sigma}_{x_{t+1}} = 0 \mid \boldsymbol{F}_{\mathrm{DC},t}\right] = \frac{1}{2}\left(1 + \mathbb{1}\{x_{t+1} \in V_0(\boldsymbol{F}_{\mathrm{DC},t})\}\right). \tag{5.30}$$

Proof. Notice that for $d < d_{\mathrm{sat}}$ the random XORSAT instance $\boldsymbol{F}$ is satisfiable w.h.p.; therefore, so is $\boldsymbol{F}_{\mathrm{DC},t}$.

We begin with the proof of (5.30). The first equality $\pi_{\boldsymbol{F}_{\mathrm{DC},t}} = \mathbb{P}\left[\boldsymbol{\sigma}_{x_{t+1}} = 0 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]$ follows from the fact that the set of solutions of $\boldsymbol{F}_{\mathrm{DC},t}$ is an affine translation of $\ker(A(\boldsymbol{F}_{\mathrm{DC},t}))$. Moreover, the second equality sign follows from the well known fact that the marginal $\mathbb{P}\left[\boldsymbol{\sigma}_{x_{t+1}} = 0 \mid \boldsymbol{F}_{\mathrm{DC},t}\right]$ is equal to $1/2$ or to 1.

Moving on to (5.29), we recall from Lemma 4.1 that the depth-2ℓ neighbourhood $\partial^{\le \ell} x_{t+1}$ of x_{t+1} in $\boldsymbol{F}_{\mathrm{DC},t}$ is acyclic w.h.p. Furthermore, we can think of $\mathbb{P}\left[\boldsymbol{\sigma}_{x_{t+1}} = 0 \mid \boldsymbol{F}_{\mathrm{DC},t}, \boldsymbol{\sigma}_{\partial^{\le \ell} x_{t+1}}\right]$ as the marginal probability that x_{t+1} receives the value zero under a random vector from the kernel of the check matrix of $\partial^{\le \ell} x_{t+1}$, subject to imposing the values $\boldsymbol{\sigma}_{\partial^{\le \ell} x_{t+1}}$ upon the variable at distance exactly 2ℓ from x_{t+1}. Let $\boldsymbol{F}^{(\ell)}_{\mathrm{DC},t}$ signify the XORSAT instance thus obtained. Then we conclude that $\mathbb{P}\left[\boldsymbol{\sigma}_{x_{t+1}} = 0 \mid \boldsymbol{F}_{\mathrm{DC},t}, \boldsymbol{\sigma}_{\partial^{\le \ell} x_{t+1}}\right] = 1$ iff $x_{t+1} \in V_0(\boldsymbol{F}^{(\ell)}_{\mathrm{DC},t})$. Furthermore, because BP is exact on acyclic factor graphs, we have $x_{t+1} \in V_0(\boldsymbol{F}^{(\ell)}_{\mathrm{DC},t})$ iff $x_{t+1} \in V_{V_0,\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cup V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})$. Thus, we obtain (5.29). $\square$

Proof of Theorem 1.2. We begin with claim (i) concerning $d < d_{\min}$. As Proposition 2.2 (i) shows, in this case we have $\alpha_* = \alpha^*$. Furthermore, Proposition 2.5 shows that $||V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t})| - \alpha_* n| < \varepsilon n$ and $|V_{V_0,\ell}(\boldsymbol{F}_{\mathrm{DC},t})| < \varepsilon n$ for large enough ℓ w.h.p. Moreover, Proposition 2.8 yields $|V_0(\boldsymbol{F}_{\mathrm{DC},t})| = \alpha_* n + o(n)$ w.h.p. Therefore, Proposition 2.7 implies that $|V_0(\boldsymbol{F}_{\mathrm{DC},t}) \triangle V_{\mathrm{n},\ell}| < \varepsilon n$ w.h.p. for large enough ℓ. Hence, Lemma 5.7 shows that the non-reconstruction property (1.7) holds w.h.p.

Similarly, towards the proof of (ii) assume that $d_{\min} < d < d_{\mathrm{sat}}$ and $\theta < \theta^*$. Then Proposition 2.2 (ii) shows that $\alpha_* = \alpha^*$ is the unique (stable) fixed point of $\phi_{d,k,\lambda}$. Therefore, the argument from the previous paragraph shows that (1.7) holds w.h.p.Further, suppose that $d_{\min} < d < d_{\mathrm{sat}}$ and $\theta > \theta_{\mathrm{cond}}$. Then Corollary 2.10 (ii) shows that $|(V_{\mathrm{n},\ell}(\boldsymbol{F}_{\mathrm{DC},t}) \cup V_{\mathtt{f},\ell}(\boldsymbol{F}_{\mathrm{DC},t})) \triangle V_{0,\ell}(\boldsymbol{F}_{\mathrm{DC},t})| < \varepsilon n$ w.h.p. Therefore, Lemma 5.7 implies non-reconstruction property, and thus the proof of (ii) is complete.

Finally, suppose that $d_{\min} < d < d_{\text{sat}}$ and $\theta^* < \theta < \theta_{\text{cond}}$. Then Proposition 2.5 shows that $||V_{\mathtt{n},\ell}(\boldsymbol{F}_{\text{DC},t}) - \alpha_* n| < \varepsilon n$ and $|V_{V_0,\mathtt{n}}| - (\alpha^* - \alpha_*)n| < \varepsilon n$ for large enough ℓ w.h.p. Moreover, Corollary 2.10 shows that $|V_{\mathtt{f},\mathtt{n}} \cap V_0(\boldsymbol{F}_{\text{DC},t})| < \varepsilon n$ w.h.p. Consequently, Lemma 5.7 demonstrates that the reconstruction condition (1.8) holds w.h.p. □

Proof of Theorem 1.3. Part (i) regarding the case $d < d_{\min}$ is an immediate consequence of Fact 2.4 (the equivalence of WP and BP), Corollary 2.9 (i) and Lemma 5.7. The same is true of part (ii) concerning $d_{\min} < d < d_{\text{sat}}$ and $\theta < \theta_{\text{cond}}$ or $\theta > \theta_*$. Furthermore, (iii) follows from Corollary 2.9 (ii) and Lemma 5.7. □

6. Belief Propagation Guided Decimation

In this section we prove Theorem 1.1. We begin by arguing that `BPGD` is actually equivalent to the simple combinatorial Unit Clause Propagation algorithm. Then we prove the 'positive' part, i.e., the formula (1.6) for the success probability for $d < d_{\min}$. Subsequently we prove the second part of the theorem concerning $d_{\min} < d < d_{\text{sat}}$.

6.1. **Unit Clause Propagation redux.** The simple-minded Unit Clause Propagation algorithm attempts to assign random values to as yet unassigned variables one after the other. After each such random assignment the algorithm pursues the 'obvious' implications of its decisions. Specifically, the algorithm substitutes its chosen truth values for all occurrences of the already assigned variables. If this leaves a clause with only a single unassigned variable, a so-called 'unit clause', the algorithm assigns that variable so as to satisfy the unit clause. If a conflict occurs because two unit clauses impose opposing values on a variable, the algorithm declares that a conflict has occurred, sets the variable to false and continues; of course, in the event of a conflict the algorithm will ultimately fail to produce a satisfying assignment. The pseudocode for the algorithm is displayed in Algorithm 3.

1 Let $U = \emptyset$ and let $\boldsymbol{\sigma}_{\text{UC}} : U \to \{0,1\}$ be the empty assignment;
2 **for** $t = 0, \dots, n-1$ **do**
3   **if** $x_{t+1} \notin U$ **then**
4   add x_{t+1} to U;
5   choose $\boldsymbol{\sigma}_{\text{UC}}(x_{t+1}) \in \{0,1\}$ uniformly at random;
6   **while** $\boldsymbol{F}[\boldsymbol{\sigma}_{\text{UC}}]$ *contains a unit clause* a **do**
7   let x be the variable in a;
8   let $s \in \{0,1\}$ be the truth value that x needs to take to satisfy a;
9   **if** *another unit clause* a' *exists that requires* x *be set to* $1-s$ **then**
10   output 'conflict' and let $\boldsymbol{\sigma}_{\text{UC}}(x) = 0$;
11   **else**
12   add x to U and let $\boldsymbol{\sigma}_{\text{UC}}(x) = s$;
13 **return** $\boldsymbol{\sigma}_{\text{UC}}$;

Algorithm 3: The `UCP` algorithm.

Let $\boldsymbol{F}_{\text{UC},t}$ denote the simplified formula obtained after the first t iterations (in which the truth values chosen for $x_1, \dots, x_t$ and any values implied by Unit Clauses have been substituted). We notice that the values assigned during Steps 6–12 are deterministic consequences of the choices in Step 5. In particular, the order in which unit clauses are processed Steps 6–12 does not affect the output of the algorithm.

Proposition 6.1. *We have*

$$\mathbb{P}\left[\texttt{BPGD} \textit{ outputs a satisfying assignment of } \boldsymbol{F}\right] = \mathbb{P}\left[\texttt{UCP} \textit{ outputs a satisfying assignment of } \boldsymbol{F}\right].$$

Proof. We employ the following coupling. Let $\boldsymbol{\tau} \in \{0,1\}^n$ be a uniformly random vector. The `BPGD` algorithm sets $\boldsymbol{\sigma}_{\text{BP}}(x_{t+1}) = \boldsymbol{\tau}_{t+1}$ if $\mu_{\boldsymbol{F}_{\text{BP},t}} = 1/2$. Analogously, `UCP` sets $\boldsymbol{\sigma}_{\text{UC}}(x_{t+1}) = \boldsymbol{\tau}_{t+1}$ in Step 5 (if $x_{t+1} \notin U$). Hence, because (1.1) guarantees that the BP marginals $\mu_{\boldsymbol{F}_{\text{BP},t}}$ are half-integral, the coupling ensures that the "free steps" of the two algorithms pick the same truth values.

We now proceed by induction on $0 \le t \le n$ to prove the following two statements.

UCP1: unless `UCP` encountered a conflict before time t we have $\boldsymbol{\sigma}_{\text{BP}}(x_i) = \boldsymbol{\sigma}_{\text{UC}}(x_i)$ for $i = 1, \dots, t$.
UCP2: if $t < n$ and there has been no conflict before time t we have we have $\mu_{\boldsymbol{F}_{\text{BP},t+1}} = 1/2$ iff $x_{t+1} \notin U$.

For $t = 0$ both of these statements are clearly correct because $\mu_{\boldsymbol{F}_{\mathrm{BP},0}} = 1/2$ and $x_1 \not\in U$.

Now assume that **UCP1**–**UCP2** hold at time $t-1$ and that no conflict has occurred yet. Then we already know that $\boldsymbol{\sigma}_{\mathrm{BP}}(x_i) = \boldsymbol{\sigma}_{\mathrm{UC}}(x_i)$ for $i = 1, \ldots, t-1$. Furthermore, since **UCP2** is correct at time $t-1$ we have $\mu_{\boldsymbol{F}_{\mathrm{BP},t}} = 1/2$ iff $x_t \not\in U$. Consequently, if $x_t \not\in U$ then $\boldsymbol{\sigma}_{\mathrm{BP}}(x_t) = \boldsymbol{\sigma}_{\mathrm{UC}}(x_t)$. Hence, suppose that $x_t \not\in U$ and thus $\mu_{\boldsymbol{F}_{\mathrm{BP},t}} \in \{0,1\}$. Then given $\boldsymbol{\sigma}_{\mathrm{BP}}(x_1) = \boldsymbol{\sigma}_{\mathrm{UC}}(x_1), \ldots, \boldsymbol{\sigma}_{\mathrm{BP}}(x_{t-1}) = \boldsymbol{\sigma}_{\mathrm{UC}}(x_{t-1})$ the value $\boldsymbol{\sigma}_{\mathrm{UC}}(x_t)$ is implied by unit clause propagation. But a glimpse at the BP update rules (2.7)–(2.8) shows that these encompass the unit clause rule. Specifically, if x is the only remaining variable in clause a, then (2.7) ensures that the message from a to x gives probability one to the value that satisfies clause a. Therefore, the definition (2.9) of the BP marginal demonstrates that $\mu_{\boldsymbol{F}_{\mathrm{BP},t}} = \boldsymbol{\sigma}_{\mathrm{UC}}(x_1)$ and thus $\boldsymbol{\sigma}_{\mathrm{BP}}(x_t) = \boldsymbol{\sigma}_{\mathrm{UC}}(x_t)$. Thus, **UCP1** continues to hold for t.

Similar reasoning yields **UCP2**. Indeed, revisiting (2.7), we see that the BP message that clause a sends to variable x equals $1/2$ unless a is a unit clause. In effect, (2.9) shows that the BP marginal $\mu_{\boldsymbol{F}_{\mathrm{BP},t+1}}$ is equal to $1/2$ unless the value of x_{t+1} is implied by the unit clause rule. This completes the induction.

To complete the proof assume that UCP manages to find a satisfying assignment. Then **UCP1** applied to $t = n$ demonstrates that BPGD outputs the very same satisfying assignment. Conversely, if UCP encounters a conflict at some time t, then **UCP1** shows that BPGD chose the same assignment up to time t. Therefore, it is not possible to extend the partial assignment $\boldsymbol{\sigma}_{\mathrm{BP}}(x_1), \ldots, \boldsymbol{\sigma}_{\mathrm{BP}}(x_t)$ to a satisfying assignment of $\boldsymbol{F}$ and thus BPGD will ultimately fail to output a satisfying assignment. □

In light of Proposition 6.1 we are left to study the success probability of UCP. The following two subsections deal with this task for $d < d_{\min}$ and $d > d_{\min}$, respectively.

6.2. **The success probability of** UCP **for** $d < d_{\min}$**.** We continue to denote by $\boldsymbol{F}_{\mathrm{UC},t}$ the sub-formula obtained after the first t iterations of UCP. Let $\boldsymbol{V}(t) \subseteq \{x_{t+1}, \ldots, x_n\}$ be the set of variables of $\boldsymbol{F}_{\mathrm{UC},t}$. Thus, $\boldsymbol{V}(t)$ contains those variables among $x_{t+1}, \ldots, x_n$ whose values are not implied by the assignment of $x_1, \ldots, x_t$ via unit clauses. Also let $\boldsymbol{C}(t)$ be the set of clauses of $\boldsymbol{F}_{\mathrm{UC},t}$; these clauses contain variables from $\boldsymbol{V}(t)$ only, and each clause contains at least two variables. Let $\bar{\boldsymbol{V}}(t) = V_n \setminus \boldsymbol{V}(t)$ be the set of assigned variables. Thus, after its first t iterations UCP has constructed an assignment $\boldsymbol{\sigma}_{\mathrm{UC}} : \bar{\boldsymbol{V}}(t) \to \{0,1\}$. Moreover, let $\boldsymbol{V}'(t+1) = \boldsymbol{V}(t) \setminus \boldsymbol{V}(t+1)$ be the set of variables that receive values in the course of the iteration $t+1$ for $0 \le t < n$. Additionally, let $\boldsymbol{C}'(t+1)$ be the set of clauses of $\boldsymbol{F}_{\mathrm{UC},t}$ that consists of variables from $\boldsymbol{V}'(t+1)$ only. Finally, let $\boldsymbol{F}'_{\mathrm{UC},t+1}$ be the formula comprising the variables $\boldsymbol{V}'(t+1)$ and the clauses $\boldsymbol{C}'(t+1)$.

To characterise the distribution of $\boldsymbol{F}_{\mathrm{UC},t}$ let $\boldsymbol{n}(t) = |\boldsymbol{V}(t)|$ and let $\boldsymbol{m}_\ell(t)$ be the number of clauses of length ℓ, i.e., clauses that contain precisely ℓ variables from $\boldsymbol{V}(t)$. Observe that $\boldsymbol{m}_1(t) = 0$ because unit clauses get eliminated. Let $\mathfrak{F}_t$ be the σ-algebra generated by $\boldsymbol{n}(t)$ and $(\boldsymbol{m}_\ell(t))_{2 \le \ell \le k}$.

Fact 6.2. *The XORSAT formula $\boldsymbol{F}_{\mathrm{UC},t}$ is uniformly random given $\mathfrak{F}_t$. In other words, the variables that appear in each clause are uniformly random and independent, as are their signs.*

Proof. This follows from the principle of deferred decisions. □

We proceed to estimate the random variables $\boldsymbol{n}(t), \boldsymbol{m}_\ell(t)$. Let $\boldsymbol{\alpha}(t) = |\bar{\boldsymbol{V}}(t)|/n$ so that $\boldsymbol{n}(t) = n(1 - \boldsymbol{\alpha}(t))$. Let $\lambda = \lambda(\theta) = -\log(1-\theta)$ with $\theta \sim t/n$ and recall that $\alpha_* = \alpha_*(d,k,\lambda)$ denotes the smallest fixed point of $\phi_{d,k,\lambda}$. The proof of the following proposition proof can be found in Section 6.2.1.

Proposition 6.3. *Suppose that $d < d_{\min}(k)$. There exists a function $\delta = \delta(n) = o(1)$ such that for all $0 \le t < n$ and all $2 \le \ell \le k$ we have*

$$\mathbb{P}\left[|\boldsymbol{\alpha}(t) - \alpha_*| > \delta\right] = O(n^{-2}), \qquad \mathbb{P}\left[\left|\boldsymbol{m}_\ell(t) - \frac{dn}{k}\binom{k}{\ell}(1-\alpha_*)^\ell \alpha_*^{k-\ell}\right| > \delta n\right] = O(n^{-2}). \tag{6.1}$$

Proposition 6.3 paves the way for the actual computation of the success probability of UCP. Let $\mathscr{R}_t$ be the event that a conflict occurs in iteration t. The following proposition gives us the correct value of $\mathbb{P}[\mathscr{R}_t \mid \mathfrak{F}_t]$ w.h.p. Since $\mathfrak{F}_t$ is a random variable the value for the probability $\mathbb{P}[\mathscr{R}_t \mid \mathfrak{F}_t]$ is random as well.

Proposition 6.4. *Fix $\varepsilon > 0$, let $0 \le t < (1-\varepsilon)n$ and define*

$$f_n(t) = d(k-1)(1-\alpha_*)\alpha_*^{k-2}. \tag{6.2}$$

Then with probability $1-o(1/n)$ *we have*

$$\mathbb{P}[\mathscr{R}_t \mid \mathfrak{F}_t] = \frac{f_n(t)^2}{4(n-t)(1-f_n(t))^2} + o(1/n).$$

The proof of Proposition 6.4 can be found in Section 6.2.2. Moreover, in Section 6.2.3 we prove the following.

Proposition 6.5. *Fix* $\varepsilon > 0$ *and* $\ell \geq 1$. *For any* $0 \leq t_1 < \cdots < t_\ell < (1-\varepsilon)n$ *we have*

$$\mathbb{P}\left[\bigcap_{i=1}^{\ell} \mathscr{R}_{t_i}\right] \sim \prod_{i=1}^{\ell} \frac{f_n(t_i)^2}{4(n-t_i)(1-f_n(t_i))^2}. \tag{6.3}$$

Finally, the following statement deals with the εn final steps of the algorithm.

Proposition 6.6. *For any* $\delta > 0$ *there exists* $\varepsilon > 0$ *such that* $\mathbb{P}\left[\bigcup_{(1-\varepsilon)n<t<n} \mathscr{R}_t\right] < \delta$.

Before we proceed we notice that Propositions 6.4–6.6 imply the first part of Theorem 1.1.

Proof of Theorem 1.1 (i). Pick $\delta > 0$, fix a small enough $\varepsilon = \varepsilon(\delta) > 0$ and let $\boldsymbol{R} = \sum_{t=0}^{n-1} \mathbb{1}\{\mathscr{R}_t\}$ be the total number of times at which conflicts occur. Proposition 6.1 shows that the probability that BPGD succeeds equals $\mathbb{P}[\boldsymbol{R} = 0]$. In order to calculate $\mathbb{P}[\boldsymbol{R} = 0]$, let $\boldsymbol{R}_\varepsilon = \sum_{0 \leq t \leq (1-\varepsilon)n} \mathbb{1}\{\mathscr{R}_t\}$ be the number of failures before time $(1-\varepsilon)n$. Proposition 6.5 shows that for any fixed $\ell \geq 1$ we have

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{\ell}(\boldsymbol{R}_\varepsilon - i + 1)\right] &= \ell! \sum_{0 \leq t_1 < \cdots < t_\ell \leq (1-\varepsilon)n} \mathbb{P}\left[\bigcap_{i=1}^{\ell} \mathscr{R}_{t_i}\right] \sim \ell! \sum_{0 \leq t_1 < \cdots < t_\ell \leq (1-\varepsilon)n} \prod_{i=1}^{\ell} \frac{f_n(t_i)^2}{4(n-t_i)(1-f_n(t_i))^2} \\
&= (1+o(1)) \sum_{0 \leq t_1, \ldots, t_\ell \leq (1-\varepsilon)n} \prod_{i=1}^{\ell} \frac{f_n(t_i)^2}{4(n-t_i)(1-f_n(t_i))^2} \sim \mathbb{E}[\boldsymbol{R}_\varepsilon]^\ell.
\end{aligned} \tag{6.4}$$

Hence, the inclusion/exclusion principle (e.g., [4, Theorem 1.21]) implies that

$$\mathbb{P}[\boldsymbol{R}_\varepsilon = 0] \sim \exp(-\mathbb{E}[\boldsymbol{R}_\varepsilon]). \tag{6.5}$$

Further, using Proposition 6.4 and the linearity of expectation, we obtain with $\lambda(\theta) = -\log(1-\theta)$

$$\begin{aligned}
\mathbb{E}[\boldsymbol{R}_\varepsilon] &= \sum_{0 \leq t \leq (1-\varepsilon)n} \mathbb{P}[\mathscr{R}_t] \sim \sum_{0 \leq t \leq (1-\varepsilon)n} \frac{f_n(t)^2}{4(n-t)(1-f_n(t))^2} \sim \frac{1}{4n} \int_0^{1-\varepsilon} \frac{f_n(\theta n)^2}{(1-\theta)(1-f_n(\theta n))^2} \mathrm{d}\theta \\
&= \frac{1}{4n} \int_0^{1-\varepsilon} \frac{f_n(\theta n)^2}{(1-\alpha_*)(1-f_n(\theta n))} \frac{\partial \alpha_*}{\partial \lambda} \frac{\partial \lambda(\theta)}{\partial \theta} \mathrm{d}\theta && \text{[by (3.11)]} \\
&= \frac{d^2(k-1)^2}{4} \int_0^{1-\varepsilon} \frac{z^{2k-4}(1-z)}{1-d(k-1)z^{k-2}(1-z)} \mathrm{d}z && \text{[by (6.2)].}
\end{aligned} \tag{6.6}$$

Finally, Proposition 6.6 implies that

$$\mathbb{P}[\boldsymbol{R} > \boldsymbol{R}_\varepsilon] < \delta. \tag{6.7}$$

Thus, the assertion follows from (6.5)–(6.7) upon taking the limit $\delta \to 0$. □

6.2.1. *Proof of Proposition 6.3.* The proof of Proposition 6.3 is based on the method of differential equations. Specifically, based on Fact 6.2 we derive a system of ODEs that track the random variables $\boldsymbol{\alpha}(t), \boldsymbol{m}_2(t), \ldots, \boldsymbol{m}_k(t)$. We will then identify the unique solution to this system. As a first step we work out the conditional expectations of $\boldsymbol{\alpha}(t+1), \boldsymbol{m}_2(t+1), \ldots, \boldsymbol{m}_k(t+1)$ given $\mathfrak{F}_t$.

Lemma 6.7. *If* $2\boldsymbol{m}_2(t)/\boldsymbol{n}(t) < 1 - \Omega(1)$ *and* $\boldsymbol{n}(t) = \Omega(n)$, *then*

$$\mathbb{E}[\boldsymbol{n}(t) - \boldsymbol{n}(t+1) \mid \mathfrak{F}_t] = \frac{\boldsymbol{n}(t)^2}{(n-t)(\boldsymbol{n}(t) - 2\boldsymbol{m}_2(t))} + o(1), \tag{6.8}$$

$$\mathbb{E}[\boldsymbol{m}_\ell(t+1) - \boldsymbol{m}_\ell(t) \mid \mathfrak{F}_t] = \frac{\boldsymbol{n}(t)^2}{(n-t)(\boldsymbol{n}(t) - 2\boldsymbol{m}_2(t))} \cdot \frac{(\ell+1)\boldsymbol{m}_{\ell+1}(t) - \ell \boldsymbol{m}_\ell(t)}{\boldsymbol{n}(t)} + o(1) \qquad (2 \leq \ell < k), \tag{6.9}$$

$$\mathbb{E}[\boldsymbol{m}_k(t+1) \mid \mathfrak{F}_t] = -\frac{\boldsymbol{n}(t)^2}{(n-t)(\boldsymbol{n}(t) - 2\boldsymbol{m}_2(t))} \cdot \frac{k\boldsymbol{m}_k(t)}{\boldsymbol{n}(t)} + o(1). \tag{6.10}$$

Proof. Going from time t to time $t+1$ involves the express assignment of variable x_{t+1}, unless it had already been assigned a value due to previous decisions, and the subsequent pursuit of unit clause implications. The probability given $\mathfrak{F}_t$ that x_{t+1} was set in a previous iteration equals

$$q_{t+1} = 1 - \frac{\boldsymbol{n}(t)}{n-t}. \tag{6.11}$$

Indeed, the first t iterations assigned values to a total of $n - \boldsymbol{n}(t)$ variables, including $x_1, \ldots, x_t$, and Fact 6.2 shows that the identities of the assigned variables among $x_{t+1}, \ldots, x_n$ are random.

Let $\mathcal{Q}_{t+1}$ be the event that x_{t+1} was not assigned previously. Given $\mathcal{Q}_{t+1}$ we need to pursue unit clause implications. To this end, recall the bipartite graph representation $G(\boldsymbol{F}_{\mathrm{UC},t})$ of the formula $\boldsymbol{F}_{\mathrm{UC},t}$. Let $G_2(\boldsymbol{F}_{\mathrm{UC},t})$ be the subgraph of $G(\boldsymbol{F}_{\mathrm{UC},t})$ obtained by removing all clauses of length greater than two. Then Fact 6.2 shows that $G_2(\boldsymbol{F}_{\mathrm{UC},t})$ is a uniformly random bipartite graph with $\boldsymbol{n}(t)$ nodes on one side and $\boldsymbol{m}_2(t)$ nodes of degree two on the other side. Furthermore, the number of variables whose values are implied by unit clause propagation is lower bounded by the number of variable nodes in the component of x_{t+1} in $G_2(\boldsymbol{F}_{\mathrm{UC},t})$. The expected size of this component can be computed as the expected progeny of a branching process with offspring $\mathrm{Po}(2\boldsymbol{m}_2(\ell)/\boldsymbol{n}(t))$. As is well known, under the assumption $2\boldsymbol{m}_2(t)/\boldsymbol{n}(t) < 1 - \Omega(1)$ that the branching process is sub-critical, the expected progeny comes to $(1 - 2\boldsymbol{m}_2(t)/\boldsymbol{n}(t))^{-1}$. Hence, we obtain

$$\mathbb{E}\left[n(\boldsymbol{\alpha}(t+1) - \boldsymbol{\alpha}(t)) \mid \mathfrak{F}_t\right] \geq \frac{1 - q_{t+1}}{1 - 2\boldsymbol{m}_2(t)/\boldsymbol{n}(t)}. \tag{6.12}$$

Strictly speaking, (6.12) only gives a lower bound on $\mathbb{E}\left[n(\boldsymbol{\alpha}(t+1) - \boldsymbol{\alpha}(t)) \mid \mathfrak{F}_t\right]$ because additional unit clause implications could arise from clauses of length greater than two. However, for this to happen a clause would have to contain at least two variables that are set in iteration $t+1$ (i.e., either x_{t+1} itself or a variable whose value is implied due to unit clause propagation). But since $2\boldsymbol{m}_2(t)/\boldsymbol{n}(t) < 1 - \Omega(1)$, the expected number of such implications is bounded, and thus the expected number of longer clauses that turn into unit clauses is of order $O(1/n)$. Consequently, the lower bound (6.12) is tight up to an $O(1/n)$ error term, whence we obtain (6.8).

Moving on to (6.9)–(6.10) we notice that for $2 \leq \ell < k$ there are two ways in which the number of clauses of length ℓ can change from iteration t to iteration $t+1$. First, it could be that clauses of length ℓ contain one variable that gets a value assigned. Any such clauses shorten to length $\ell - 1$ (if $\ell > 2$) or become unit clauses and subsequently disappear ($\ell = 2$). In light of Fact 6.2, the probability that a given clause of length ℓ suffers this fate comes to $\ell(\boldsymbol{n}(t) - \boldsymbol{n}(t+1))/\boldsymbol{n}(t) + o(1)$. Conversely, if $\ell < k$ additional clauses of length ℓ may result from the shortening of clauses of length $\ell + 1$. Analogously to the previous computation, the probability that a given clause of length $\ell + 1$ shortens to length ℓ comes to $(\ell+1)(\boldsymbol{n}(t) - \boldsymbol{n}(t+1))/\boldsymbol{n}(t) + o(1)$. Of course, there could also be clauses that contain more than one variable that receives a value during iteration $t+1$. However, the probability of this event is of order $O(1/n^2)$. Hence, (6.8) implies (6.9) and (6.10). $\square$

Lemma 6.7 puts us in a position to derive a system of ODEs to track the random variables $\boldsymbol{n}(t), \boldsymbol{m}_2(t), \ldots, \boldsymbol{m}_k(t)$. Specifically, we obtain the following.

Corollary 6.8. *Let* $\mathfrak{n}, \mathfrak{m}_2, \ldots, \mathfrak{m}_k : [0,1] \to \mathbb{R}$ *be continuously differentiable functions such that*

$$\mathfrak{n}(0) = 1, \qquad \mathfrak{m}_k(0) = \frac{d}{k}, \tag{6.13}$$

$$\frac{\partial \mathfrak{n}}{\partial \theta} = -\frac{\mathfrak{n}^2}{(1-\theta)(\mathfrak{n} - 2\mathfrak{m}_2)}, \tag{6.14}$$

$$\frac{\partial \mathfrak{m}_\ell}{\partial \theta} = \frac{\mathfrak{n}((\ell+1)\mathfrak{m}_{\ell+1} - \ell \mathfrak{m}_\ell)}{(1-\theta)(\mathfrak{n} - 2\mathfrak{m}_2)} \quad (2 \leq \ell < k), \qquad \frac{\partial \mathfrak{m}_k}{\partial \theta} = -\frac{k \mathfrak{n} \mathfrak{m}_k}{(1-\theta)(\mathfrak{n} - 2\mathfrak{m}_2)}. \tag{6.15}$$

Assume, furthermore, that

$$\sup_{\theta \in [0,1]} 2\mathfrak{m}_2(\theta)/\mathfrak{n}(\theta) < 1. \tag{6.16}$$

Then with probability $1 - o(n^{-2})$ *for all* $0 \leq t \leq n$ *we have*

$$\boldsymbol{n}(t)/n = \mathfrak{n}(t/n) + o(1), \qquad \boldsymbol{m}_\ell(t)/n = \mathfrak{m}_\ell(t/n) + o(1) \quad (2 \leq \ell \leq k).$$

Proof. This follows from Lemma 6.7 in combination with [26, Theorem 2]. $\square$

As a next step we construct an explicit solution to the system (6.13)–(6.15).

Lemma 6.9. *If $d < d_{\min}$, then the functions*

$$\mathfrak{n}^*(\theta) = 1 - \alpha_*(\lambda(\theta)), \qquad \mathfrak{m}^*_\ell(\theta) = \frac{d}{k}\binom{k}{\ell}(1-\alpha_*(\lambda(\theta)))^\ell \alpha_*(\lambda(\theta))^{k-\ell}. \tag{6.17}$$

satisfy (6.13)–(6.16).

Proof. The initial condition (6.13) is satisfied because $\alpha_*(\lambda(0)) = 0$. Furthermore, (3.11) shows that

$$\frac{\partial \mathfrak{n}^*}{\partial \theta} = -\frac{\partial \alpha_*}{\partial \lambda}\cdot\frac{\partial \lambda}{\partial \theta} = -\frac{1-\alpha_*}{1-d(k-1)\alpha_*^{k-2}(1-\alpha_*)}\cdot\frac{1}{1-\theta} = -\frac{\mathfrak{n}^*}{(1-\theta)(1-2\mathfrak{m}^*_2/\mathfrak{n}^*)}. \tag{6.18}$$

Hence, (6.14) is satisfied. Furthermore, (6.18) implies that for $2 \le \ell < k$ we have

$$\begin{aligned}\frac{\partial \mathfrak{m}^*_\ell}{\partial \theta} &= \frac{d}{k}\cdot\frac{\partial \lambda}{\partial \theta}\cdot\frac{\partial \alpha_*}{\partial \lambda}\cdot\binom{k}{\ell}\left[(k-\ell)\alpha_*^{k-\ell-1}(1-\alpha_*)^\ell - \ell\alpha_*^{k-\ell}(1-\alpha_*)^{\ell-1}\right]\\ &= \frac{\mathfrak{n}^*}{(1-\theta)(1-2\mathfrak{m}^*_2/\mathfrak{n}^*)}\cdot\frac{d}{k(1-\alpha_*)}\cdot\binom{k}{\ell}\left[(\ell+1)(1-\alpha_*)^{\ell+1}\alpha_*^{k-\ell-1} - \ell\alpha_*^{k-\ell}(1-\alpha_*)^\ell\right]\\ &= \frac{\mathfrak{n}^*}{(1-\theta)(\mathfrak{n}^*-2\mathfrak{m}^*_2)}\cdot\left[(\ell+1)\mathfrak{m}^*_{\ell+1} - \ell\mathfrak{m}^*_\ell\right],\end{aligned}$$

which is the first part of (6.15). An analogous computation yields the second part of (6.15). Finally, (6.16) follows from (3.11). □

Proof of Proposition 6.3. The proposition is an immediate consequence of Corollary 6.8 and Lemma 6.9. □

6.2.2. *Proof of Proposition 6.4.* Recall that $\boldsymbol{F}'_{\mathrm{UC},t+1}$ is the XORSAT formula that contains the variables $\boldsymbol{V}'(t+1)$ that get assigned during iteration $t+1$ and the clauses $\boldsymbol{C}'(t+1)$ of $\boldsymbol{F}_{\mathrm{UC},t}$ that contain variables from $\boldsymbol{V}'(t+1)$ only. Also recall that $G(\boldsymbol{F}'_{\mathrm{UC},t+1})$ signifies the graph representation of this XORSAT formula. Unless $\boldsymbol{V}'(t+1) = \emptyset$, the graph $G(\boldsymbol{F}'_{\mathrm{UC},t+1})$ is connected.

Lemma 6.10. *Fix $\varepsilon > 0$ and let $0 \le t \le (1-\varepsilon)n$. With probability $1-o(1/n)$ the graph $G(\boldsymbol{F}'_{\mathrm{UC},t+1})$ satisfies*

$$|E(G(\boldsymbol{F}'_{\mathrm{UC},t+1}))| \le |V(G(\boldsymbol{F}'_{\mathrm{UC},t+1}))|.$$

Proof. We recall from the proof of Lemma 6.7 that iteration $t+1$ of `UCP` can be described by a branching process on the random graph $G(\boldsymbol{F}_{\mathrm{UC},t})$. Given that x_{t+1} is still unassigned, the offspring distribution of the branching process has mean $2\boldsymbol{m}_2(t)/\boldsymbol{n}(t)$. Moreover, Proposition 6.3 shows that with probability $1-O(n^{-2})$ we have $2\boldsymbol{m}_2(t)/\boldsymbol{n}(t) \sim d(k-1)(1-\alpha_*)\alpha_*^{k-2} < 1$ (as $d < d_{\min}$). Hence, the branching process is sub-critical. As a consequence, with probability $1-O(n^{-2})$ we have

$$\mathbb{P}\left[|V(G(\boldsymbol{F}'_{\mathrm{UC},t+1}))| \ge \log^2 n\right] = O(n^{-2}). \tag{6.19}$$

Each step of the branching process corresponds to pursuing the unit clause implications of assigning a truth value to a single variable x. A cycle in $G(\boldsymbol{F}'_{\mathrm{UC},t+1})$ can only ensue if a clause that contains x also contains a variable that has already been set previously during iteration $t+1$. In light of (6.19), with probability $1-O(n^{-2})$ there are no more than $\log^2 n$ such variables. Hence, the probability that the assignment of x closes a cycle is of order $O(\log^2 n/n)$. Additionally, by the principle of deferred decisions the events that two different clauses processed by unit clause propagation close cycles is of order $O(\log^4 n/n^2)$. Finally, since by (6.19) we may assume that the total number of clauses does not exceed $O(\log^2 n)$, we conclude that

$$\mathbb{P}\left[|E(G(\boldsymbol{F}'_{\mathrm{UC},t+1}))| > |V(G(\boldsymbol{F}'_{\mathrm{UC},t+1}))|\right] = O(\log^6 n/n^2) = o(1/n),$$

as desired. □

Thus, with probability $1-o(1/n)$ the graph $G(\boldsymbol{F}'_{\mathrm{UC},t+1})$ contains at most one cycle. While it is easy to check that no conflict occurs in iteration $t+1$ if $G(\boldsymbol{F}'_{\mathrm{UC},t+1})$ is acyclic, in the case that $G(\boldsymbol{F}'_{\mathrm{UC},t+1})$ contains a single cycle there is a chance of a conflict. The following definition describes the type of cycle that poses an obstacle.

Definition 6.11. *For a XORSAT formula F we call a sequence of variables and clauses $\mathscr{C} = (v_1, c_1, \dots, v_\ell, c_\ell, v_\ell + 1 = v_1)$ a* toxic cycle *of length ℓ if*

TOX1: *c_i contains the variables x_i, x_{i+1} only, and*
TOX2: *the total number of negations in $c_1, \dots c_\ell$ is odd iff ℓ is even.*

Lemma 6.12. *(i) If $\boldsymbol{F}'_{\mathrm{UC},t+1}$ contains a toxic cycle, then a conflict occurs in iteration $t+1$.*
(ii) If $\boldsymbol{F}'_{\mathrm{UC},t+1}$ contains no toxic cycle and $|E(G(\boldsymbol{F}'_{\mathrm{UC},t+1}))| \le |V(G(\boldsymbol{F}'_{\mathrm{UC},t+1}))|$, then no conflict occurs in iteration $t+1$.

Proof. Towards (i) we show that $\boldsymbol{F}'_{\mathrm{UC},t+1}$ is not satisfiable if there is a toxic cycle $\mathscr{C} = (v_1, c_1, \dots, c_\ell, v_{\ell+1} = v_1)$; then UCP will, of course, run into a contradiction. To see that $\boldsymbol{F}'_{\mathrm{UC},t+1}$ is unsatisfiable, we transform each of the clauses $c_1, \dots, c_\ell$ into a linear equation $c_i \equiv (v_i + v_{i+1} = y_i)$ over $\mathbb{F}_2$. Here $y_i \in \mathbb{F}_2$ equals 1 iff c_i contains an even number of negations. Adding these equations up yields $\sum_{i=1}^{\ell} y_i = 0$ in $\mathbb{F}_2$. This condition is violated if $\mathscr{C}$ is toxic.

Let us move on to (ii). Assume for contradiction that there exists a formula F without a toxic cycle such that $|V(G(F))| \le |E(G(F))|$ and such that given $\boldsymbol{F}'_{\mathrm{UC},t+1} = F$, UCP may run into a conflict. Consider such a formula F that minimises $|V(F)| + |C(F)|$. Since UCP succeeds on acyclic F, we have $|V(G(F))| = |E(G(F))|$. Thus, $G(F)$ contains a single cycle $\mathscr{C} = (v_1, c_1, \dots, v_\ell, c_\ell, v_{\ell+1} = v_1)$. Apart from the cycle, F contains (possibly empty) acyclic formulas $F'_1, \dots, F'_\ell$ attached to $v_1, \dots, v_\ell$ and $F''_1, \dots, F''_\ell$ attached to $c_1, \dots, c_\ell$. The formulas $F'_1, F''_1, \dots, F'_\ell, F''_\ell$ are mutually disjoint and do not contain unit clauses.

We claim that $F'_1, \dots, F'_\ell$ are empty because $|V(F)| + |C(F)|$ is minimum. This is because given any truth assignment of $v_1, \dots, v_\ell$, UCP will find a satisfying assignment of the acyclic formulas $F'_1, \dots, F'_\ell$.

Further, assume that one of the formulas $F''_1, \dots, F''_\ell$ is non-empty; say, F''_1 is non-empty. If the start variable that UCP assigns were to belong to F''_1, then c_1, containing x_1 and x_2, would not shrink to a unit clause, and thus UCP would not assign values to these variables. Hence, UCP starts by assigning a truth value to one of the variables $v_1, \dots, v_\ell$; say, UCP starts with v_1. We claim that then UCP does not run into a conflict. Indeed, the clauses $c_2, \dots, c_\ell$ may force UCP to assign truth values to $x_2, \dots, x_\ell$, but no conflict can ensue because UCP will ultimately satisfy c_1 by assigning appropriate truth values to the variables of F''_1.

Thus, we may finally assume that all of $F'_1, F''_1, \dots, F'_\ell, F''_\ell$ are empty. In other words, F consists of the cycle $\mathscr{C}$ only. Since $\mathscr{C}$ is not toxic, **TOX2** does not occur. Consequently, UCP will construct an assignment that satisfies all clauses $c_1, \dots, c_\ell$. This final contradiction implies (ii). □

Corollary 6.13. *Fix $\varepsilon > 0$ and let $0 \le t \le (1-\varepsilon)n$. Then*

$$\mathbb{P}[\mathscr{R}_{t+1}] = \mathbb{P}\left[\boldsymbol{F}'_{\mathrm{UC},t+1} \text{ contains a toxic cycle}\right] + o(1/n).$$

Proof. This is an immediate consequence of Lemma 6.10 and Lemma 6.12. □

Thus, we are left to calculate the probability that $\boldsymbol{F}'_{\mathrm{UC},t+1}$ contains a toxic cycle. To this end, we estimate the number of toxic cycles in the 'big' formula $\boldsymbol{F}_{\mathrm{UC},t}$. Let $\boldsymbol{T}_{t,\ell}$ be the number of toxic cycles of length ℓ in $\boldsymbol{F}_{\mathrm{UC},t}$.

Lemma 6.14. *Fix $\varepsilon > 0$ and let $1 \le t \le (1-\varepsilon)n$.*
(i) For any fixed ℓ, with probability $1 - O(n^{-2})$ we have

$$\mathbb{E}[\boldsymbol{T}_t(\ell) \mid \mathfrak{F}_t] = \beta_\ell + o(1), \qquad \text{where } \beta_\ell = \frac{1}{4\ell}\left(d(k-1)(1-\alpha_*)\alpha_*^{k-2}\right)^\ell = \frac{1}{4\ell}\left(f_n(t)\right)^\ell.$$

(ii) For any $1 \le \ell \le n$, with probability $1 - O(n^{-2})$ we have $\mathbb{E}[\boldsymbol{T}_t(\ell) \mid \mathfrak{F}_t] \le \beta_\ell \exp(\varepsilon\ell)$.

Proof. In light of Fact 6.2, the calculation of the expected number of toxic cycles is straightforward. Indeed, we just need to pick sequences of ℓ distinct variables and clauses, place the variables into the clauses in a cyclic fashion, and multiply by the probability that the clauses contain no other variables and that the parity of the signs of the clauses works out as per **TOX2**. Of course, in this way we over count toxic cycles 2ℓ times (due to the choice of the starting point and the orientation). Hence, we obtain

$$\mathbb{E}[\boldsymbol{T}_t(\ell) \mid \mathfrak{F}_t] = \frac{(n)_\ell (m)_\ell}{4\ell n^{2\ell}} (k(k-1))^\ell (1-\boldsymbol{\alpha}(t))^\ell \boldsymbol{\alpha}(t)^{\ell(k-2)}. \tag{6.20}$$

Thus, (i) follows from (6.20) and Proposition 6.3. Further, (6.20) demonstrates that

$$\mathbb{E}[\boldsymbol{T}_t(\ell) \mid \mathfrak{F}_t] \le \frac{1}{4\ell}\left(d(k-1)(1-\boldsymbol{\alpha}(t))\boldsymbol{\alpha}(t)^{k-2}\right)^\ell. \tag{6.21}$$

Finally, combining (6.21) with Proposition 6.3, we obtain (ii). □

Proof of Proposition 6.4. In light of Corollary 6.13 we just need to calculate the probability that $\boldsymbol{F}'_{\mathrm{UC},t+1}$ contains a toxic cycle. Clearly, if during iteration $t+1$ `UCP` encounters a variable of $\boldsymbol{F}_{\mathrm{UC},t}$ that lies on a toxic cycle, `UCP` will proceed to add the entire toxic cycle to $\boldsymbol{F}'_{\mathrm{UC},t+1}$ (and run into a contradiction). Furthermore, Lemma 6.14 shows that with probability $1-O(n^{-2})$ given $\mathfrak{F}_t$ the probability that a random variable of $\boldsymbol{F}_{\mathrm{UC},t}$ belongs to a toxic cycle comes to

$$\bar{\beta}=\sum_{\ell\geq 2}\ell\beta_\ell+o(1)=\sum_{\ell\geq 2}\frac{1}{4}\left(f_n(t)\right)^\ell=\frac{f_n(t)^2}{4(1-f_n(t))}+o(1)=O(1). \tag{6.22}$$

We now use (6.22) to calculate the desired probability of encountering a toxic cycle. To this end we recall from the proof of Lemma 6.7 that the $(t+1)$-st iteration of `UCP` corresponds to a branching process with expected offspring $f_n(t)$, unless the root variable x_{t+1} has already been assigned. Due to (6.11) and Proposition 6.3, with probability $1-O(n^{-2})$ the conditional probability of this latter event equals $(n\alpha_*-t)/(n-t)+o(1)$. Further, given that the root variable has not been assigned previously, the expected progeny of the branching process, i.e., the expected number of variables in $\boldsymbol{F}'_{\mathrm{UC},t+1}$, equals $1/(1-f_n(t))+o(1)$. Since with probability $1-O(n^{-2})$ given $\mathfrak{F}_t$ there remain $\boldsymbol{n}(t)=(1-\alpha_*+o(1))n$ unassigned variables in total, (6.22) implies that with probability $1-o(1/n)$,

$$\mathbb{P}\left[\mathscr{R}_{t+1}\mid\mathfrak{F}_t\right]\sim\frac{\bar{\beta}}{(1-\alpha_*)n}\cdot\frac{1-\alpha_*}{1-t/n}\cdot\frac{1}{1-f_n(t)}=\frac{f_n(t)^2}{4(1-f_n(t))^2(n-t)}+o(1/n),$$

as claimed. □

6.2.3. *Proof of Proposition 6.5.* We combine Fact 6.2 with the tower rule. Specifically, let $0\leq t_1<\cdots<t_h<(1-\varepsilon)n$ be distinct time indices. Then repeated application of the tower rule gives

$$\begin{aligned}\mathbb{P}\left[\bigcap_{i=1}^h\mathscr{R}_{t_i}\right]&=\mathbb{E}\left[\prod_{i=1}^h\mathbb{1}\{\mathscr{R}_{t_i}\}\right]=\mathbb{E}\left[\mathbb{E}\left[\prod_{i=1}^h\mathbb{1}\{\mathscr{R}_{t_i}\}\mid\mathfrak{F}_{t_i-1}\right]\right]\\&=\mathbb{E}\left[\left(\prod_{i=1}^{h-1}\mathbb{1}\{\mathscr{R}_{t_i}\}\right)\mathbb{P}\left[\mathscr{R}_{t_h}\mid\mathfrak{F}_{t_h-1}\right]\right]=\cdots=\mathbb{E}\left[\prod_{i=1}^h\mathbb{P}\left[\mathscr{R}_{t_i}\mid\mathfrak{F}_{t_i-1}\right]\right].\end{aligned} \tag{6.23}$$

Furthermore, Proposition 6.4 shows that with probability $1-o(1/n)$,

$$\mathbb{P}\left[\mathscr{R}_{t_i}\mid\mathfrak{F}_{t_i-1}\right]=\frac{f_n(t_i)^2}{4(n-t_i)(1-f_n(t_i))^2}+o(1/n)\qquad\text{for all }1\leq i\leq h. \tag{6.24}$$

Combining (6.23)–(6.24) completes the proof.

6.2.4. *Proof of Proposition 6.6.* Given $\delta>0$ pick $\varepsilon>0$ small enough and let $t=\lceil(1-\varepsilon)n\rceil$. We are going to show that the graph $G(\boldsymbol{F}_{\mathrm{UC},t})$ is acyclic with probability at least $1-\delta$. Since all clauses of $\boldsymbol{F}_{\mathrm{UC},t}$ contain at least two variables, `UCP` will find a satisfying assignment if $G(\boldsymbol{F}_{\mathrm{UC},t})$ is acyclic.

To show that $G(\boldsymbol{F}_{\mathrm{UC},t})$ is acyclic, we observe that $\alpha_*\geq t/n$. Hence, α_* approaches one as $t/n\to 1$. Further, Fact 6.2 shows that $G(\boldsymbol{F}_{\mathrm{UC},t})$ is uniformly random given the degree distribution (6.1) of the clause nodes. Indeed, the expression (6.1) shows that with probability $1-O(n^{-2})$ the expected size of the second neighbourhood of a given variable node is asymptotically equal to

$$\gamma=\gamma(\varepsilon)=\frac{1}{(1-\alpha_*)n}\cdot\frac{dn}{k}\sum_{\ell=2}^k\ell\binom{k}{\ell}(1-\alpha_*)^\ell\alpha_*^{k-\ell}=d(1-\alpha_*^{k-1}).$$

Hence, as $\lim_{\varepsilon\to 0}\gamma=0$, the average degree of the random graph $G(\boldsymbol{F}_{\mathrm{UC},t})$ tends to zero as $\varepsilon\to 0$. Therefore, for small enough $\varepsilon>0$ the random graph $\boldsymbol{G}(\boldsymbol{F}_{\mathrm{UC},t})$ is acyclic with probability greater than $1-\delta$.

6.3. **Failure of `UCP` for $d_{\min}<d<d_{\mathrm{sat}}$.** In this section we assume that $d_{\min}<d<d_{\mathrm{sat}}$. As in Section 6.2 we are going to trace `UCP` via the method of differential equations. In particular, we keep the notation from Section 6.2. Thus, $\boldsymbol{n}(t)$ signifies the number of unassigned variables after t iterations, and $\boldsymbol{m}_\ell(t)$ denotes the number of clauses that contain precisely $2\leq\ell\leq k$ unassigned variables. Moreover, $\boldsymbol{F}_{\mathrm{UC},t}$ is the formula comprising these variables and clauses. The following statement is the analogue of Proposition 6.3 for $d_{\min}<d<d_{\mathrm{sat}}$. Its proof relies on similar arguments as the proof of Proposition 6.3.

Proposition 6.15. *Suppose that $d_{\min}(k)<d<d_{\mathrm{sat}}(k)$, fix $\varepsilon,\delta>0$ and let $0<t<(1-\varepsilon)\theta_*n$. Then* (6.1) *holds with probability* $1-O(n^{-2})$.

Proof. The formulas (6.8)–(6.10) for the conditional expected changes $\boldsymbol{n}(t+1)-\boldsymbol{n}(t), \boldsymbol{m}_\ell(t+1)-\boldsymbol{m}(t)$ continue to hold for $d_{\min} < d < d_{\text{sat}}$, so long as we assume that $2\boldsymbol{m}_2(t)/\boldsymbol{n}(t) < 1-\Omega(1)$ and $\boldsymbol{n}(t) = \Omega(n)$. Indeed, the proof of Lemma 6.7 only hinges on these assumptions on $\boldsymbol{n}(t), \boldsymbol{m}_2(t)$, irrespective of d. Hence, if $\mathfrak{n}, \mathfrak{m}_2, \ldots, \mathfrak{m}_k : [0, \theta_* - \delta] \to \mathbb{R}$ are functions that satisfy the conditions (6.13)–(6.15) and that satisfy

$$\sup_{\theta \in [0,\theta_*-\delta]} 2\mathfrak{m}_2(\theta)/\mathfrak{n}(\theta) < 1, \tag{6.25}$$

then [26, Theorem 2] implies that for all $0 \le t < (1-\delta)\theta_* n$ we have

$$\boldsymbol{n}(t)/n = \mathfrak{n}(t/n) + o(1), \qquad \boldsymbol{m}_\ell(t)/n = \mathfrak{m}_\ell(t/n) + o(1) \quad (2 \le \ell \le k).$$

Finally, we claim that the functions $\mathfrak{n}^* : [0, \theta_* - \delta] \to \mathbb{R}$, $\mathfrak{m}^*_\ell : [0, \theta_* - \delta] \to \mathbb{R}$ defined by (6.17) satisfy (6.13)–(6.15) and (6.25). In fact, the same manipulations as in the proof of Lemma 6.9 yield (6.13)–(6.15). Additionally, (6.25) follows from Lemma 3.5 (ii) and Proposition 2.2 (ii), which shows that α_* is a stable fixed point and therefore

$$2\mathfrak{m}_2(\theta)/\mathfrak{n}(\theta) = d(k-1)(1-\alpha_*)\alpha_*^{k-2} < 1 \qquad \text{for } 0 \le \theta \le \theta_* - \delta.$$

Thus, we obtain (6.1) for $0 \le \theta < \theta_*$. □

Proof of Theorem 1.1 (ii). Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in \{0,1\}$ be uniformly distributed, mutually independent and independent of all other randomness. We couple the execution of the decimation process and of the UCP algorithm on a random formula $\boldsymbol{F}$ as follows. At every time t where $\pi_{\boldsymbol{F}_{\text{DC},t}} = 1/2$, the decimation process sets $\boldsymbol{\sigma}_{\text{DC}}(x_{t+1}) = \boldsymbol{u}_{t+1}$. Similarly, whenever UCP executes Step 5 we set $\boldsymbol{\sigma}_{\text{UC}}(x_{t+1}) = \boldsymbol{u}_{t+1}$. Let $\boldsymbol{\Delta}$ be the first time $0 \le t < n$ such that $\boldsymbol{\sigma}_{\text{DC}}(x_{t+1}) \ne \boldsymbol{\sigma}_{\text{UC}}(x_{t+1})$; if $\boldsymbol{\sigma}_{\text{DC}}(x_{t+1}) = \boldsymbol{\sigma}_{\text{UC}}(x_{t+1})$ for all t, we set $\boldsymbol{\Delta} = n$.

We claim that UCP encounters a conflict if $\boldsymbol{\Delta} < n$. To see this, assume that $0 \le t < n$ satisfies $\boldsymbol{\sigma}_{\text{DC}}(x_{t+1}) \ne \boldsymbol{\sigma}_{\text{UC}}(x_{t+1})$ but $\boldsymbol{\sigma}_{\text{DC}}(x_{s+1}) \ne \boldsymbol{\sigma}_{\text{UC}}(x_{s+1})$ for all $0 \le s < t$ and that UCP did not encounter a conflict at any time $s \le t$. Then $\pi_{\boldsymbol{F}_{\text{DC},t}} \in \{0,1\}$ but Step 5 of UCP sets $\boldsymbol{\sigma}_{\text{UC}}(x_{t+1}) = \boldsymbol{u}_{t+1} \ne \boldsymbol{\sigma}_{\text{DC}}(x_{t+1})$. Consequently, $\boldsymbol{F}$ possesses no satisfying assignment σ such that $\boldsymbol{\sigma}_{\text{UC}}(x_i) = \sigma(x_i)$ for $1 \le i \le t+1$, and thus UCP will ultimately encounter a conflict.

To complete the proof we claim that $\mathbb{P}[\boldsymbol{\Delta} < n] = 1 - o(1)$. To verify this consider a time $(1+\varepsilon)\theta_{\text{cond}} < t/n < (1-\varepsilon)\theta_* n$. Then Proposition 2.2 and Proposition 2.8 show that $|V_0(\boldsymbol{F}_{\text{DC},t})| = \alpha^* n + o(n)$ w.h.p., while Proposition 6.15 shows that $\boldsymbol{\alpha}(t) = \alpha_* + o(1)$ w.h.p. In particular, even if $\boldsymbol{\Delta} \ge (1+\varepsilon)\theta_{\text{cond}}$, the probability that $\pi_{\boldsymbol{F}_{\text{DC},t}} \in \{0,1\}$ while UCP assigns x_{t+1} randomly is $\Omega(1)$. Therefore, $\boldsymbol{\Delta} < \theta_* n$ w.h.p. □

References

[1] D. Achlioptas, A. Coja-Oghlan: Algorithmic barriers from phase transitions. Proc. 49th FOCS (2008) 793–802.
[2] D. Achlioptas, M. Molloy: The solution space geometry of random linear equations. Random Structures and Algorithms **46** (2015) 197–231.
[3] P. Ayre, A. Coja-Oghlan, P. Gao, N. Müller: The satisfiability threshold for random linear equations. Combinatorica **40** (2020) 179–235.
[4] B. Bollobás: Random graphs. Cambridge University Press (2001).
[5] A. Braunstein, M. Mézard, R. Zecchina: Survey propagation: An algorithm for satisfiability. Random Structures and Algorithms **27** (2005) 201–226.
[6] A. Coja-Oghlan: Belief Propagation fails on random formulas. Journal of the ACM **63** (2017) #49.
[7] A. Coja-Oghlan: A better algorithm for random k-SAT. SIAM J. Computing **39** (2010) 2823–2864.
[8] A. Coja-Oghlan, A. Ergür, P. Gao, S. Hetterich, M. Rolvien: The rank of sparse random matrices. Proc. 31st SODA (2020) 579–591.
[9] A. Coja-Oghlan, P. Gao, M. Hahn-Klimroth, J. Lee, N. Müller, M. Rolvien: The full rank condition for sparse random matrices. Combinatorics, Probability and Computing **33** (2024), 643–707.
[10] A. Coja-Oghlan, A. Pachon-Pinzon: The decimation process in random k-SAT. SIAM Journal on Discrete Mathematics **26** (2012) 1471–1509.
[11] C. Deroulers, R. Monasson: Criticality and universality in the unit-propagation search rule. Eur. Phys. J. B **49** (2006) 339–369.
[12] O. Dubois, J. Mandler: The 3-XORSAT threshold. Proc. 43rd FOCS (2002) 769–778.
[13] A. Frieze, S. Suen: Analysis of two simple heuristics on a random instance of k-SAT. J. Algorithms **20** (1996) 312–355.
[14] D. Gamarnik: The overlap gap property: a topological barrier to optimizing over random structures. PNAS **118** (2021) e2108492118.
[15] S. Hetterich: Analysing survey propagation guided decimation on random formulas. Proc. 43rd ICALP (2016) #65.
[16] M. Ibrahimi, Y. Kanoria, M. Kraning, A. Montanari: The set of solutions of random XORSAT formulae. Annals of Applied Probability **25** (2015) 2743–2808.
[17] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborová: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. National Academy of Sciences **104** (2007) 10318–10323.
[18] A. Maier, F. Behrens, L. Zdeborová: Dynamical cavity method for hypergraphs and its application to quenches in the k-XOR-SAT problem. arxiv 2412.14794 (2024).
[19] M. Mézard, A. Montanari: Information, physics and computation. Oxford University Press 2009.
[20] M. Mézard, T. Mora, R. Zecchina: Clustering of solutions in the random satisfiability problem. Phys. Rev. Lett. **94** (2005) 197205
[21] M. Mézard, G. Parisi, R. Zecchina: Analytic and algorithmic solution of random satisfiability problems. Science **297** (2002) 812–815.

[22] M. Mézard, F. Ricci-Tersenghi, R. Zecchina: Two solutions to diluted p-spin models and XORSAT problems. Journal of Statistical Physics **111** (2003) 505–533.
[23] M. Molloy: Cores in random hypergraphs and Boolean formulas. Random Structures and Algorithms **27** (2005) 124–135.
[24] B. Pittel, G. Sorkin: The satisfiability threshold for k-XORSAT. Combinatorics, Probability and Computing **25** (2016) 236–268.
[25] F. Ricci-Tersenghi, G. Semerjian: On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms. J. Stat. Mech. (2009) P09001.
[26] N. Wormald: Differential equations for random processes and random graphs. Ann. Appl. Probab. **5** (1995) 1217–1235.
[27] K. Yung: Limits of sequential local algorithms on the random k-XORSAT problem. Proc. 51st ICALP (2024) #123.

ARNAB CHATTERJEE, `arnab.chatterjee@tu-dortmund.de`, TU DORTMUND, FACULTY OF COMPUTER SCIENCE, 12 OTTO-HAHN-ST, DORTMUND 44227, GERMANY.

AMIN COJA-OGHLAN, `amin.coja-oghlan@tu-dortmund.de`, TU DORTMUND, FACULTY OF COMPUTER SCIENCE AND FACULTY OF MATHEMATICS, 12 OTTO-HAHN-ST, DORTMUND 44227, GERMANY.

MIHYUN KANG, `kang@math.tugraz.at`, TU GRAZ, INSTITUTE OF DISCRETE MATHEMATICS, STEYRERGASSE 30, 8010 GRAZ, AUSTRIA.

LENA KRIEG, `lena.krieg@tu-dortmund.de`, TU DORTMUND, FACULTY OF COMPUTER SCIENCE, 12 OTTO-HAHN-ST, DORTMUND 44227, GERMANY.

MAURICE ROLVIEN, `maurice.rolvien@tu-dortmund.de`, TU DORTMUND, FACULTY OF COMPUTER SCIENCE, 12 OTTO-HAHN-ST, DORTMUND 44227, GERMANY.

GREGORY B. SORKIN, `g.b.sorkin@lse.ac.uk`, THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE, DEPARTMENT OF MATHEMATICS, COLUMBIA HOUSE, HOUGHTON ST, LONDON WC2A 2AE, UNITED KINGDOM

THE RANDOM k-SAT GIBBS UNIQUENESS THRESHOLD REVISITED

ARNAB CHATTERJEE, AMIN COJA-OGHLAN, CATHERINE GREENHILL, VINCENT PFENNINGER, MAURICE ROLVIEN, PAVEL ZAKHAROV, KOSTAS ZAMPETAKIS

ABSTRACT. We prove that for any $k \geq 3$ for clause/variable ratios up to the Gibbs uniqueness threshold of the corresponding Galton-Watson tree, the number of satisfying assignments of random k-SAT formulas is given by the 'replica symmetric solution' predicted by physics methods [Monasson, Zecchina: Phys. Rev. Lett. **76** (1996)]. Furthermore, while the Gibbs uniqueness threshold is still not known precisely for any $k \geq 3$, we derive new lower bounds on this threshold that improve over prior work [Montanari and Shah: SODA (2007)]. The improvement is significant particularly for small k.
MSc: 68Q87, 60C05, 68R07

1. INTRODUCTION

1.1. **Background and motivation.** Going back to experimental work from the 1990s, the most prominent question concerning random k-SAT has been to pinpoint the satisfiability threshold, defined as the largest density m/n of clauses m to variables n up to which satisfying assignments likely exist [6, 18]. Currently, the satisfiability threshold is known precisely in the case of $k = 2$ [21, 40] and for $k \geq k_0$ with k_0 an undetermined (large) constant [33]. The latter result confirms 'predictions' based on an analytic but non-rigorous physics technique called the 'cavity method'. Indeed, the cavity method predicts the satisfiability threshold for every $k \geq 3$ [51], but random k-SAT for 'small' $k \geq 3$ appears to be a particularly hard nut to crack. Additionally, according to the cavity method several phase transitions precede the satisfiability threshold and are expected to impact, among other things, the performance of algorithms [47]. One of these phase transitions, the *Gibbs uniqueness transition*, pertains to a spatial mixing property that also plays a pivotal role in the computational complexity of counting and sampling [61].

From a statistical physics viewpoint, the satisfiability threshold is only the second most important quantity associated with random k-SAT. The first place firmly belongs to the typical number of satisfying assignments, known as the *partition function* in physics parlance [50]. All the other predictions, including the location of the satisfiability threshold, ultimately derive from the formula for the number of satisfying assignments or closely related variables [49]. Yet there has been little progress on confirming the physics formula for the number of satisfying assignments rigorously.

Three prior contributions stand out. First, a proof technique called the 'interpolation method' turns the physics prediction into a rigorous upper bound [35, 41, 60].[1] Second, in the case $k = 2$, conceptually much simpler than $k \geq 3$, the physics formula has been proved correct [3]. Third, Montanari and Shah [56] proved that also for $k \geq 3$ for certain clause/variable densities the 'replica symmetric solution' from physics correctly approximates the number of 'good' assignments that satisfy all but $o(n)$ clauses. However, it seems difficult to estimate the gap between the number of such 'good' assignments and the number of actual satisfying assignments. A rigorous method to this effect would likely imply the existence of uniform satisfiability thresholds for all $k \geq 3$, thereby resolving a long-standing conundrum [12, 36]. The proof of Montanari and Shah is based on the aforementioned Gibbs uniqueness property.

The aim of the present paper is to determine the number of *actual* satisfying assignments of random k-SAT formulas for clause/variable densities up to the Gibbs uniqueness threshold. Specifically, we verify that the 'replica symmetric solution' from [54, 55] yields the correct answer for any $k \geq 3$ right up to the Gibbs uniqueness threshold, even though the precise value of this threshold is not currently known. Additionally, we derive a new lower bound on the Gibbs uniqueness threshold. The improvement is particularly significant for 'small' $k \geq 3$. Combining these two results, we obtain the first rigorous formula for the number of satisfying assignments of random k-SAT formula for a non-trivial regime of clause/variable densities. Crucially, the result covers meaningful clause/variable densities even for small $k \geq 3$.

[1]Strictly speaking, the contributions [35, 41, 60] deal with the 'random k-SAT model at positive tempertature', see Section 2.6. In Corollary 2.2 below we combine the interpolation method with a concentration argument to bound the number of actual satisfying assignments.

1.2. **Results.** Let $\mathbf{\Phi} = \mathbf{\Phi}_{d,k}(n)$ be the random k-CNF on n Boolean variables $x_1,\dots,x_n$ with $\boldsymbol{m} = \boldsymbol{m}_n \sim \mathrm{Po}(dn/k)$ clauses $a_1,\dots,a_{\boldsymbol{m}}$. The clauses a_i are drawn independently and uniformly from the set of all $2^k\binom{n}{k}$ possible clauses with k distinct variables. Hence, the parameter d prescribes the expected number of clauses in which a given variable appears. Let $S(\mathbf{\Phi})$ be the set of satisfying assignments of $\mathbf{\Phi}$ and let $Z(\mathbf{\Phi}) = |S(\mathbf{\Phi})|$. We encode the Boolean values 'true' and 'false' by $+1$ and -1, respectively. Since right up to the satisfiability threshold $Z(\mathbf{\Phi})$ is of order $\exp(\Theta(n))$ w.h.p. for trivial reasons[2], our objective is to study the random variable $n^{-1}\log Z(\mathbf{\Phi})$ as $n \to \infty$.

1.2.1. *The number of satisfying assignments up to the Gibbs uniqueness threshold.* The first main result vindicates the 'replica symmetric solution' for values of d up to the Gibbs uniqueness threshold of the Galton-Watson tree that mimics the local topology of $\mathbf{\Phi}$. Let us define these concepts precisely.

We begin with the Galton-Watson tree $\mathbb{T} = \mathbb{T}_{d,k}$, which is generated by a two-type branching process. The two types are *variable nodes* and *clause nodes*. The process starts with a single root variable node $\mathfrak{x}$. The offspring of any variable node is a $\mathrm{Po}(d)$ number of clause nodes, while every clause node begets precisely $k-1$ variable nodes. Additionally, independently for each clause node a and every variable node x that is either a child or the parent of a a *sign*, denoted $\mathrm{sign}(x,a) \in \{\pm 1\}$, is chosen uniformly at random. The resulting random tree $\mathbb{T}$ models the local structure of the random formula $\mathbf{\Phi}$ in the sense of local weak convergence [9, 48].[3]

Next, we define the Gibbs uniqueness property on the tree $\mathbb{T}$. For an integer $\ell \geq 0$ let $\mathbb{T}^{(\ell)}$ be the finite tree obtained by removing all variable and clause nodes at a distance greater than 2ℓ from the root $\mathfrak{x}$. We identify the finite tree $\mathbb{T}^{(\ell)}$ with a Boolean formula whose variables/clauses are precisely the variable/clause nodes of $\mathbb{T}^{(\ell)}$. Let $S(\mathbb{T}^{(\ell)}) \neq \emptyset$ be the set of satisfying assignments of this formula and let $\boldsymbol{\tau}^{(\ell)} \in S(\mathbb{T}^{(\ell)})$ be a uniformly random satisfying assignment. Moreover, let $\partial^{2\ell}\mathfrak{x}$ be the set of variable nodes of $\mathbb{T}^{(\ell)}$ at distance precisely 2ℓ from the root $\mathfrak{x}$. Then for given d,k the tree $\mathbb{T} = \mathbb{T}_{d,k}$ has the *Gibbs uniqueness property* if

$$\lim_{\ell\to\infty} \mathbb{E}\left[\max_{\tau\in S(\mathbb{T}^{(\ell)})}\left|\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{x}) = 1 \mid \mathbb{T}\right] - \mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{x}) = 1 \mid \mathbb{T}, \forall x \in \partial^{2\ell}\mathfrak{x} : \boldsymbol{\tau}^{(\ell)}(x) = \tau(x)\right]\right|\right] = 0 \qquad \text{(see [47]).} \tag{1.1}$$

In words, in the limit of large ℓ the truth value $\boldsymbol{\tau}^{(\ell)}(\mathfrak{x})$ of the root $\mathfrak{x}$ is asymptotically independent of the truth values $\{\boldsymbol{\tau}^{(\ell)}(x)\}_{x\in\partial^{2\ell}\mathfrak{x}}$ of the variables at distance 2ℓ from $\mathfrak{x}$. In light of the above, for any $k \geq 2$ we further define $d_{\mathrm{uniq}}(k)$ as

$$d_{\mathrm{uniq}}(k) = \inf\{d > 0 : \text{ condition (1.1) fails to hold for } d,k\}\ . \tag{1.2}$$

It is easy to see that $d_{\mathrm{uniq}}(k)$ is strictly positive and finite for any $k \geq 2$. Indeed, in Theorem 1.2 we will derive explicit lower bounds on $d_{\mathrm{uniq}}(k)$. However, the exact value of $d_{\mathrm{uniq}}(k)$ is not currently known for any $k \geq 3$.

As a final preparation we need to spell out the 'replica symmetric solution' from [54]. This prediction comes in terms of a distributional fixed point problem, i.e., a fixed point problem on the space $\mathscr{P}(0,1)$ of probability measures on the open unit interval. Specifically, consider the *Belief Propagation operator*

$$\mathrm{BP}_{d,k} : \mathscr{P}(0,1) \to \mathscr{P}(0,1), \qquad\qquad \pi \mapsto \hat{\pi} = \mathrm{BP}_{d,k}(\pi) \tag{1.3}$$

defined as follows. Let $\boldsymbol{d}^+, \boldsymbol{d}^- \sim \mathrm{Po}(d/2)$ be Poisson variables with expectation $d/2$. Moreover, let $(\boldsymbol{\mu}_{\pi,i,j})_{i,j\geq 1}$ be a sequence of i.i.d. random variables, each following distribution π. All these random variables are mutually independent. Further, let

$$\boldsymbol{\mu}_{\pi,i} = 1 - \prod_{j=1}^{k-1}\boldsymbol{\mu}_{\pi,i,j} \quad \text{for } i \geq 1, \qquad\qquad \text{and} \qquad\qquad \hat{\boldsymbol{\mu}}_\pi = \frac{\prod_{i=1}^{\boldsymbol{d}^-}\boldsymbol{\mu}_{\pi,2i-1}}{\prod_{i=1}^{\boldsymbol{d}^-}\boldsymbol{\mu}_{\pi,2i-1} + \prod_{i=1}^{\boldsymbol{d}^+}\boldsymbol{\mu}_{\pi,2i}}. \tag{1.4}$$

Then $\hat{\pi}$ is the distribution of $\hat{\boldsymbol{\mu}}_\pi$. Furthermore, for a probability measure $\pi \in \mathscr{P}(0,1)$ define the *Bethe free entropy*[4]

$$\mathfrak{B}_{d,k}(\pi) = \mathbb{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}^-}\boldsymbol{\mu}_{\pi,2i-1} + \prod_{i=1}^{\boldsymbol{d}^+}\boldsymbol{\mu}_{\pi,2i}\right) - \frac{d(k-1)}{k}\log\left(1 - \prod_{j=1}^{k}\boldsymbol{\mu}_{\pi,1,j}\right)\right], \tag{1.5}$$

provided that the expectation on the r.h.s. exists. Finally, let $\delta_{1/2} \in \mathscr{P}(0,1)$ be the atom at $1/2$ and let us write $\mathrm{BP}_{d,k}^{\ell}$ for the ℓ-fold application of the operator $\mathrm{BP}_{d,k}$.

[2]For example, w.h.p. there are $\Omega(n)$ variables that do not appear in any clause.

[3]Corollary 3.7 below provides a precise statement to this effect.

[4]Throughout the paper log refers to the natural logarithm.

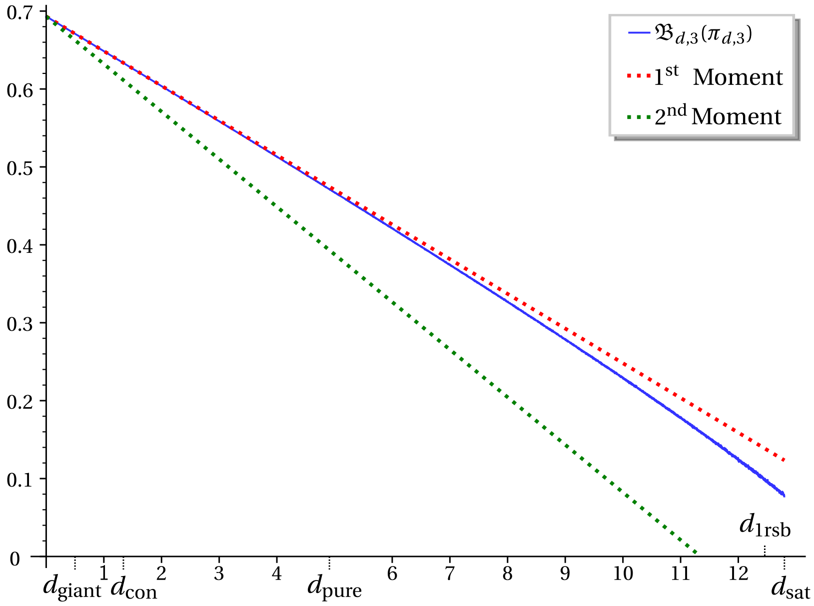


FIGURE 1. Comparison of $\mathfrak{B}_{d,k}(\pi_{d,k})$ with known bounds for $\lim_{n\to\infty}\frac{1}{n}\log Z(\mathbf{\Phi})$ for $k=3$. The red dotted line depicts the first moment upper bound (1.13), while the green dotted line represents the lower bound provided by (1.14). The blue line displays a numerical approximation of $\mathfrak{B}_{d,3}(\pi_{d,3})$. To obtain our values, we generated 10^6 samples from $\pi \approx \mathrm{BP}^{25}_{d,3}(\delta_{1/2})$ and then evaluated the corresponding empirical average of the expression in (1.5).

Theorem 1.1. *Let $k \geq 3$ and assume that $0 < d < d_{\text{uniq}}(k)$. Then the weak limit*

$$\pi_{d,k} = \lim_{\ell\to\infty} \mathrm{BP}^{\ell}_{d,k}(\delta_{1/2}) \in \mathscr{P}(0,1) \tag{1.6}$$

exists and

$$\lim_{n\to\infty} \frac{1}{n} \log Z(\mathbf{\Phi}) = \mathfrak{B}_{d,k}(\pi_{d,k}) \qquad \textit{in probability.} \tag{1.7}$$

The formula (1.7) matches the prediction from [54] precisely. Of course, part of the assertion of Theorem 1.1 is that the Bethe free entropy $\mathfrak{B}_{d,k}(\pi_{d,k})$ is well defined. Admittedly, the formula (1.7) is not 'explicit'. But the proof of Theorem 1.1 evinces that the convergence (1.6) occurs rapidly. Therefore, a randomised algorithm called 'population dynamics' [50] can be used to approximate (1.7) within any desired numerical accuracy.

1.2.2. *An improved lower bound on Gibbs uniqueness.* The obvious next task is to determine the Gibbs uniqueness threshold $d_{\text{uniq}}(k)$. Currently, its value is known precisely only in the case $k=2$, where $d_{\text{uniq}}(2)=2$ coincides with the random 2-SAT satisfiability threshold [3, 21, 40]. Furthermore, Montanari and Shah [56] proved that the *pure literal threshold*[5] $d_{\text{pure}}(k)$ upper bounds $d_{\text{uniq}}(k)$ for all $k \geq 2$.[6] The value of $d_{\text{pure}}(k)$ admits a neat formula [16, 53]:

$$d_{\text{uniq}}(k) \leq d_{\text{pure}}(k) = \min_{z>0} \frac{z}{(1-\exp(-z/2))^{k-1}}. \tag{1.8}$$

Complementing the upper bound (1.8), Montanari and Shah derived a lower bound $d_{\text{MS}}(k)$:

$$d_{\text{MS}}(k) = \sup\left\{d>0 : d(k-1)\left(1-\exp(-d/2)/4\right)\left(1-\exp(-d/2)/2\right)^{k-2} < 1\right\} \leq d_{\text{uniq}}(k). \tag{1.9}$$

[5]This marks the threshold up to which the pure literal algorithm–which repeatedly assigns the preferred value to all variables appearing with a single sign–produces a satisfying assignment w.h.p.

[6]To be precise, Montanari and Shah established an upper bound on the Gibbs uniqueness threshold that turns out to coincide with the pure literal threshold, albeit without pointing out this identity.

Unfortunately, the bound (1.9) is tight not even in the case $k = 2$, where $d_{\text{uniq}}(2) = 2$ while $d_{\text{MS}}(2) \approx 1.16$. That said, the lower and upper bounds $d_{\text{MS}}(k)$ and $d_{\text{pure}}(k)$ match asymptotically in the limit of large k, as

$$d_{\text{MS}}(k), d_{\text{pure}}(k) = (2 + o_k(1)) \log k, \tag{1.10}$$

with $o_k(1)$ hiding a term that vanishes as $k \to \infty$. The following theorem yields an improved lower bound $d_{\text{con}}(k)$ on $d_{\text{uniq}}(k)$.

Theorem 1.2. *For all $k \geq 3$ we have*

$$d_{\text{uniq}}(k) \geq d_{\text{con}}(k) := \sup\left\{ d > 0 : \frac{d(k-1)}{2} \left(1 - \exp(-d/2)/2\right)^{k-2} < 1 \right\}. \tag{1.11}$$

An easy calculation reveals that

$$d_{\text{MS}}(k) < d_{\text{con}}(k) \qquad \text{for every } k \geq 2. \tag{1.12}$$

Moreover, it is satisfactory that the formula (1.11) reproduces the correct (previously known) threshold $d_{\text{uniq}}(2) = d_{\text{con}}(2) = d_{\text{pure}}(2) = 2$. That said, we have no reason to believe that (1.11) is tight for any $k \geq 3$.

k	2	3	4	5
d_{giant}	1.0000	0.5000	0.3333	0.2500
d_{MS}	1.1625	0.8792	0.8695	0.9236
d_{con}	2.0000	1.3431	1.2451	1.2635
d_{pure}	2.0000	4.9108	6.1782	7.0178
d_{sat}	2.0000	12.801	39.724	105.585

TABLE 1. The values of $d_{\text{MS}}(k), d_{\text{con}}(k)$, and $d_{\text{pure}}(k)$ for $2 \leq k \leq 5$. Additionally, $d_{\text{giant}}(k) = 1/(k-1)$ marks the giant component threshold of the hypergraph induced by the random k-CNF formula. Moreover, $d_{\text{sat}}(k)$ is the satisfiability threshold according to physics predictions [49]. It is not hard to show that $d_{\text{giant}}(k) \leq d_{\text{MS}}(k) \leq d_{\text{con}}(k) \leq d_{\text{uniq}}(k) \leq d_{\text{pure}}(k) \leq d_{\text{sat}}(k)$, for all $k \geq 2$.

Combining Theorems 1.1 and 1.2, we obtain the following.

Corollary 1.3. *Let $k \geq 3$. If $d < d_{\text{con}}(k)$ then (1.7) holds.*

Corollary 1.3 constitutes the first rigorous result to determine the precise asymptotic value of $\log Z(\mathbf{\Phi})$ for a non-trivial regime of d for any $k \geq 3$. To elaborate, the formula (1.7) is trivially true for $d < 1/(k-1)$ because for such d the k-uniform hypergraph induced by the clauses of $\mathbf{\Phi}$ has no giant component and Belief Propagation is exact on acyclic graphical models [50]. But Corollary 1.3 applies to d well beyond this threshold, as displayed in Table 1. In particular, in contrast to much of the prior work on random k-SAT, Corollary 1.3 applies to a non-trivial regime of d even for 'small' $k \geq 3$.

Although Table 1 contains the values $d_{\text{MS}}(k)$ from [56] for comparison, we emphasise that Montanari and Shah's result only yields the number of 'good' assignments satisfying all but $o(n)$ clauses, rather than of actual satisfying assignments. In fact, the best prior rigorous bounds on the number of satisfying assignments for $d > 1/(k-1)$ derive from the first and the second moment methods. Specifically, the folklore first moment bound reads

$$\frac{1}{n} \log Z(\mathbf{\Phi}) \leq \log 2 + \frac{d}{k} \log(1 - 2^{-k}) + o(1) \quad \text{w.h.p.} \tag{1.13}$$

Furthermore, Achlioptas and Peres [7] perform a second moment argument on the number of *balanced* satisfying assignments, i.e., satisfying assignments that enjoy a peculiar additional condition required to keep the second moment under control. They show that w.h.p.

$$\frac{1}{n} \log Z(\mathbf{\Phi}) \geq (1-d)\log 2 + \frac{d}{k} \log\left[\left(\lambda^{1/2} + \lambda^{-1/2}\right)^k - \lambda^{-k/2}\right] + o(1), \quad \text{where} \quad (1-\lambda)(1+\lambda)^{k-1} = 1, \lambda > 0. \tag{1.14}$$

Figure 1 illustrates the bounds (1.13)–(1.14) along with (1.7) for $k = 3$. As the figure shows, the correct value (1.7) is quite close to the first moment bound. That said, the first moment bound strictly exceeds $\mathfrak{B}_{d,k}(\pi_{d,k})$ for all $d > 0$, $k \geq 3$ [24]. On the other hand, Figure 1 demonstrates that the 'balanced second moment bound' (1.14) significantly undershoots $\mathfrak{B}_{d,3}(\pi_{d,3})$. Recall that Figure 1 is on a logarithmic scale; thus, even small differences translate into exponentially large errors.

1.3. **Preliminaries and notation.** Let Φ be a Boolean expression in conjunctive normal such that no clause contains the same variable twice. We write $V(\Phi)$ for the set of Boolean variables of Φ and $F(\Phi)$ for the set of clauses. The formula Φ gives rise to a bipartite graph $G(\Phi)$ on the vertex set $V(\Phi) \cup F(\Phi)$ in which a variable x and a clause a are adjacent iff variable x appears in clause a (either positively or negatively). Let $E(\Phi)$ denote the edge set of the graph $G(\Phi)$. Furthermore, for a vertex $v \in V(\Phi) \cup F(\Phi)$ let $\partial_\Phi v$ be the set of neighbours of v; where the reference to Φ is self-evident, we just write ∂v.

The graph $G(\Phi)$ induces a metric on $V(\Phi) \cup F(\Phi)$ by letting $\mathrm{dist}_\Phi(v, w)$ equal the length of the shortest path from v to w. For a vertex v and an integer $\ell \geq 0$ let $\partial_\Phi^\ell v = \partial^\ell v$ be the set of all vertices w at distance precisely ℓ from v.

For a clause a and a variable $x \in \partial a$ we define $\mathrm{sign}_\Phi(x, a) = 1$ if a contains x as a positive literal, and $\mathrm{sign}_\Phi(x, a) = -1$ if a contains the negation $\neg x$. (This is unambiguous because clause a is not allowed to contain both x and $\neg x$.) For a variable $x \in V(\Phi)$ and $s \in \{\pm 1\}$ we let $\partial_\Phi^s x = \partial^s x$ be the set of clauses $a \in \partial_\Phi x$ such that $\mathrm{sign}_\Phi(x, a) = s$. Where convenient we use the shorthand $\partial^\pm x = \partial^{\pm 1} x$. We say that a variable x is *pure* in Φ if $\mathrm{sign}_\Phi(x, a) = \mathrm{sign}_\Phi(x, b)$ for all $a, b \in \partial x$. More specifically, say that x is a *pure literal* of Φ if $\partial^- x = \emptyset$. Similarly, $\neg x$ is called a pure literal if $\partial^+ x = \emptyset$. A variable or literal that fails to be pure is called *mixed*.

For a literal $l \in \{x, \neg x : x \in V(\Phi)\}$ we let $|l|$ denote the underlying variable; thus, $|x| = |\neg x| = x$ for $x \in V(\Phi)$. Moreover, we define $\mathrm{sign}(x) = 1$ and $\mathrm{sign}(\neg x) = -1$. Further, for a literal l we define $1 \cdot l = l$ and $(-1) \cdot l = \neg l$.

If Φ is satisfiable, then $\boldsymbol{\sigma}_\Phi = (\boldsymbol{\sigma}_\Phi(x))_{x \in V(\Phi)}$ denotes a uniformly random satisfying assignment of Φ. Where the reference to Φ is obvious we just write $\boldsymbol{\sigma}$.

Let μ, ν be two probability measures on $\mathbb{R}^h$, let $q \geq 1$ and assume that $\int_{\mathbb{R}^h} \|x\|_q^q \mathrm{d}\mu(x), \int_{\mathbb{R}^h} \|x\|_q^q \mathrm{d}\nu(x) < \infty$. We recall that the L_q*-Wasserstein distance* of μ, ν is defined as

$$W_q(\mu, \nu) = \inf_{(\boldsymbol{\xi}, \boldsymbol{\zeta})} \mathbb{E}\left[\|\boldsymbol{\xi} - \boldsymbol{\zeta}\|_q^q\right]^{1/q},$$

where the infimum is taken over all pairs $(\boldsymbol{\xi}, \boldsymbol{\zeta})$ of random variables defined on the same probability space Ω such that $\boldsymbol{\xi}$ has distribution μ and $\boldsymbol{\zeta}$ has distribution ν. If $\boldsymbol{X}, \boldsymbol{Y}$ are random variables with distributions μ, ν, it is convenient to use the shorthand $W_q(\boldsymbol{X}, \boldsymbol{Y}) = W_q(\mu, \nu)$, provided that $\mathbb{E}[\|\boldsymbol{X}\|_q^q], \mathbb{E}[\|\boldsymbol{Y}\|_q^q] < \infty$.

For two random variables $\boldsymbol{X}, \boldsymbol{Y}$ we write $\boldsymbol{X} \sim \boldsymbol{Y}$ if $\boldsymbol{X}, \boldsymbol{Y}$ are identically distributed. Moreover, for a probability distribution μ and a random variable $\boldsymbol{X}$ we write $\boldsymbol{X} \sim \mu$ if $\boldsymbol{X}$ has distribution μ.

We will make repeated use of the following tail bound for Poisson variables.

Lemma 1.4 (Bennett's inequality [14, Theorem 2.9])**.** *Suppose that* $\boldsymbol{X} \sim \mathrm{Po}(\lambda)$ *with* $\lambda > 0$ *and let* $\varphi(x) = (1+x)\log(1+x) - x$ *for* $x > -1$. *Then*

$$\mathbb{P}[\boldsymbol{X} \geq \lambda + t] \leq \exp(-\lambda\varphi(t/\lambda)) \qquad \text{for any } t > 0,$$
$$\mathbb{P}[\boldsymbol{X} \leq \lambda - t] \leq \exp(-\lambda\varphi(-t/\lambda)) \qquad \text{for any } 0 < t < \lambda.$$

For reals a, b we write

$$a \vee b = \max\{a, b\}, \qquad a \wedge b = \min\{a, b\}.$$

Unless specified otherwise asymptotic notation $o(\cdot)$, $O(\cdot)$, etc. is understood to refer to the limit $n \to \infty$. The symbol $\tilde{O}(\cdot)$ is understood to swallow $\mathrm{polylog}(n)$ terms. Throughout we tacitly assume that n is sufficiently large so that the various estimates are valid. We use the conventions $\log 0 = -\infty$ and $\log \infty = \infty$. Finally, throughout the paper we assume that $k \geq 3$ is a fixed integer.

2. Overview

In this section we survey the proofs of the main results. Subsequently, we discuss further related work. The proof details are deferred to the remaining sections; see Section 2.7 for pointers. We assume throughout that $k \geq 3$.

2.1. **Existence of the fixed point and upper bound.** As a first step towards the proof of Theorem 1.1 we prove that the limit (1.6) exists for $d < d_{\mathrm{uniq}}(k)$. More precisely, we will establish the following statement.

Proposition 2.1. *For every* $k \geq 3$ *and every* $d < d_{\mathrm{uniq}}(k)$, *the* W_1*-limit* $\pi_{d,k} = \lim_{\ell \to \infty} \mathrm{BP}_{d,k}^\ell(\delta_{1/2})$ *exists and*

$$\mathbb{E}\left[\log^2 \boldsymbol{\mu}_{\pi_{d,k},1,1}\right] + \mathbb{E}\left|\log\left(\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\pi_{d,k},2i} + \prod_{i=1}^{\boldsymbol{d}^+} \boldsymbol{\mu}_{\pi_{d,k},2i-1}\right)\right| + \mathbb{E}\left|\log\left(1 - \prod_{j=1}^{k} \boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right| < \infty. \tag{2.1}$$

In addition, $\boldsymbol{\mu}_{\pi_{d,k},1,1}$ *and* $1 - \boldsymbol{\mu}_{\pi_{d,k},1,1}$ *are identically distributed.*

The existence of the limit $\pi_{d,k}$ is an easy consequence of the Gibbs uniqueness property. As an aside, the limit $\pi_{d,k} = \lim_{\ell\to\infty} \mathrm{BP}^{\ell}_{d,k}(\delta_{1/2})$ is a fixed point of the Belief Propagation operator, i.e.,

$$\pi_{d,k} = \mathrm{BP}_{d,k}(\pi_{d,k}). \tag{2.2}$$

The proof of the bound (2.1) is a bit more subtle and requires a few preparations, but we will come to that. The upshot of (2.1) is that the Bethe free entropy $\mathfrak{B}_{d,k}(\pi_{d,k})$ is well defined.

With the fixed point $\pi_{d,k}$ in hand we can bring to bear the 'interpolation method' to the upper bound the likely value of $\log Z(\mathbf{\Phi})$.

Corollary 2.2. *If $d < d_{\mathrm{uniq}}(k)$ then w.h.p. we have $\frac{1}{n}\log Z(\mathbf{\Phi}) \le \mathfrak{B}_{d,k}(\pi_{d,k}) + o(1)$.*

The interpolation method is a mainstay of the study of disordered systems in mathematical physics and has also been used to investigate random constraint satisfaction problems. In particular, the variant of the interpolation method from [60] (in combination with Proposition 2.1) easily implies that

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\log(Z(\mathbf{\Phi}) \vee 1)\right] \le \mathfrak{B}_{d,k}(\pi_{d,k}) \ ;$$

taking the logarithm of $Z(\mathbf{\Phi}) \vee 1$ ensures that the expectation is well defined, as it is possible (albeit unlikely for $d < d_{\mathrm{uniq}}(k)$) that $\mathbf{\Phi}$ is unsatisfiable. The added value of Corollary 2.2 is that we obtain a bound that holds *with high probability*, rather than just a bound on the expectation. The interpolation method was used in [3] in a similar fashion to prove a 'with high probability' bound on the number of satisfying assignments of random 2-CNFs. The proof of Corollary 2.2 is an adaptation of that argument to $k \ge 3$.

2.2. **A matching lower bound.** The key step towards Theorem 1.1 is to establish a lower bound on $\log Z(\mathbf{\Phi})$ that matches the upper bound from Corollary 2.2. To accomplish this task we employ a coupling argument known as the 'Aizenman-Sims-Starr scheme' in mathematical physics. Its original version was intended to estimate the partition function of the Sherrington-Kirkpatrick model, a spin glass model [8]. But the technique has since been employed in probabilistic combinatorics (e.g., [24, 25, 65]). By comparison to the mathematical physics context, the crucial difference is that here our objective is to count actual satisfying assignments where every single clause imposes a hard constraint, whereas in spin glass theory constraints are soft. The same issue occurred in previous work on the random 2-SAT problem [3]. However, in that case a relatively simple percolation argument was sufficient to deal with the ensuing complications. As we will see, for $k \ge 3$ considerably more care is needed.

But first things first. The basic idea behind the Aizenman-Sims-Starr argument is to perform a kind of induction. Translated to random k-SAT this means that we couple the random k-CNF $\mathbf{\Phi}_{d,k}(n)$ with n variables with the random k-CNF $\mathbf{\Phi}_{d,k}(n+1)$ with $n+1$ variables. Recall that $\mathbf{\Phi}_{d,k}(n)$ comprises $\boldsymbol{m}_n \sim \mathrm{Po}(dn/k)$ independent random clauses. Ultimately Theorem 1.1 is going to be a consequence of Corollary 2.2 and the following statement.

Proposition 2.3. *If $d < d_{\mathrm{uniq}}(k)$ then $\mathbb{E}\left[\log(Z(\mathbf{\Phi}_{d,k}(n+1)) \vee 1)\right] - \mathbb{E}\left[\log(Z(\mathbf{\Phi}_{d,k}(n)) \vee 1)\right] = \mathfrak{B}_{d,k}(\pi_{d,k}) + o(1)$.*

Once again we work with $Z(\mathbf{\Phi}_{d,k}(n)) \vee 1$ and $Z(\mathbf{\Phi}_{d,k}(n+1)) \vee 1$ to ensure that the expectations are well defined.

To prove Proposition 2.3 we couple the random formulas $\mathbf{\Phi}_{d,k}(n+1)$ and $\mathbf{\Phi}_{d,k}(n)$ as follows.

CPL1: Let $\mathbf{\Phi}'$ be a random k-CNF with variables $x_1,\dots,x_n$ and $\boldsymbol{m}' \sim \mathrm{Po}(d(n-k+1)/k)$ clauses.

CPL2: Obtain $\mathbf{\Phi}''$ from $\mathbf{\Phi}'$ by adding another $\mathbf{\Delta}'' \sim \mathrm{Po}(d(k-1)/k)$ independent random clauses.

CPL3: Obtain $\mathbf{\Phi}'''$ from $\mathbf{\Phi}'$ by adding one new variable x_{n+1} and $\mathbf{\Delta}''' \sim \mathrm{Po}(d)$ independent random clauses that each contain x_{n+1} and $k-1$ other variables from $\{x_1,\dots,x_n\}$.

Observe that $\mathbf{\Phi}''$ ultimately has variables $x_1,\dots,x_n$ and a total of $\boldsymbol{m}_n \sim \mathrm{Po}(dn/k)$ random clauses. Thus, $\mathbf{\Phi}''$ is identical to the random formula $\mathbf{\Phi}_{d,k}(n)$. Similarly, $\mathbf{\Phi}'''$ has the same distribution as $\mathbf{\Phi}_{d,k}(n+1)$. Consequently, we obtain the following.

Fact 2.4. *For any $d > 0$ we have $Z(\mathbf{\Phi}_{d,k}(n)) \sim Z(\mathbf{\Phi}'')$ and $Z(\mathbf{\Phi}_{d,k}(n+1)) \sim Z(\mathbf{\Phi}''')$.*

The coupling **CPL1–CPL3** reduces the proof of Proposition 2.3 to getting a handle on the differences $\log(Z(\mathbf{\Phi}'')\vee 1) - \log(Z(\mathbf{\Phi}')\vee 1)$ and $\log(Z(\mathbf{\Phi}''')\vee 1) - \log(Z(\mathbf{\Phi}')\vee 1)$. More precisely, recalling (1.4)–(1.5), we see that Proposition 2.3 is a consequence of the following two statements.

Proposition 2.5. *If $d < d_{\mathrm{uniq}}(k)$ then*

$$\mathbb{E}\left[\log \frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right] = \frac{d(k-1)}{k}\mathbb{E}\left[\log\left(1 - \prod_{j=1}^{k} \boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right] + o(1). \tag{2.3}$$

Proposition 2.6. *If $d < d_{\mathrm{uniq}}(k)$ then*

$$\mathbb{E}\left[\log\frac{Z(\mathbf{\Phi}''')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right] = \mathbb{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}^-}\boldsymbol{\mu}_{\pi_{d,k},2i}+\prod_{i=1}^{\boldsymbol{d}^+}\boldsymbol{\mu}_{\pi_{d,k},2i-1}\right)\right]+o(1). \tag{2.4}$$

To prove Propositions 2.5–2.6 we effectively need to trace the impact that *local* changes have on the number of satisfying assignments. Indeed, under the coupling **CPL1–CPL3**, the formula $\mathbf{\Phi}''$ is obtained from the 'base formula' $\mathbf{\Phi}'$ by adding just a bounded expected number of random clauses. Thus, if we imagine that, as both the first moment upper bound (1.13) and the balanced second moment lower bound (1.14) suggest, each additional random clause typically reduces the number of satisfying assignments by a constant factor, then the quantity $|\log(Z(\mathbf{\Phi}'')/Z(\mathbf{\Phi}'))|$ should be bounded with probability close to one. Similar reasoning applies to $\mathbf{\Phi}'''$.

Yet while with high probability the local changes that turn $\mathbf{\Phi}'$ into $\mathbf{\Phi}''$ or $\mathbf{\Phi}'''$ are indeed benign, because we are dealing with hard constraints there is a non-negligible probability that $\log(Z(\mathbf{\Phi}'')/Z(\mathbf{\Phi}'))$ and $\log(Z(\mathbf{\Phi}''')/Z(\mathbf{\Phi}'))$ could be large. Indeed, a single extra clause might wipe out all satisfying assignments of $\mathbf{\Phi}'$, in which case

$$\log\frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1} = -\log Z(\mathbf{\Phi}') = -\Omega(n).$$

Hence, we need to argue that such drastic changes are sufficiently rare. The following statement furnishes the necessary tail bound.

Proposition 2.7. *For $d < d_{\mathrm{uniq}}(k)$ we have*

$$\mathbb{E}\left[\left|\log\frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}+\left|\log\frac{Z(\mathbf{\Phi}''')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right] = O(1). \tag{2.5}$$

2.3. **Pure literal pursuit.** The proof of Proposition 2.7 constitutes the main technical challenge towards the proof of Theorem 1.1. The linchpin of the proof is an algorithm that we call *Pure Literal Pursuit* ('PULP'). Its purpose is to trace the repercussions of setting a relatively small number of variables to specific truth values. More precisely, PULP will allow us to compare the number of satisfying assignments that set a few chosen variables to specific values to the total number of satisfying assignments.

To this end PULP attempts to solve the following optimisation task. Suppose we are given a k-CNF Φ and a set $\mathscr{L}$ of literals of Φ that we deem to be set to 'true'. We would like to identify a superset $\bar{\mathscr{L}} \supseteq \mathscr{L}$ of literals with the following properties; think of $\bar{\mathscr{L}}$ as a 'closure' of $\mathscr{L}$.

- **PULP1:** every clause a of Φ that contains a literal from $\neg\bar{\mathscr{L}} = \{\neg l : l \in \bar{\mathscr{L}}\}$ also contains a literal from $\bar{\mathscr{L}}$.
- **PULP2:** there is no literal l such that $l, \neg l \in \bar{\mathscr{L}}$.

Of course, it may be impossible to satisfy **PULP1** and **PULP2** simultaneously. In this case we ask PULP to report a 'contradiction'. But if **PULP1–PULP2** can be satisfied, we aim to find a closure $\bar{\mathscr{L}}$ of as small size $|\bar{\mathscr{L}}|$ as possible.

The combinatorial idea behind **PULP1–PULP2** is as follows. Deeming the literals from the initial set $\mathscr{L}$ 'true', our goal is to reconcile this assumption with the formula Φ. To this end we enhance the set $\mathscr{L}$. Clearly, any clause that contains the negation $\neg l$ of a literal l that we deem true also needs to contain another literal l' that is set to true. This is what **PULP1** asks. Furthermore, it would be contradictory to deem both l and its negation $\neg l$ true; this is **PULP2**.

The size of the closure $\bar{\mathscr{L}}$ yields a bound on the reduction in the number of satisfying assignments if we indeed insist on all literals $l \in \mathscr{L}$ being set to true. Formally, let $S(\Phi, \mathscr{L})$ be the set of all satisfying assignments $\sigma \in S(\Phi)$ under which all literals $l \in \mathscr{L}$ evaluate to 'true'. Also set $Z(\Phi, \mathscr{L}) = |S(\Phi, \mathscr{L})|$.

Lemma 2.8. *For any $\Phi, \mathscr{L}$ and any $\bar{\mathscr{L}} \supseteq \mathscr{L}$ that satisfies* **PULP1–PULP2** *we have $Z(\Phi) \leq 2^{|\bar{\mathscr{L}}|} Z(\Phi, \mathscr{L})$.*

In order to identify a 'small' closure $\bar{\mathscr{L}}$ the PULP algorithm resorts to *pure literal elimination*, a simple trick commonplace to satisfiability algorithms. A variable x is *pure* in a CNF formula Φ if $\mathrm{sign}(x, a) = \mathrm{sign}(x, b)$ for any two clauses $a, b \in \partial x$. Clearly, if our objective is to construct a satisfying assignment, we might as well set all pure variables x to the value that satisfies all clauses $a \in \partial x$ and disregard these clauses henceforth. In light of this observation, pure literal elimination repeatedly removes all clauses that contain a pure variable. Naturally, every round of clause removals may create new pure variables, and thus more clauses may be ripe for removal in the next round. For a clause a of the original formula Φ let $\mathfrak{h}_a(\Phi) \geq 1$ be the number of the round at which pure literal elimination removes a. If a is never removed then we set $\mathfrak{h}_a(\Phi) = \infty$.

The PULP algorithm invokes a slightly modified version of pure literal elimination to accommodate the initial set $\mathscr{L}$ of literals. Specifically, for a variable x of a CNF Φ and $s \in \{\pm 1\}$ let $\Phi[x \mapsto s]$ be the CNF obtained by removing all clauses $a \in \partial x$ with $\mathrm{sign}(x,a) = s$ and removing the literal $-s \cdot x$ from all $a \in \partial x$ with $\mathrm{sign}(x,a) = -s$. The definition reflects that if we set x to value s, all $a \in \partial^s x$ will be satisfied, while all $a \in \partial^{-s} x$ will have to be satisfied by one of their other constituent literals. Further, let

$$\mathfrak{h}_x(s,\Phi) = \begin{cases} 0 & \text{if } \partial_\Phi^{-s} x = \emptyset, \\ \max\left\{\mathfrak{h}_a(\Phi[x \mapsto s]) : a \in \partial_\Phi^{-s} x\right\} & \text{otherwise.} \end{cases} \in [0,\infty]. \tag{2.6}$$

We refer to $\mathfrak{h}_x(s,\Phi)$ as the *height of literal $s \cdot x$ in Φ*.

The PULP algorithm, displayed as Algorithm 1, harnesses the heights as follows. In its attempt to precipitate **PULP1** and **PULP2** the algorithm iteratively enhances the set $\mathscr{L}$ of literals deemed to be 'true'. For any clause a that violates **PULP1** and that contains a literal $l \notin \neg\mathscr{L}$ the algorithm adds one such literal l *of minimum height* to $\mathscr{L}$. This choice is intended to keep the ultimate size of the closure small; one could say that PULP uses height as a proxy of 'size'. If at any point the algorithm encounters a clause a that consists of literals from $\neg\mathscr{L}$ only, the algorithm reports a contradiction and aborts.

Input: A k-CNF Φ and a set $\mathscr{L}$ of literals.
1 Let $\bar{\mathscr{L}} = \mathscr{L}$;
2 **while** *there is a clause a that contains a literal from $\neg\bar{\mathscr{L}}$ but no literal from $\bar{\mathscr{L}}$* **do**
3   Pick such a clause a that minimises the distance from the initial set $\mathscr{L} = \{|l| : l \in \mathscr{L}\}$;
4   **if** *a consists of literals $l \in \neg\bar{\mathscr{L}}$ only* **then**
5   return 'contradiction' and halt;
6   **else**
7   choose $x \in \partial a$ with $x, \neg x \notin \bar{\mathscr{L}}$ that minimises $\mathfrak{h}_x(\mathrm{sign}(x,a),\Phi)$ and add $\mathrm{sign}(x,a) \cdot x$ to $\bar{\mathscr{L}}$;
8 **return** $\bar{\mathscr{L}}$

Algorithm 1: The PULP algorithm

Remark 2.9. *To break ties that may occur in the execution of Steps 3 and 7 of* PULP *we assume that the variables and clauses of Φ are numbered so that Steps 3 and 7 can choose the clause/variable with the smallest number that satisfies the respective requirements. In due course we will run* PULP *on (finite subtrees of) the Galton-Watson tree $\mathbb{T}$. To number the variables and clauses of $\mathbb{T}$ we equip each of them with an independent Gaussian label. Since $\mathbb{T}$ comprises a countable number of clauses/variables, these labels will almost surely be pairwise distinct.*

From here on we write $\bar{\mathscr{L}}$ for the set of literals returned by PULP if the algorithm does not encounter a contradiction; in the event of a a contradiction we let $\bar{\mathscr{L}} = \{x, \neg x : x \in V(\Phi)\}$ be the set of all literals. Where the reference to the formula Φ is not entirely obvious, we write $\bar{\mathscr{L}}_\Phi$. The analysis of PULP on the random formula $\mathbf{\Phi}'$ furnishes the following bound on $|\bar{\mathscr{L}}|$ in terms of the size of the initial set $\mathscr{L}$. This bound is the key ingredient towards the proof of Proposition 2.7.

Lemma 2.10. *There exists $C = C(d,k) > 0$ such that the following is true. Let $\mathscr{L}$ be a set of literals of $\mathbf{\Phi}'$ such that $1 \le |\mathscr{L}| \le \log^2 n$ and such that $\{x_i, \neg x_i\} \not\subseteq \mathscr{L}$ for all $1 \le i \le n$. Then $\mathbb{E}[|\bar{\mathscr{L}}|^{3/2}] \le C|\mathscr{L}|^{3/2}$.*

The proof of Lemma 2.10 is one of the main technical challenges of the present work. The difficulty stems from the stochastic dependencies that are inherent to the PULP algorithm. Specifically, in order to decide which literals to add to the set $\mathscr{L}$, PULP requires knowledge of the heights $\mathfrak{h}_x(\pm 1, \mathbf{\Phi}')$. But these heights depend on the other variables $y \in \partial a \setminus \{x\}$, the clauses that these variables y appear in, etc. Furthermore, in its subsequent iterations the algorithm is apt to revisit some of these variables and clauses at a point when their heights have already been revealed. These repetitions rule out an analysis of PULP by way of routine techniques such as the principle of deferred decisions or the differential equations method. The reason why we manage to cope with these complicated dependencies at all is that, remarkably, the heights $\mathfrak{h}_x(\pm 1, \mathbf{\Phi}')$ have only a tiny upper tail. More precisely, as we will see the tails of these random variables decay at a *doubly* exponential rate.

Proposition 2.7 follows from the analysis of PULP. The basic idea is to apply the algorithm to an initial set $\mathscr{L}$ of literals that contain one literal from each of the extra clauses that are present in $\mathbf{\Phi}''$ or $\mathbf{\Phi}'''$ but not in $\mathbf{\Phi}'$. With a bit of care the bounds from Lemmas 2.8 and 2.10 then imply (2.5). Finally, the analysis of PULP that leads up to the proof of Lemma 2.10 also implies the necessary tail bounds to verify the bounds from (2.1). Specifically, the proof of Lemma 2.10 proceeds by way of analysing PULP on the Galton-Watson tree $\mathbb{T}_{d,k}$, and the bounds (2.1) come out as a byproduct of that analysis.

2.4. **Completing the Aizenman-Sims-Starr scheme.** To obtain Propositions 2.5–2.6 we combine Proposition 2.7 with an analysis of the quotients $Z(\mathbf{\Phi}'')/Z(\mathbf{\Phi}')$ and $Z(\mathbf{\Phi}''')/Z(\mathbf{\Phi}')$ on a likely 'good' event. On this good event the empirical distribution of the marginal probabilities $(\mathbb{P}[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i)=1\mid\mathbf{\Phi}'])_{1\le i\le n}$ of the different variables x_i receiving the value 'true' under a random satisfying assignment is 'close' to the limiting distribution $\pi_{d,k}$ from Proposition 2.1. Additionally, on the good event the joint distribution of the truth values assigned to a moderate number of variables is well approximated by a product measure. Of course, to make this precise we need to investigate the empirical distribution

$$\boldsymbol{\pi}'_n=\frac{1}{n}\sum_{i=1}^n\delta_{\mathbb{P}[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i)=1\mid\mathbf{\Phi}']}\in\mathscr{P}(0,1) \tag{2.7}$$

of the marginals $(\mathbb{P}[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i)=1\mid\mathbf{\Phi}'])_{1\le i\le n}$.

Proposition 2.11. *Assume that $d<d_{\mathrm{uniq}}(k)$. Then $\mathbb{E}\left[W_1(\boldsymbol{\pi}'_n,\pi_{d,k})\right]=o(1)$ and for any $\ell=O(1)$ we have*

$$\sum_{\sigma\in\{\pm1\}^\ell}\mathbb{E}\left|\mathbb{P}\left[\forall 1\le i\le\ell:\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i)=\sigma_i\mid\mathbf{\Phi}'\right]-\prod_{i=1}^{\ell}\mathbb{P}\left[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i)=\sigma_i\mid\mathbf{\Phi}'\right]\right|=o(1).$$

The proof of Proposition 2.11 hinges on the Gibbs uniqueness property and the convergence of the local topology of the random formula $\mathbf{\Phi}'$ to the Galton-Watson tree $\mathbb{T}_{d,k}$. Together with careful coupling arguments Propositions 2.7–2.11 imply Propositions 2.5–2.6. Moreover, in combination with Fact 2.4 these two propositions yield Proposition 2.3. We complete this paragraph by showing how Theorem 1.1 follows from Corollary 2.2 and Proposition 2.3.

Proof of Theorem 1.1. The existence of the limit (1.6) follows from Proposition 2.1. With respect to (1.7), we apply Proposition 2.3 to obtain

$$\begin{aligned}\frac{1}{n}\mathbb{E}\left[\log(1\vee Z(\mathbf{\Phi}_{d,k}(n)))\right]&=\frac{1}{n}\sum_{N=0}^{n-1}\left(\mathbb{E}\left[\log(1\vee Z(\mathbf{\Phi}_{d,k}(N+1))\right]-\mathbb{E}\left[\log(1\vee Z(\mathbf{\Phi}_{d,k}(N))\right]\right)\\&=\mathfrak{B}_{d,k}(\pi_{d,k})+o(1)\ .\end{aligned} \tag{2.8}$$

Since, conversely, Corollary 2.2 shows that $\frac{1}{n}\log Z(\mathbf{\Phi})\le\mathfrak{B}_{d,k}(\pi_{d,k})+o(1)$ w.h.p. and since $\log Z(\mathbf{\Phi})\le n\log 2$ deterministically, the assertion follows from (2.8). □

2.5. **Lower-bounding the uniqueness threshold.** The proof of Theorem 1.2 combines three ingredients. From the work [3] on random 2-SAT we borrow the idea of constructing an explicit extremal boundary configuration. In effect, in order to prove Gibbs uniqueness we just have to consider one single boundary configuration, instead of an enormous number of possible configurations τ that grows quickly with the height ℓ as in the original definition (1.1). Second, from the work [56] of Montanari and Shah we borrow the idea of expressly considering the effect of pure literals. As it turns out, without explicit consideration of pure literals it seems difficult to even recover the correct asymptotic order (1.8) of the Gibbs uniqueness threshold. Third, and most importantly, the improvement over the bound from [56] stems from a new subtle coupling argument that we will explain in due course.

2.5.1. *The extremal boundary condition.* An obvious challenge associated with establishing the Gibbs uniqueness property (1.1) seems to be that we need to estimate the marginal of the root variable given *any* possible boundary condition, i.e., given any assignment of the variables at distance 2ℓ from $\mathfrak{r}$. As we expect to see $(d(k-1))^\ell$ variables at distance 2ℓ from $\mathfrak{r}$, we thus face a doubly exponential number $2^{(d(k-1))^\ell}$ of possible boundary conditions. But fortunately, following [3] we may confine ourselves to just a single, explicit boundary configuration $\boldsymbol{\tau}^+$ that satisfies

$$\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(r)=1\mid\mathbb{T},\forall x\in\partial^{2\ell}r:\boldsymbol{\tau}^{(\ell)}(x)=\boldsymbol{\tau}^+(x)\right]=\max_{\tau\in S(\mathbb{T}^{(\ell)})}\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(r)=1\mid\mathbb{T},\forall x\in\partial^{2\ell}r:\boldsymbol{\tau}^{(\ell)}(x)=\tau(x)\right]. \tag{2.9}$$

Due to the inherent symmetry of the distribution of $\mathbb{T}$ with respect to the signs of the clauses, towards the proof of (1.1) it is sufficient to show that the difference (2.9) vanishes as $\ell\to\infty$.

The extremal boundary condition can be constructed explicitly. Specifically, given $\mathbb{T}^{(\ell)}$ we construct a satisfying assignment $\boldsymbol{\tau}^+ \in S(\mathbb{T}^{(\ell)})$ by working our way down the tree $\mathbb{T}^{(\ell)}$. We begin by setting $\boldsymbol{\tau}^+(\mathfrak{x}) = 1$. Now suppose that for $q \geq 1$, the values of the variables at distance $2(q-1)$ from $\mathfrak{x}$ have been already determined. Let w be a variable at distance $2q$ from $\mathfrak{x}$ with parent clause a and grandparent variable u. Then we define

$$\boldsymbol{\tau}^+(w) = \mathrm{sign}(w,a) \cdot \mathbb{1}\{\mathrm{sign}(u,a) \neq \boldsymbol{\tau}^+(u)\} - \mathrm{sign}(w,a) \cdot \mathbb{1}\{\mathrm{sign}(u,a) = \boldsymbol{\tau}^+(u)\} \ . \tag{2.10}$$

The idea behind (2.10) is for $\boldsymbol{\tau}^+(w)$ to "nudge" u towards $\boldsymbol{\tau}^+(u)$ by making sure that w satisfies clause a if setting u to $\boldsymbol{\tau}^+(u)$ does not, and conversely making sure that w fails to satisfy clause a if setting u to $\boldsymbol{\tau}^+(u)$ does. A simple induction on ℓ shows that $\boldsymbol{\tau}^+$ is a satisfying assignment for which (2.9) holds.

Lemma 2.12. *For any integer $\ell \geq 0$ the assignment $\boldsymbol{\tau}^+$ defined via (2.10) satisfies (2.9).*

Hence, proving Theorem 1.2 reduces to establishing the following.

Proposition 2.13. *For $d < d_{\mathrm{con}}(k)$ we have that*

$$\lim_{\ell\to\infty} \mathbb{E}\left[\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{x}) = 1 \mid \mathbb{T}, \forall x \in \partial^{2\ell}\mathfrak{x} : \boldsymbol{\tau}^{(\ell)}(x) = \boldsymbol{\tau}^+(x)\right] - \mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{x}) = 1 \mid \mathbb{T}\right]\right] = 0. \tag{2.11}$$

The proof of Proposition 2.13 may seem delicate because the boundary condition $\boldsymbol{\tau}^+$ depends on the tree $\mathbb{T}^{(\ell)}$. To sidestep this problem, we generalise another idea from the work [3] on random 2-SAT to $k \geq 3$ by introducing a quantity that allows us to prove (2.13) but that behaves 'Markovian' as we pass up and down the tree. Specifically, for a variable x of $\mathbb{T}^{(\ell)}$ let $\mathbb{T}_x^{(\ell)}$ be the sub-formula of $\mathbb{T}^{(\ell)}$ comprising x and its progeny. Moreover, for a satisfying assignment $\tau \in S(\mathbb{T}^{(\ell)})$ let

$$S(\mathbb{T}_x^{(\ell)}, \tau) = \left\{\chi \in S(\mathbb{T}_x^{(\ell)}) : \forall y \in V(\mathbb{T}_x^{(\ell)}) \cap \partial_{\mathbb{T}}^{2\ell}\mathfrak{x} : \chi_y = \tau_y\right\} \ , \qquad Z(\mathbb{T}_x^{(\ell)}, \tau) = \left|S(\mathbb{T}_x^{(\ell)}, \tau)\right| \ .$$

In words, $S(\mathbb{T}_x^{(\ell)}, \tau)$ contains all satisfying assignments of $\mathbb{T}_x^{(\ell)}$ that comply with the boundary condition induced by τ. Additionally, for $t = \pm 1$ let

$$S(\mathbb{T}_x^{(\ell)}, \tau, t) = \left\{\chi \in S(\mathbb{T}_x^{(\ell)}, \tau) : \chi_x = t\right\}, \qquad Z(\mathbb{T}_x^{(\ell)}, \tau, t) = \left|S(\mathbb{T}_x^{(\ell)}, \tau, t)\right|$$

be the set and number of satisfying assignments of $\mathbb{T}_x^{(\ell)}$ that agree with τ on the boundary and assign value t to x. Finally, let

$$\boldsymbol{\eta}_x^{(\ell)} = \log \frac{Z(\mathbb{T}_x^{(\ell)}, \boldsymbol{\tau}^+, \boldsymbol{\tau}^+(x))}{Z(\mathbb{T}_x^{(\ell)}, \boldsymbol{\tau}^+, -\boldsymbol{\tau}^+(x))} \in \mathbb{R} \cup \{\pm\infty\} \tag{2.12}$$

be the log-likelihood ratio that gauges how likely a random satisfying assignment $\boldsymbol{\tau}$ of $\mathbb{T}_x^{(\ell)}$ subject to the $\boldsymbol{\tau}^+$-boundary condition is to set x to its designated value $\boldsymbol{\tau}^+(x)$ from (2.10). In terms of (2.12), the proof of Proposition 2.13 comes down to showing that for $d < d_{\mathrm{con}}(k)$,

$$\lim_{\ell\to\infty} \boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)} = \log\left(\frac{\boldsymbol{\mu}_{\pi_{d,k},1,1}}{1 - \boldsymbol{\mu}_{\pi_{d,k},1,1}}\right) \qquad \text{in distribution.} \tag{2.13}$$

For a start, the following lemma bounds the tails of $\boldsymbol{\eta}_x^{(\ell)}$ for large enough ℓ and x reasonably close to the root variable $\mathfrak{x}$.

Lemma 2.14. *For every $0 < d < d_{\mathrm{con}}(k)$ there exist $c = c(d,k)$ and a sequence $(\varepsilon_t)_t$ with $\lim_{t\to\infty} \varepsilon_t = 0$ such that for any $t > 0$, $\ell > ct^c$ we have*

$$\mathbb{P}\left[\max_{x \in \partial^{2t}\mathfrak{x}} \left|\boldsymbol{\eta}_x^{(\ell)}\right| \leq 2t^c\right] > 1 - \varepsilon_t \ . \tag{2.14}$$

The proof of Lemma 2.14 rests on combinatorial arguments reminiscent of the analysis of PULP.

A key feature of the definition (2.12) is that the random variables $\boldsymbol{\eta}_x^{(\ell)}$ exhibit a 'reverse Markovian' behaviour. This is because $\boldsymbol{\eta}_x^{(\ell)}$ depends only on $\boldsymbol{\tau}^+(x)$ and the part $\mathbb{T}_x^{(\ell)}$ of the tree pending on x. Furthermore, because the distribution of the random tree $\mathbb{T}_x^{(\ell)}$ is symmetric with respect to sign flips, even the dependence on the value $\boldsymbol{\tau}^+(x)$ can be eliminated. All we need to keep in mind is that the values $\boldsymbol{\tau}^+(y)$ for $y \in V(\mathbb{T}_x^{(\ell)})$ are constructed from the value $\boldsymbol{\tau}^+(x)$ in accordance with the recurrence (2.10). Thus, by flipping all signs in the tree $\mathbb{T}_x^{(\ell)}$ if necessary, we could assume without loss that $\boldsymbol{\tau}^+(x) = 1$ without changing the distribution of $\boldsymbol{\eta}_x^{(\ell)}$ with respect to the randomness of $\mathbb{T}_x^{(\ell)}$. As a consequence, it is possible to set up a recurrence that expresses the log-likelihood ratios $\boldsymbol{\eta}_x^{(\ell)}$ of variables x at distance q from $\mathfrak{x}$ in terms of the $\boldsymbol{\eta}_y^{(\ell)}$ for y at distance $q+2$ from $\mathfrak{x}$.

Due to the recursive nature of the random tree $\mathbb{T}$, it suffices to set up this recurrence for the root $\mathfrak{x}$ of the tree. In other words, to prove (2.13) we just need a recurrence that expresses the distribution of the random variable $\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell+1)}$ in terms of the law of $\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)}$ for $\ell \geq 0$. A bit of reflection (see Claim 7.1), reveals that the corresponding distributional operator

$$\mathrm{LL}^+_{d,k} : \mathscr{P}((-\infty,\infty]) \to \mathscr{P}((-\infty,\infty]) \ , \qquad\qquad \rho \mapsto \hat{\rho} = \mathrm{LL}^+_{d,k}(\rho)$$

has the following shape. For a distribution $\rho \in \mathscr{P}((-\infty,\infty])$ let $(\boldsymbol{\eta}_{\rho,i,j})_{i,j\geq 1}$ be a family of random variables with distribution ρ. Moreover, let $(\boldsymbol{s}_i)_{i\geq 1}$ be a sequence of uniformly random ± 1-valued random variables and let $\boldsymbol{d} \sim \mathrm{Po}(d)$. All of these random variables are mutually independent. Additionally, for $q \geq 0$ and $z_1,\ldots,z_q \in \mathbb{R}\cup\{\pm\infty\}$ define

$$\Gamma(z_1,\ldots,z_q) = \prod_{i=1}^{q} \frac{1+\tanh(z_i/2)}{2} \ . \tag{2.15}$$

Then $\hat{\rho} = \mathrm{LL}^+_{d,k}(\rho)$ is the distribution of the random variable

$$-\sum_{i=1}^{\boldsymbol{d}} \boldsymbol{s}_i \cdot \log\Big(1-\Gamma\Big(\boldsymbol{s}_i\cdot(\boldsymbol{\eta}_{\rho,i,1},\ldots,\boldsymbol{\eta}_{\rho,i,k-1})\Big)\Big) \ . \tag{2.16}$$

Ultimately we will derive (2.13), and thereby Proposition 2.13, from Lemma 2.14 and a contraction argument. However, this is not quite as straightforward as one might be inclined to expect. Indeed, at first glance, a natural approach to proving (2.13) from Lemma 2.14 seems to be to show that $\mathrm{LL}^+_{d,k}$ is a contraction, say, with respect to the W_1-metric. This is indeed carried out in [3] for $k=2$, where it is shown that $\mathrm{LL}^+_{d,2}$ contracts for all $0<d<2$, i.e., right up to the random 2-SAT satisfiability threshold. However, for $k\geq 3$ we can only show that $\mathrm{LL}^+_{d,k}$ contracts for $d < 2/(k-1)$, a value well below $d_{\mathrm{con}}(k)$ and short of the correct asymptotic order (1.10).

2.5.2. *Pure and mixed literals.* To cover a larger range of d we borrow from [56] the idea of expressly taking into account pure literals. To elaborate, while $\mathrm{LL}^+_{d,k}$ describes how the law of $\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell+1)}$ results from that of $\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)}$, the operator fails to take into account that $\mathfrak{x}$ itself as well as some of the grandchildren of $\mathfrak{x}$ in $\mathbb{T}$ may be pure literals. However, the pure literal property has a marked effect on the log-likelihood ratios. For if, say, $\mathfrak{x}$ only appears positively, then a simple double counting argument shows that $\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)} \geq 0$ for all ℓ. By extension, pure literals among the grandchildren of $\mathfrak{x}$ have a 'dampening' effect and may thus improve the range of d for which we can establish contraction.

For a variable node x of $\mathbb{T}$, let us denote by $\mathbb{T}_x$ the subtree of $\mathbb{T}$ rooted at x and containing its progeny. Leveraging the above observation, we classify a variable x of $\mathbb{T}$ as $\bullet$, $\oplus$, $\ominus$, or $\circ$, depending on whether x appears both positively and negatively in $\mathbb{T}_x$, only positively, only negatively, or whether x has no children at all, respectively. Furthermore, instead of just tracing the law of $\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)}$ for $\ell \geq 0$, we study the four separate conditional distributions given the type $\bullet, \oplus, \ominus$ or $\circ$ of $\mathfrak{x}$. Of course, the distribution of $\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)}$ given type $\circ$ (i.e., $\mathfrak{x}$ has no children) is just the atom at zero for all ℓ.

To describe the evolution of the other distributions we introduce the operator

$$\begin{aligned} \mathrm{LL}^\star_{d,k} : \mathscr{P}(-\infty,\infty] \times \mathscr{P}(0,+\infty] \times \mathscr{P}(-\infty,0] &\to \mathscr{P}(-\infty,\infty] \times \mathscr{P}(0,+\infty] \times \mathscr{P}(-\infty,0] \ , \\ \text{with } (\rho_\bullet,\rho_\oplus,\rho_\ominus) &\mapsto (\hat\rho_\bullet,\hat\rho_\oplus,\hat\rho_\ominus) = \mathrm{LL}^\star_{d,k}(\rho_\bullet,\rho_\oplus,\rho_\ominus) \end{aligned} \tag{2.17}$$

defined as follows. Let $\boldsymbol{d}^\star_+, \boldsymbol{d}^{\star\prime}_+, \boldsymbol{d}^\star_-, \boldsymbol{d}^{\star\prime}_-$ be Poisson variables with parameter $d/2$, conditioned on being positive. Moreover, let $\boldsymbol{r}_1 = \big(\boldsymbol{r}_{\bullet,1}, \boldsymbol{r}_{\oplus,1}, \boldsymbol{r}_{\ominus,1}, \boldsymbol{r}_{\circ,1}\big)$, $\boldsymbol{r}_2 = \big(\boldsymbol{r}_{\bullet,2}, \boldsymbol{r}_{\oplus,2}, \boldsymbol{r}_{\ominus,2}, \boldsymbol{r}_{\circ,2}\big), \ldots$ be multinomial variables with $k-1$ trials and probabilities

$$p_\bullet = (1-e^{-d/2})^2, \qquad p_\oplus = p_\ominus = e^{-d/2}(1-e^{-d/2}), \qquad p_\circ = e^{-d} \ . \tag{2.18}$$

For $i, j \geq 1$ let $\boldsymbol{\eta}_{\bullet,i,j}$, $\boldsymbol{\eta}_{\oplus,i,j}$, $\boldsymbol{\eta}_{\ominus,i,j}$ be random variables with law $\rho_\bullet$, $\rho_\oplus$, $\rho_\ominus$, respectively. All of the aforementioned random variables are mutually independent. Further, for a sign $\varepsilon \in \{\pm 1\}$ and a vector $r = (r_\bullet, r_\oplus, r_\ominus, r_\bigcirc)$ of non-negative integers with $r_\bullet + r_\oplus + r_\ominus + r_\bigcirc = k-1$ and $i \geq 0$, $1 \leq j \leq 4$ we let

$$\Xi_{i,j}(\varepsilon, r) = 1 - \frac{1}{2^{r_\bigcirc}} \cdot \Gamma\left(\varepsilon\left(\boldsymbol{\eta}_{\bullet,4i+j,1}, \ldots, \boldsymbol{\eta}_{\bullet,4i+j,r_\bullet}\right)\right) \Gamma\left(\varepsilon\left(\boldsymbol{\eta}_{\oplus,4i+j,1}, \ldots, \boldsymbol{\eta}_{\oplus,4i+j,r_\oplus}\right)\right) \Gamma\left(\varepsilon\left(\boldsymbol{\eta}_{\ominus,4i+j,1}, \ldots, \boldsymbol{\eta}_{\ominus,4i+j,r_\ominus}\right)\right). \tag{2.19}$$

The r.h.s. of (2.19) amounts to rewriting the argument of the logarithm in (2.16) when the number of variables of each type is distributed according to r. Finally, let

$$\Xi_{i,j} = \Xi_{i,j}\left((-1)^{j+1}, \boldsymbol{r}_{4i+j}\right) \ . \tag{2.20}$$

Then the operator (2.17) maps $\rho_\bullet, \rho_\oplus, \rho_\ominus$ to the distributions $\hat{\rho}_\bullet, \hat{\rho}_\oplus, \hat{\rho}_\ominus$ of the random variables

$$\hat{\rho}_\bullet \sim -\sum_{i=1}^{\boldsymbol{d}_+^\star} \log \Xi_{i,1} + \sum_{i=1}^{\boldsymbol{d}_-^\star} \log \Xi_{i,2} \ , \qquad \hat{\rho}_\oplus \sim -\sum_{i=1}^{\boldsymbol{d}_+^{\star\prime}} \log \Xi_{i,3} \ , \qquad \hat{\rho}_\ominus \sim +\sum_{i=1}^{\boldsymbol{d}_-^{\star\prime}} \log \Xi_{i,4} \ . \tag{2.21}$$

2.5.3. *Coupling and contraction.* While Montanari and Shah [56] do not write their proof of the lower bound $d_{\mathrm{MS}}(k) \leq d_{\mathrm{uniq}}(k)$ in the language of distributional recurrences, translating their argument to the current formalism evinces two key differences by comparison to the approach that we are going to take. First, Montanari and Shah establish contraction with respect to messages *from clauses to variables*, instead of messages from variables to clauses as considered here.

While this change of perspective may seem innocuous at first, working with respect to variables provides us with greater control over how the change in log-likelihood ratios propagates. In particular, working with variable-to-clause messages and taking into account the four variable types $\bullet, \oplus, \ominus, \bigcirc$ allows us to optimise the metric with respect to which we establish contraction. Hence, for $t > 0$ we endow the space $\mathscr{P}(-\infty, \infty] \times \mathscr{P}(0, +\infty] \times \mathscr{P}(-\infty, 0]$ with the metric

$$\mathrm{dist}_t\left((\rho_\bullet, \rho_\oplus, \rho_\ominus), (\rho'_\bullet, \rho'_\oplus, \rho'_\ominus)\right) = \left(1 - e^{-t/2}\right) \cdot W_1\left(\rho_\bullet, \rho'_\bullet\right) + e^{-t/2} \cdot W_1\left(\rho_\oplus, \rho'_\oplus\right) + e^{-t/2} \cdot W_1\left(\rho_\ominus, \rho'_\ominus\right) \ . \tag{2.22}$$

The following proposition summarises the main step towards the proof of Theorem 1.2.

Proposition 2.15. *For every $d < d_{\mathrm{con}}(k)$, the operator $\mathrm{LL}^\star_{d,k}$ is a contraction with respect to the metric dist_d.*

The second key difference between [56] and the present approach will emerge in the proof of Proposition 2.15 itself. As we are about to see, leveraging the four variable types enables us to carry out a sharper bound on the derivative of our operator $\mathrm{LL}^\star_{d,k}$. This comes in the form of a subtle combinatorial coupling between variable types among clauses with opposite signs. To explain this, we recall that $\mathrm{LL}^\star_{d,k}$ describes how the laws of the log-likelihood ratios $\rho_\bullet, \rho_\oplus$, and $\rho_\ominus$, evolve given the corresponding laws of the variables in one generation below. Recall also that we are always considering the positive boundary condition, i.e., the one maximising the value of each log-likelihood ratio.

Let us write $\rho = (\rho_\bullet, \rho_\oplus, \rho_\ominus)$, $\rho' = (\rho'_\bullet, \rho'_\oplus, \rho'_\ominus)$, and $\hat{\rho}, \hat{\rho}'$ for their corresponding images under the operator $\mathrm{LL}^\star_{d,k}$. We wish to establish that $\mathrm{dist}_d(\hat{\rho}, \hat{\rho}') < c \cdot \mathrm{dist}_d(\rho, \rho')$, for some constant $c = c(d,k) < 1$. We call a clause a positive if it contains its parent variable as a direct literal; otherwise, we call a negative. The change between the output distributions $\hat{\rho}, \hat{\rho}'$ describing the log-likelihood law of, say, variable $\mathfrak{x}$, comes from two sources: the positive and the negative children of $\mathfrak{x}$. Observe that there is no obvious symmetry between the two, as we have imposed the positive boundary condition, and therefore, the influence of positive clauses is typically more pronounced. In turn, the change caused by each clause can be further attributed to that of the $k-1$ grandchildren variables it features. To be more precise, let us consider the contribution of a single positive clause a. Let us write $\boldsymbol{r} = (\boldsymbol{r}_\bullet, \boldsymbol{r}_\oplus, \boldsymbol{r}_\ominus, \boldsymbol{r}_\bigcirc)$ for the type-distribution of the children variables of a, where $\boldsymbol{r}$ follows the law described in (2.18). Consider also an arbitrary enumeration of the variables of each type $t \in \{\bullet, \oplus, \ominus\}$, and write $\mathscr{D}_i^t(z, \boldsymbol{r}; +1)$ for the magnitude of the partial derivative of the message clause a sends to $\mathfrak{x}$, with respect to the message clause a receives from its i-th variable of type t. Then, the expected contribution of clause a to the distance $\mathrm{dist}(\hat{\rho}, \hat{\rho}')$ is bounded in terms of $\mathscr{D}_i^t(z, \boldsymbol{r}; +)$'s as follows

$$\mathbb{E}\left[\sum_{i=1}^{\boldsymbol{r}_\bullet} \left| \int_{\boldsymbol{\eta}_{\bullet,i}}^{\boldsymbol{\eta}'_{\bullet,i}} \mathscr{D}_i^\bullet(w_i, \boldsymbol{r}; +1) \mathrm{d}w_i \right| + \sum_{j=1}^{\boldsymbol{r}_\oplus} \left| \int_{\boldsymbol{\eta}_{\oplus,j}}^{\boldsymbol{\eta}'_{\oplus,j}} \mathscr{D}_j^\oplus(y_j, \boldsymbol{r}; +1) \mathrm{d}y_j \right| + \sum_{\ell=1}^{\boldsymbol{r}_\ominus} \left| \int_{\boldsymbol{\eta}_{\ominus,\ell}}^{\boldsymbol{\eta}'_{\ominus,\ell}} \mathscr{D}_\ell^\ominus(z_\ell, \boldsymbol{r}; +1) \mathrm{d}z_\ell \right| \right] \ , \tag{2.23}$$

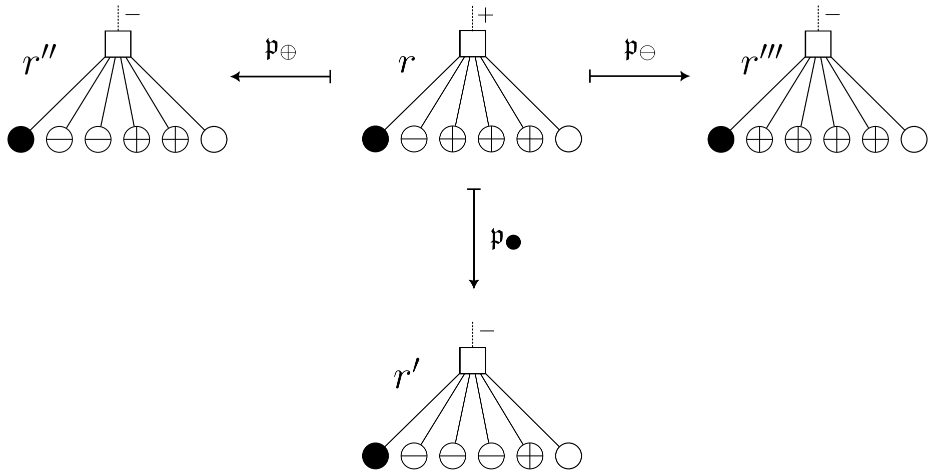


FIGURE 2. Example of a coupling between derivative terms in (2.24)–(2.25). For vector r and type $t \in \{\bullet, \oplus, \ominus\}$, we pair the term $\mathscr{D}^t(z, r; +1)$ in (2.24) with the term $\mathscr{D}^t(z, \mathfrak{p}_t(r); -1)$ in (2.25).

where $\boldsymbol{\eta}_{t,i}, \boldsymbol{\eta}'_{t,i}$ follow the law of ρ_t, ρ'_t, respectively. Expanding the expectation with respect to the type-distribution $\boldsymbol{r}$, and writing $P(r) = \mathbb{P}[\boldsymbol{r} = r]$, for the probability of a vector $r = (r_\bullet, r_\oplus, r_\ominus, r_\bigcirc)$, we rewrite (2.23) as

$$\sum_r P(r)\left(r_\bullet \cdot \mathbb{E}\left|\int_{\boldsymbol{\eta}_{\bullet,1}}^{\boldsymbol{\eta}'_{\bullet,1}} \mathscr{D}_1^{\bullet}(z, r; +1)\mathrm{d}z\right| + r_\oplus \cdot \mathbb{E}\left|\int_{\boldsymbol{\eta}_{\oplus,1}}^{\boldsymbol{\eta}'_{\oplus,1}} \mathscr{D}_1^{\oplus}(z, r; +1)\mathrm{d}z\right| + r_\ominus \cdot \mathbb{E}\left|\int_{\boldsymbol{\eta}_{\ominus,1}}^{\boldsymbol{\eta}'_{\ominus,1}} \mathscr{D}_1^{\ominus}(z, r; +1)\mathrm{d}z\right|\right). \tag{2.24}$$

The expected contribution of a negative clause is given by an expression similar to (2.24), albeit in terms of $\mathscr{D}_1^t(z, \boldsymbol{r}; -)$, i.e., the partial derivative of the message $a \to \mathfrak{x}$, with respect to the message from a variable of type t to clause a. Specifically, the expected contribution of a negative clause reads:

$$\sum_{r'} P(r')\left(r'_\bullet \cdot \mathbb{E}\left|\int_{\boldsymbol{\eta}_{\bullet,1}}^{\boldsymbol{\eta}'_{\bullet,1}} \mathscr{D}_1^{\bullet}(z, r'; -1)\mathrm{d}z\right| + r'_\oplus \cdot \mathbb{E}\left|\int_{\boldsymbol{\eta}_{\oplus,1}}^{\boldsymbol{\eta}'_{\oplus,1}} \mathscr{D}_1^{\oplus}(z, r'; -1)\mathrm{d}z\right| + r'_\ominus \cdot \mathbb{E}\left|\int_{\boldsymbol{\eta}_{\ominus,1}}^{\boldsymbol{\eta}'_{\ominus,1}} \mathscr{D}_1^{\ominus}(z, r'; -1)\mathrm{d}z\right|\right). \tag{2.25}$$

It is not hard to see that pure literals have a 'dampening' effect on each partial derivative $\mathscr{D}_1^t(z, r; \pm)$. Consider a clause a whose children variables are distributed among the different types according to r. Then each derivative in (2.24)–(2.25), can be bounded in terms of the number of pure literals featured in a, excluding the variable of type t with respect to which the derivative is taken. Notice that the operator $\mathrm{LL}^\star_{d,k}$ effectively incorporates the positive boundary condition by imposing the sign of each variable with respect to its parent clause, a, to be + if a is positive, and − if a is negative. With that in mind, we see that if a is a positive clause, the total number of pure literals it contains is just $r_\ominus + r_\bigcirc$. On the other hand, if a is negative, then the total number of pure literals it contains is $r_\oplus + r_\bigcirc$. Bounding *separately* each derivative $\mathscr{D}_i^t(z, r; \pm 1)$ in (2.24)–(2.25), and invoking the mean value theorem, yields an upper bound on for the contraction constant c.

However, we can do better by partitioning the derivatives in (2.24)–(2.25) into groups, and optimising them *jointly*. Indeed, a careful examination of the expression (2.19), reveals that, for example, any sum of the form $\mathscr{D}^\oplus(z, (*, *, r_\ominus, r_\bigcirc); +1) + \mathscr{D}^\oplus(z, (*, r_\ominus + 1, *, r_\bigcirc); -1)$ can be explicitly maximised, and the resulting maximum is smaller than the sum of maxima of the parts. At first sight, this seems to be of little use, if any, as in order to implement such a coupling between terms (2.24)–(2.25), we should also match their coefficients, that is, the quantity $P(r) \cdot r_\oplus$, must remain invariant under the coupling. Somewhat unexpectedly, it turns out that the coupling $r \mapsto r'$ with $r' = (r_\bullet, r_\ominus + 1, r_\oplus - 1, r_\bigcirc)$, enjoys both features. Similar couplings strategies (depicted in Figure 2) facilitate the maximisation of $\oplus, \ominus$-terms. The full proof of Proposition 2.15 can be found in Section 7.

We conclude the section explaining how Theorem 1.2 follows from the above.

Proof of Theorem 1.2. From the triangle inequality, and Lemma 2.12, it is immediate to obtain Theorem 1.2 from Proposition 2.13. □

2.6. **Discussion.** The location of the random 2-SAT satisfiability threshold was pinpointed already in the 1990s [21, 40] essentially because the threshold coincides with the giant component phase transition of a directed random graph whose edges correspond to the clauses. This argument also implies that both the pure literal algorithm and another efficient algorithm called unit clause propagation find satisfying assignments up to the satisfiability threshold w.h.p. By contrast, in the case of random k-SAT with $k \geq 3$ the satisfiability threshold is known only for k exceeding an undetermined (but large) constant k_0 [33]. The proof is based on a sophisticated, physics-inspired second moment argument that significantly extends ideas from earlier work [5, 7, 27]. Asymptotically in the limit of large k the satisfiabiliy threshold reads

$$d_{\mathrm{sat}}(k) = k\left[2^k \log 2 - \frac{1+\log 2}{2}\right] + \varepsilon_k, \qquad \text{where } \lim_{k\to\infty} \varepsilon_k = 0. \tag{2.26}$$

Even though [5, 7, 33, 27] rely on the second moment method, they do not yield asymptotically tight estimates of the number of satisfying assignments for any regime of d. This is because the second moment method is applied not to the number of satisfying assignments, but to another, exponentially smaller random variable. The assumption that k exceeds a large constant is used critically in [27, 33] to ensure certain concentration and expansion properties.

For $3 \leq k < k_0$ even the *existence* of a uniform satisfiability threshold remains an open problem, although a sharp threshold sequence that may vary with n is known to exist [36]. That said, an upper bound on the satisfiability threshold (sequence) that matches the so-called '1-step replica symmetry breaking' prediction from statistical physics can be verified using the interpolation method from mathematical physics [35, 49, 51, 60]. However, the currently known lower bounds for small k (say, $k = 3,4,5$) fall short of this upper bound [7, 42, 46]. For example, in the case $k = 3$ the best current lower bound is $d_{\mathrm{sat}}(3) \geq 10.56$, while $d_{\mathrm{sat}}(3) \approx 12.801$ according to physics predictions [49, 51].

Thus, the satisfiability of random formulas continues to pose a substantial challenge for 'small' $3 \leq k < k_0$. In light of this, a particularly satisfactory aspect of the present results is that they apply and are meaningful for all $k \geq 3$. In fact, comparing the asymptotic bounds (1.10) and (2.26), we see that the Gibbs uniqueness threshold $d_{\mathrm{uniq}}(k)$ is much smaller than $d_{\mathrm{sat}}(k)$ for large k . Thus, Theorems 1.1 and 1.2 cover larger shares of the satisfiable regime of d for smaller values of k; cf. Table 1.

The best current lower bounds on the satisfiability thresholds for $k \geq 4$ are non-constructive. With respect to the algorithmic problem of finding a satisfying assignment of a random k-CNF the best current results for 'small' k are based on simple combinatorial algorithms, analysed via the method of differential equations [37, 42, 46]. Asymptotically for large k the best known efficient algorithm [22] succeeds up to

$$d_{\mathrm{alg}}(k) = (1-\varepsilon_k)2^k \log k \qquad \text{where } \lim_{k\to\infty} \varepsilon_k = 0, \tag{2.27}$$

about a factor of $\log(k)/k$ below (2.26). There is evidence that certain types of algorithms do not succeed for much larger values of d, at least for enough large k [2, 15, 23, 44]. Apart from the task of finding a satisfying assignment, an important line of work deals with the problem of counting and sampling satisfying assignments of random k-CNFs for large k [19, 20, 43]. The best current result [20] covers the regime $d \leq 2^k/k^c$ for an undetermined (large enough) constant $c > 0$. Since for large k the bound $2^k/k^c$ significantly exceeds the pure literal threshold (1.10), it might be an interesting question whether ideas from [19, 20, 43] can be used to verify the replica symmetric solution (1.7) for d beyond the Gibbs uniqueness threshold for large k.

Most of the prior work on the rigorous verification of the replica symmetric solution focuses on a soft version of random k-SAT, the so-called random k-SAT model at *inverse termperature* $\beta > 0$ [57]. The partition function $Z_\beta(\mathbf{\Phi})$ of this model, its the key quantity of interest, is defined as follows. For a clause a of the random formula $\mathbf{\Phi}$ and a truth assignment σ write $\sigma \models a$ if σ satisfies clause a. Then

$$Z_\beta(\mathbf{\Phi}) = \sum_{\sigma\in\{\pm 1\}^n} \exp\left(-\beta \sum_{i=1}^{\boldsymbol{m}} \mathbb{1}\{\sigma \not\models a_i\}\right). \tag{2.28}$$

Thus, each assignment σ contributes a summand equal to $\exp(-\beta)$ raised to the power of the number of clauses that σ fails to satisfy. In effect,

$$Z(\mathbf{\Phi}) = \lim_{\beta\to\infty} Z_\beta(\mathbf{\Phi}). \tag{2.29}$$

A line of prior work [13, 59, 62] deals with the derivation of the 'thermodynamic limit'

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[\log Z_\beta(\mathbf{\Phi})\right] \tag{2.30}$$

for small d and/or small β. Specifically, these works verify that (2.30) is given by the replica symmetric solution at inverse temperature β from [54, 55] under the assumption

$$d(k-1)\min\left\{1,6\beta\exp(4\beta)\right\}<1. \tag{2.31}$$

We observe that for large β the bound (2.31) holds only up to the giant component threshold $d=1/(k-1)$, where the replica symmetric solution trivially follows from the fact that Belief Propagation is exact on acyclic graphical models [50, Theorem 4.1]. That said, a technique called the interpolation method shows that the replica symmetric solution yields an upper bound on (2.30) for all $d,\beta>0$ [35, 60]. In particular, we will combine the interpolation method with a concentration argument in order to prove Corollary 2.2.

According to physics predictions the 'replica symmetric solution' from [54, 55] yields the correct value of both $\lim_{n\to\infty}n^{-1}\log Z_\beta(\mathbf{\Phi})$ for all $\beta>0$ and of $\lim_{n\to\infty}n^{-1}\log Z(\mathbf{\Phi})$ for all d up to a threshold $d_{\mathrm{rsb}}(k)$ close to but strictly below the satisfiability threshold $d_{\mathrm{sat}}(k)$ for all $k\geq 3$ [47]. The threshold $d_{\mathrm{rsb}}(k)$ is known as the '1-step replica symmetry breaking phase transition' in physics jargon; its asymptotic value is predicted as

$$d_{\mathrm{rsb}}(k)=k\left[2^k\log 2-2\log 2\right]+\varepsilon_k, \qquad\qquad \text{where } \lim_{k\to\infty}\varepsilon_k=0. \tag{2.32}$$

Indeed, the interpolation method can be used to verify that the replica symmetric solution ceases to be correct for $d_{\mathrm{rsb}}(k)+\varepsilon_k<d<d_{\mathrm{sat}}(k)$ with $\varepsilon_k\to 0$. Conversely, the replica symmetric solution is known to be correct for all d and $\beta>0$ where a certain correlation decay condition is satisfied [26], provided that k is large enough. Physics methods predict that this condition holds for all $\beta>0$ and all $d<d_{\mathrm{rsb}}(k)$ [47].

The aforementioned work of Montanari and Shah [56] also deals with the soft variant of random k-SAT (2.28), but allows for an inverse temperature $\beta=\beta(n)$ that tends to infinity slowly as $n\to\infty$. Specifically, considering a small power $\beta=n^\delta$ enables Montanari and Shah to estimate the number of assignments that satisfy all but $o(n)$ clauses. The proof combines an interpolation on $0\leq\beta\leq n^\delta$ with a contraction argument that improves over the previous contraction estimates from [13, 59, 62]. Instead of the interpolation on β, in order to prove Theorem 1.1 we use the Aizenman-Sims-Starr scheme. Because we count actual satisfying assignments, this requires the careful combinatorial analysis of tail events, which is where the PULP algorithm and its analysis come in. Additionally, towards the proof of Theorem 1.2 we devise an improved version of the contraction argument from [56]. Following Montanari and Shah, we also take advantage of the impact of pure literals on the Belief Propagation operator. But we develop an improved coupling scheme that yields a better range of d for which contraction occurs. Additionally, once again because we deal with actual satisfying assignments, the proof of the Gibbs uniqueness property involves the analysis of the PULP algorithm on a Galton-Watson tree in order to cope with unlikely events.

By comparison to the 'soft' random k-SAT model (2.28), few prior contributions deal with the actual number $Z(\mathbf{\Phi})$ of satisfying assignments. A result of Abbe and Montanari [1] implies that a deterministic limit (in probability)

$$\lim_{n\to\infty}\frac{1}{n}\log Z(\mathbf{\Phi}) \tag{2.33}$$

exists for Lesbegue-almost all $0<d<d_{\mathrm{pure}}(k)$ for all $k\geq 2$. However, the proof, which is based on the interpolation method, does not reveal the value of (2.33). In fact, prior to the present work the limit (2.33) was known only in two cases. First, in the trivial regime $d<1/(k-1)$ below the giant component threshold. Second, in the case $k=2$ for $0<d<d_{\mathrm{sat}}(2)=2$ [3]. In both cases the limit (2.33) coincides with the replica symmetric solution from [54]. Beyond the convergence in probability, $\log Z(\mathbf{\Phi})$ is known to satisfy a central limit theorem in the case $k=2$ [17].

To compute the limit (2.33) in the case $k=2$ the contribution [3] employs the Aizenman-Sims-Starr scheme. The couplings that we use towards the proofs of Proposition 2.5–2.6 generalise the argument from [3] to $k\geq 3$. The main technical novelty lies in the way that moderately unlikely events are treated. Specifically, in the case $k=2$ the simple Unit Clause propagation algorithm, which essentially boils down to directed reachability, was sufficient to derive a tail bound similar to (and actually stronger than) (2.5). By contrast, since in the case $k\geq 3$ the clauses "branch out", the analysis of tail events and, accordingly, the derivation of (2.5) is far more delicate. The core of this derivation is the detailed analysis of the PULP algorithm right up to $d_{\mathrm{pure}}(k)$.

Finally, by contrast to random k-SAT the validity of the replica symmetric solution is known for the optimal parameter range for several other random constraint satisfaction problems that enjoy certain symmetry properties.

Examples include random graph colouring or random k-NAESAT [11, 25]. Due to the symmetry property [7] the replica symmetric solution simply coincides with the first moment of the number of solutions. In effect, in many symmetric problems it is even possible to precisely determine the limiting distribution of the number of solutions, which superconcentrates on the first moment [24]. By contrast, in random k-SAT the first moment overshoots the typical number of satisfying assignments by an exponential factor [5], which is why random k-SAT is so much more delicate than symmetric problems. That said, there is a regular variant of random k-SAT (where every variable appears an equal number of times positively and negatively) where symmetry and superconcentration are recovered [29].

2.7. **Organisation.** In the remaining sections we work our way through the proofs of Theorems 1.1 and 1.2. Specifically, in Section 3 we analyse the PULP algorithm introduced in Section 2.3, proving Lemmas 2.8-2.10, which facilitates many of the subsequent results.

In Section 4 we establish Proposition 2.1, verifying that the quantities appearing in Theorem 1.1 are well-defined. The proof of Corollary 2.2 follows in Section 5. Section 6 is devoted to the Aizenman-Sims-Starr scheme, and in particular the proof of Proposition 2.3. There we also complete the proof of Theorem 1.1.

Our final Section 7, deals with the remaining proofs toward establishing Theorem 1.2. We begin by proving Lemma 2.14, showing that the log-likelihood ratios of the random Galton-Watson formula close to the root are bounded w.h.p. This enables us to compare the output distribution of the non-random operator introduced in Section 2.5 with that of actual ratios on the random tree. We then proceed with the proof of Proposition 2.15, and conclude with the proof of (2.13), completing the proof of Theorem 1.2.

3. Analysis of PULP

This section is concerned with the analysis of PULP from Section 2.3. In particular, we prove Lemma 2.10. But let us get the proof of Lemma 2.8 out of the way first.

3.1. **Proof of Lemma 2.8.** Suppose that $\bar{\mathscr{L}} \supseteq \mathscr{L}$ satisfies **PULP1**–**PULP2**. Let $U = \{|l| : l \in \bar{\mathscr{L}}\}$ be the set of variables underlying the literals $\bar{\mathscr{L}}$. Moreover, let $\chi : U \to \{\pm 1\}$ be the truth assignments under which all literals of $l \in \bar{\mathscr{L}}$ evaluate to 'true'. Due to **PULP2**, the assignment χ is well defined. Moreover, since $\mathscr{L} \subseteq \bar{\mathscr{L}}$, under χ all literals $l \in \mathscr{L}$ evaluate to 'true'. Hence, for a satisfying assignment $\sigma \in S(\Phi)$ define an assignment σ' by letting

$$\sigma'(x) = \mathbb{1}\{x \in U\}\chi(x) + \mathbb{1}\{x \notin U\}\sigma(x) \ .$$

Because $\bar{\mathscr{L}}$ satisfies condition **PULP1**, we have $\sigma' \in S(\Phi, \mathscr{L})$. Finally, because for a satisfying assignment $\tau' \in S(\Phi, \mathscr{L})$ there are no more than $2^{|U|} = 2^{|\bar{\mathscr{L}}|}$ satisfying assignments $\tau \in S(\Phi)$ such that $\tau(x) = \tau'(x)$ for all $x \notin U$, we obtain the desired bound $Z(\Phi) \le 2^{|\bar{\mathscr{L}}|} Z(\Phi, \mathscr{L})$.

3.2. **Turning a tree to PULP.** While the ultimate goal of this section is to study the PULP algorithm on the random formula $\mathbf{\Phi}'$ to prove Lemma 2.10, a necessary preparation is to investigate the algorithm on the random Galton-Watson tree $\mathbb{T} = \mathbb{T}_{d,k}$. Of course, since $\mathbb{T}$ may be infinite we should formally confine ourselves to the finite trees $\mathbb{T}^{(\ell)}$ truncated at the 2ℓ-th level from the root $\mathfrak{r}$. Hence, recalling (2.6), we aim to estimate the height $\mathfrak{h}_{\mathfrak{r}}(s, \mathbb{T}^{(\ell)})$ for finite ℓ. That said, since these random variables are monotonically increasing in ℓ, it makes sense to define

$$\mathfrak{h}_{\mathfrak{r}}(s, \mathbb{T}) = \lim_{\ell \to \infty} \mathfrak{h}_{\mathfrak{r}}(s, \mathbb{T}^{(\ell)}) \in [0, \infty]. \tag{3.1}$$

We point out that for $d < d_{\mathrm{pure}}(k)$ the tails of $\mathfrak{h}_{\mathfrak{r}}(s, \mathbb{T})$ decay at a doubly exponential rate.

Lemma 3.1. *For any $d < d_{\mathrm{pure}}(k)$ there exist $c_1 = c_1(d,k), c_2 = c_2(d,k) > 0$ such that*

$$\mathbb{P}\left[\mathfrak{h}_{\mathfrak{r}}(\pm 1, \mathbb{T}) \ge h\right] \le c_1 \cdot \exp\left(-\exp\left(c_2 \cdot h\right)\right) \qquad \text{for every } h \ge 1 \ .$$

Proof. By symmetry it suffices to consider $\mathfrak{h}_{\mathfrak{r}}(1, \mathbb{T})$. Thus, let $p_{h,\ell} = \mathbb{P}\left[\mathfrak{h}_{\mathfrak{r}}(1, \mathbb{T}^{(\ell)}) \ge h\right]$. All variables at distance 2ℓ from $\mathfrak{r}$ are leaves and therefore pure in the tree $\mathbb{T}^{(\ell)}$. Consequently, pure literal elimination removes all clauses of $\mathbb{T}^{(\ell)}$ within at most ℓ rounds. Hence, $p_{h,\ell} = 0$ for $h > \ell$. Furthermore, we claim that

$$p_{h,\ell} = \varphi_{d,k}(p_{h-1,\ell-1}), \qquad \text{where} \quad \varphi_{d,k}(z) = 1 - \exp\left(-\frac{d}{2} z^{k-1}\right) \qquad (1 \le h \le \ell). \tag{3.2}$$

[7] A formal definition of 'symmetry' could be that the uniform distribution on spins is a fixed point of the Belief Propagation operator on the respective Galton-Watson tree.

Indeed, if $\mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T}^{(\ell)}) \ge h \ge 1$ then by (2.6) there exists a clause $a \in \partial_{\mathbb{T}} r$ with $\mathrm{sign}(\mathfrak{x},a) = -1$ such that $\mathfrak{h}_a(\mathbb{T}^{(\ell)}[\mathfrak{x} \mapsto 1]) \ge h-1$. In other words, pure literal elimination on the sub-tree $\mathbb{T}^{(\ell)}[a]$ of $\mathbb{T}^{(\ell)}$ rooted at clause a and with variable $\mathfrak{x}$ removed takes at least $h-1$ rounds to remove clause a. Consequently, pure literal elimination on $\mathbb{T}^{(\ell)}[a]$ takes at least $h-1$ rounds to remove one of the variables $x \in \partial_{\mathbb{T}} a \setminus \{\mathfrak{x}\}$. In other words, the sub-tree $\mathbb{T}^{(\ell)}[x]$ comprising x and its successors satisfies

$$\mathfrak{h}_x(\mathrm{sign}(x,a),\mathbb{T}^{(\ell)}[x]) \ge h-1 \qquad \text{for every } x \in \partial_{\mathbb{T}} a \setminus \{\mathfrak{x}\}. \tag{3.3}$$

But since $\mathbb{T}$ is a Galton-Watson tree, the sub-tree $\mathbb{T}^{(\ell)}[x]$ has the same distribution as the random tree $\mathbb{T}^{(\ell-1)}$. Hence, (3.3) implies that for every $a \in \partial_{\mathbb{T}}\mathfrak{x}$ with $\mathrm{sign}(x,a) = -1$,

$$\mathbb{P}\left[\mathfrak{h}_a(\mathbb{T}^{(\ell)}[\mathfrak{x} \mapsto 1]) \ge h-1 \mid \mathbb{T}^{(1)}\right] = p_{h-1,\ell-1}^{k-1}. \tag{3.4}$$

Finally, the construction of $\mathbb{T}$ ensures that the number of $a \in \partial_{\mathbb{T}}\mathfrak{x}$ with $\mathrm{sign}(x,a) = -1$ has distribution $\mathrm{Po}(d/2)$. Therefore, (3.4) shows that

$$p_{h,\ell} = \sum_{i=0}^{\infty} \mathbb{P}\left[\mathrm{Po}(d/2) = i\right]\left(1-\left(1-p_{h-1,\ell-1}^{k-1}\right)^i\right) = 1-\exp\left(-\frac{d}{2}p_{h-1,\ell-1}^{k-1}\right),$$

which completes the proof of (3.2).

Since the sequences $(p_{h,\ell})_\ell$ are non-decreasing, the limits $p_h = \lim_{\ell\to\infty} p_{h,\ell}$ exist. Moreover, (3.2) shows that

$$p_h = \varphi_{d,k}(p_{h-1}) \qquad (h \ge 1). \tag{3.5}$$

Hence, recalling the definition (1.8) of $d_{\mathrm{pure}}(k)$, we find

$$\left(\frac{p_{h+1}}{p_h}\right)^{k-1} = \left(\frac{\varphi_{d,k}(p_h)}{p_h}\right)^{k-1} = d\cdot\frac{\left(1-\exp\left(-dp_h^{k-1}/2\right)\right)^{k-1}}{dp_h^{k-1}} \le \frac{d}{d_{\mathrm{pure}}} < 1\ .$$

Consequently,

$$\lim_{h\to\infty} p_h = 0. \tag{3.6}$$

To complete the proof we expand $\varphi_{d,k}(z)$ around $z=0$:

$$\varphi_{d,k}(z) = \frac{d}{2}z^{k-1} + O(z^{2k-2}) \qquad \text{as } z \to 0. \tag{3.7}$$

Thus, the function $\varphi_{d,k}(z)$ is well approximated by a $(k-1)$-th power. Since $k \ge 3$, combining (3.5)–(3.7), we conclude that for sufficiently large h we have $p_h \le (d/2+1)p_{h-1}^{k-1}$. Consequently, $p_h \le c_1 \cdot \exp\left(-\exp(c_2 \cdot h)\right)$ for suitable $c_1 = c_1(d,k)$ and $c_2 = c_2(d,k)$. $\square$

We remind ourselves that $\overline{\{\pm 1\cdot\mathfrak{x}\}}_{\mathbb{T}^{(\ell)}}$ signifies the output of PULP run on the formula $\mathbb{T}^{(\ell)}$ with initial literal set $\{\pm 1\cdot\mathfrak{x}\}$. We extend the definition of the closure to the (possibly infinite) tree $\mathbb{T}$ by letting

$$\overline{\{\pm 1\cdot\mathfrak{x}\}}_{\mathbb{T}} = \bigcap_{\ell_0 \ge 1}\bigcup_{\ell \ge \ell_0} \overline{\{\pm 1\cdot\mathfrak{x}\}}_{\mathbb{T}^{(\ell)}}.$$

This definition ensures that if the height $\mathfrak{h}_{\mathfrak{x}}(\pm 1,\mathbb{T})$ from (3.1) is finite, then

$$\overline{\{\pm 1\cdot\mathfrak{x}\}}_{\mathbb{T}} = \overline{\{\pm 1\cdot\mathfrak{x}\}}_{\mathbb{T}^{(\ell)}} \qquad \text{for all } \ell \ge \mathfrak{h}_{\mathfrak{x}}(\pm 1,\mathbb{T}).$$

In order to estimate the size of this set, we combine Lemma 3.1 with a crude bound on the total number of variable nodes of the Galton-Watson tree $\mathbb{T}^{(\ell)}$. Recall that $V(\mathbb{T}^{(\ell)})$ signifies the set of variable nodes of $\mathbb{T}^{(\ell)}$.

Lemma 3.2. *Let $d > 0$. For any $\ell \ge 1$ and any $t > 100(1+d(k-1))^2$ we have*

$$\mathbb{P}\left[|V(\mathbb{T}^{(\ell)})| > t^\ell\right] \le \ell\exp(-t^{\ell/2}/4).$$

Proof. Let $\boldsymbol{N}_\ell = |V(\mathbb{T}^{(\ell)})|$ for brevity, set $g = 10(1+d(k-1))$ and notice that $t > g^2$. The construction of the Galton-Watson tree $\mathbb{T}$ ensures that $\boldsymbol{N}_0 = 1$ and that for $\ell \ge 1$ given $\boldsymbol{N}_{\ell-1}$ we have $\boldsymbol{N}_\ell \sim (k-1)\cdot\mathrm{Po}(d\boldsymbol{N}_{\ell-1})$. Therefore, Bennett's inequality shows that

$$\mathbb{P}\left[\boldsymbol{N}_h > g^{h-\ell}t^\ell \mid \boldsymbol{N}_{h-1} \le g^{h-1-\ell}t^\ell\right] \le \exp\left(-\frac{t^\ell}{4g^{\ell-h}}\right) \le \exp(-t^{\ell/2}/4) \qquad (1 \le h \le \ell). \tag{3.8}$$

Furthermore, if $\boldsymbol{N}_\ell > t^\ell$ then there exists $1 \le h \le \ell$ such that $\boldsymbol{N}_h > g^{h-\ell} t^\ell$ while $\boldsymbol{N}_{h-1} \le g^{h-1-\ell} t^\ell$. Hence, combining (3.8) with the union bound completes the proof. □

Corollary 3.3. *For any $d < d_{\mathrm{pure}}(k)$ there exists $c_3 = c_3(d,k) > 0$ such that*

$$\mathbb{P}[|\overline{\{\pm 1 \cdot \mathfrak{x}\}}_{\mathbb{T}}| > t] \le c_3 \exp(-t^{1/c_3}) \qquad \textit{for all } t > 0.$$

Proof. By symmetry it suffices to consider $\boldsymbol{N} = |\overline{\{\mathfrak{x}\}}_{\mathbb{T}}|$. Since Lemma 3.1 shows that $\mathbb{P}[\mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T}) < \infty] = 1$, we may assume from now on that indeed $\mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T}) < \infty$. Moreover, picking $c_3 = c_3(d,k) > 0$ large enough, we may assume that $t > t_0$ for a large $t_0 = t_0(d,k)$. Let $\boldsymbol{N}_\ell = |V(\mathbb{T}^{(\ell)})|$, $p_h = \mathbb{P}\left[\mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T}) = h\right]$ and $g = 10(1+d(k-1))$. It is an immediate consequence of the way that PULP proceeds that for all $l \in \overline{\{\mathfrak{x}\}}_{\mathbb{T}}$ we have $|l| \in V(\mathbb{T}^{(\mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T}))})$. Hence, $\boldsymbol{N} \le \boldsymbol{N}_{\mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T})}$. Therefore, by the law of total probability,

$$\mathbb{P}\left[\boldsymbol{N} > t\right] \le \sum_{h \ge 0} S_h, \qquad \text{where } S_h = \mathbb{P}\left[\mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T}) = h\right]\mathbb{P}\left[\boldsymbol{N}_h > t \mid \mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T}) = h\right]. \tag{3.9}$$

Depending on the value of t in relation to h, we use either Lemma 3.1 or Lemma 3.2 to bound S_h.

Case 1: $t_0 < t \le g^{2h}$**:** Lemma 3.1 shows that for certain $c_1, c_2 > 0$ we have

$$S_h \le \mathbb{P}\left[\mathfrak{h}_{\mathfrak{x}}(1,\mathbb{T}) = h\right] \le c_1 \exp(-\exp(c_2 h)) \le c_1 2^{-h} \exp(-t^{1/c_3}), \tag{3.10}$$

provided c_3 is chosen large enough.

Case 2: $t > g^{2h}$**:** we apply Lemma 3.2 to obtain

$$S_h \le \mathbb{P}\left[\boldsymbol{N}_h > t\right] \le h \exp(-\sqrt{t}/4) \le h 2^{-h} \exp(-t^{1/3}), \tag{3.11}$$

provided that $t > t_0$ is sufficiently large.

Combining the bounds (3.9)–(3.11) completes the proof. □

Corollary 3.4. *For any $d < d_{\mathrm{pure}}(k)$ we have* $\mathbb{E}[|\overline{\{\pm 1 \cdot \mathfrak{x}\}}_{\mathbb{T}}|^2] < \infty$.

Proof. This is an immediate consequence of Corollary 3.3. □

3.3. **Proof of Lemma 2.10.** Because the distribution of $\boldsymbol{\Phi}'$ is invariant under variable permutations and inversions, we may assume the initial set $\mathscr{L}$ of literals passed to PULP is just $\mathscr{L} = \{x_1, \dots, x_L\}$ for an integer $L = \tilde{O}(1)$. For an integer $\ell \ge 1$ let $\boldsymbol{\phi}'_{\ell,L}$ be the sub-formula of $\boldsymbol{\Phi}'$ comprising all clauses and variables at distance at most 2ℓ from $\mathscr{L}$. We recall that this formula has a bipartite graph representation $G(\boldsymbol{\phi}'_{\ell,L})$ with variable nodes $V(\boldsymbol{\phi}'_{\ell,L})$, clause nodes $F(\boldsymbol{\phi}'_{\ell,L})$ and edges $E(\boldsymbol{\phi}'_{\ell,L})$. The *excess* of $\boldsymbol{\phi}'_{\ell,L}$ is defined as

$$\boldsymbol{X}_{\ell,L} = |E(\boldsymbol{\phi}'_{\ell,L})| - |V(\boldsymbol{\phi}'_{\ell,L})| - |F(\boldsymbol{\phi}'_{\ell,L})|.$$

Thus, $\boldsymbol{X}_{\ell,L} = -L$ iff $G(\boldsymbol{\phi}'_{\ell,L})$ consists of L acyclic components.

Lemma 3.5. *Let $d > 0$, $c > 0$ and assume that $L \le \log^c n$ and $\ell \le c \log\log n$. Then*

$$\mathbb{P}\left[\boldsymbol{X}_{\ell,L} > -L\right] = \tilde{O}(n^{-1}), \qquad \mathbb{P}\left[\boldsymbol{X}_{\ell,L} > 1-L\right] = \tilde{O}(n^{-2}). \tag{3.12}$$

Furthermore, there exists $c_4 = c_4(c,d,k) > 0$ such that

$$\mathbb{P}\left[|V(\boldsymbol{\phi}'_{\ell,L})| + |F(\boldsymbol{\phi}'_{\ell,L})| > \log^{c_4} n\right] = O(n^{-2}). \tag{3.13}$$

Proof. We study breadth first search ('BFS') on the graph $G(\boldsymbol{\Phi}')$ from the start vertices $\mathscr{L}$ by means of a routine deferred decisions argument. Throughout the execution of BFS each variable node is in one of three possible states: unexplored, active, or finished.

Towards the proof of (3.13) we study a 'parallel' version of BFS. More precisely, let $\mathscr{A}_0 = \mathscr{L}$ be the set of initially active variables, let $\mathscr{U}_0 = \{x_1, \dots, x_n\} \setminus \mathscr{L}$ comprise the initially unexplored variables and let $\mathscr{F}_0 = \emptyset$. Further, for $t \ge 0$ define $\mathscr{A}_{t+1}, \mathscr{U}_{t+1}, \mathscr{F}_{t+1}$ as follows. If $\mathscr{A}_t = \emptyset$ then the process has stopped and we let $\mathscr{A}_{t+1} = \mathscr{A}_t = \emptyset, \mathscr{U}_{t+1} = \mathscr{U}_t, \mathscr{F}_{t+1} = \mathscr{F}_t$. Otherwise let $\mathscr{A}_{t+1}$ be the set of all variable nodes $y \in \mathscr{U}_t$ such that there exist an active variable node $x \in \mathscr{A}_t$ and a clause a that contains x and y; in symbols, $x, y \in \partial_{\boldsymbol{\Phi}'} a$. Further, let $\mathscr{F}_{t+1} = \mathscr{F}_t \cup \mathscr{A}_t$ and $\mathscr{U}_{t+1} = \mathscr{U}_t \setminus \mathscr{A}_{t+1}$. The BFS exploration occurs 'in parallel' in the sense that all active vertices activate their previously unexplored second neighbours simultaneously.

Let $\mathfrak{F}_t$ be the σ-algebra generated by the first t rounds of parallel exploring. Then the distribution of $|\mathscr{A}_{t+1}|$ given $\mathfrak{F}_t$ is stochastically dominated by a random variable with distribution $(k-1)\mathrm{Po}(d|\mathscr{A}_t|)$. This is because by the construction of the formula $\mathbf{\Phi}'$ the total number of clauses containing a given variable node has distribution $\mathrm{Po}(d(1-(k-1)/n))$. Hence, for any $u>0$ we have

$$\mathbb{P}\left[|\mathscr{A}_{t+1}|>u \mid \mathfrak{F}_t\right] \le \mathbb{P}\left[(k-1)\mathrm{Po}(d|\mathscr{A}_t|)>u\right]. \tag{3.14}$$

To complete the proof of (3.13) we mimic the argument from the proof of Lemma 3.2. Thus, let $u=\log^{c_4-3} n$ for a large enough $c_4=c_4(c,d,k)$ and set $g=10(1+d(k-1))$. Since $\ell \le c\log\log n$, the bound (3.14) and Bennett's inequality show that

$$\mathbb{P}\left[|\mathscr{A}_{t+1}|>g^{t+1-\ell}u \mid |\mathscr{A}_t|\le g^{t-\ell}u\right] \le \exp\left(-\frac{u}{4g^{\ell-t+1}}\right) \le \exp(-\sqrt{u}/4)=O(n^{-3}) \qquad (0\le t<\ell).$$

Hence, taking a union bound on $0\le t<\ell$ and observing that $|V(\boldsymbol{\phi}'_{\ell,L})|\subseteq \mathscr{A}_0\cup\cdots\cup\mathscr{A}_\ell$, we obtain

$$\mathbb{P}\left[|V(\boldsymbol{\phi}'_{\ell,L})|>u\log n\right]=\tilde{O}(n^{-3}). \tag{3.15}$$

Finally, another application of Bennett's inequality demonstrates that with probability $1-O(n^{-2})$ no variable of $\mathbf{\Phi}'$ appears in more than $\log n$ clauses. Thus, $|F(\boldsymbol{\phi}'_{\ell,L})|\le |V(\boldsymbol{\phi}'_{\ell,L})|\log n$. Hence, (3.15) implies (3.13).

We are left to establish (3.12). The way we set up the BFS process implies that there are only two ways in which excess edges can come about. First, there may be clauses a with $\partial_{\mathbf{\Phi}'}a\subseteq \mathscr{A}_t\cup\mathscr{A}_{t+1}$ such that $|\partial_{\mathbf{\Phi}'}a\cap\mathscr{A}_t|\ge 2$. Given that $|\mathscr{A}_t\cup\mathscr{A}_{t+1}|\le \log^{c_4} n$, the number of such a with $|\partial_{\mathbf{\Phi}'}a\cap\mathscr{A}_t|=2$ has distribution $\mathrm{Po}(\tilde{O}(1/n))$, and the number of a with $|\partial_{\mathbf{\Phi}'}a\cap\mathscr{A}_t|>2$ has distribution $\mathrm{Po}(\tilde{O}(1/n^2))$. The second possibility is that for a variable $x\in\mathscr{A}_{t+1}$ there exist clauses $a,b\in\partial_{\mathbf{\Phi}'}x$ with $\partial_{\mathbf{\Phi}'}a,\partial_{\mathbf{\Phi}'}b\subseteq\mathscr{A}_t\cup\mathscr{A}_{t+1}$. Once again the number of such clauses has distribution $\mathrm{Po}(\tilde{O}(1/n))$ given $|\mathscr{A}_t\cup\mathscr{A}_{t+1}|\le\log^{c_4} n$. Furthermore, excess inducing clauses occur independently at different rounds t of the BFS process. Thus, (3.12) follows from (3.13). □

We proceed to derive bounds on $|\bar{\mathscr{L}}|=|\bar{\mathscr{L}}_{\mathbf{\Phi}'}|$ depending on the value of the excess. To deal with the case of excess $-L$, let $\Lambda=\Theta(\log\log n)$ and let $(\mathbb{T}[i])_{i\ge1}$ be a sequence of independent copies of the random tree $\mathbb{T}$. In the case that the excess $\boldsymbol{X}_{\Lambda,L}$ equals $-L$, the bound on $|\bar{\mathscr{L}}|$ follows from the fact that the Galton-Watson tree $\mathbb{T}$ captures the local structure of the graph $G(\mathbf{\Phi}')$ in combination with the bound from Corollary 3.3. More precisely, the following is true.

Lemma 3.6. *For any $0<d<d_{\mathrm{uniq}}(k)$ and $c>0$ there exists $\zeta=\zeta(c,d,k)>0$ such that with $\Lambda=\lceil\zeta\log\log n\rceil$ uniformly for all $1\le L\le\log^c n$ and all $u>0$ we have*

$$\mathbb{P}\left[\mathbb{1}\{\boldsymbol{X}_{\Lambda,L}=-L\}|\bar{\mathscr{L}}|>u\right]\le\mathbb{P}\left[\sum_{i=1}^{L}|\overline{\{\mathfrak{r}\}}_{\mathbb{T}[i]}|>u\right]+O(n^{-2}).$$

Proof. We begin by coupling the random formula $\boldsymbol{\phi}'_{\ell,1}$ with the Galton-Watson tree $\mathbb{T}^{(\ell)}[1]$ for $0\le\ell\le\Lambda$. The coupling operates in accordance with the iterations of the BFS process from the proof of Lemma 3.5. Under the coupling some of the variable and clause nodes of $\boldsymbol{\phi}'_{\ell,1}$ and of the tree $\mathbb{T}^{(\ell)}[1]$ are identical, but both $\mathbb{T}^{(\ell)}[1]$ and $\boldsymbol{\phi}'_{\ell,1}$ may contain additional clauses or variables. These additional clauses/variables result from excess edges of $G(\boldsymbol{\phi}'_{\ell,1})$, i.e., edges that close cycles or merge different components in the course of the BFS process.

For $\ell=0$ we just identify the start variable x_1 with the root $\mathfrak{r}$ of the Galton-Watson tree $\mathbb{T}[1]$. Going from ℓ to $\ell+1$, we remember the sets $\mathscr{A}_\ell,\mathscr{A}_{\ell+1}$ from the proof of Lemma 3.5. For each variable $x\in\mathscr{A}_\ell$ let $\mathscr{C}_x$ be the set of clauses $a\in\partial_{\mathbf{\Phi}'}x$ such that $|\partial_{\mathbf{\Phi}'}a\cap\mathscr{A}_{\ell+1}|=k-1$ and also such that none of the variables $y\in\mathscr{A}_{\ell+1}\cap\partial_{\mathbf{\Phi}'}a$ appear in another clause $b\ne a$ with $\partial_{\mathbf{\Phi}'}b\subseteq\mathscr{A}_\ell\cup\mathscr{A}_{\ell+1}$. In other words, $\mathscr{C}_x$ contains all clauses $a\in\partial_{\mathbf{\Phi}'}x$ that do not induce excess edges. Let $\boldsymbol{d}_x=|\mathscr{C}_x|$ be the number of such clauses. As we pointed out in the proof of Lemma 3.5, $\boldsymbol{d}_x$ is stochastically dominated by a $\mathrm{Po}(d)$ variable. Hence, there is a random variable $\boldsymbol{d}'_x$ such that $\boldsymbol{d}_x+\boldsymbol{d}'_x\sim\mathrm{Po}(d)$.

For any variable $x\in\mathscr{A}_\ell$ that is also a variable node of $\mathbb{T}^{(\ell)}[1]$ we add all clauses $a\in\mathscr{C}_x$ and the $k-1$ variables $y\in\partial_{\mathbf{\Phi}'}a\cap\mathscr{A}_{\ell+1}$ to $\mathbb{T}^{(\ell+1)}[1]$. Additionally, $\mathbb{T}^{(\ell+1)}[1]$ contains $\boldsymbol{d}'_x$ independent random clauses that contain x and $k-1$ new variable nodes without a counterpart in $\boldsymbol{\phi}'_{\ell+1,1}$. Finally, to complete $\mathbb{T}^{(\ell+1)}[1]$ every variable y of $\mathbb{T}^{(\ell)}[1]$ at distance precisely 2ℓ from r such that $y\notin V(\boldsymbol{\phi}'_{\ell,1})$ independently begets $\mathrm{Po}(d)$ offspring clause nodes, each containing $k-1$ new variable nodes that do not belong to $V(\boldsymbol{\phi}'_{\ell+1,1})$.

The coupling ensures that $\boldsymbol{\phi}'_{\Lambda,1}$ is a sub-formula of $\mathbb{T}^{(\Lambda)}[1]$ unless $\boldsymbol{X}_{\Lambda,1}>-L$. The extension of this coupling to $\mathscr{L}=\{x_1,\ldots,x_L\}$ is straightforward. We simply perform BFS exploration from the start variables $x_1,\ldots,x_L$ one after the other. Given that $\boldsymbol{X}_{\Lambda,L}=-L$, we thus couple the sub-formula of $\mathbf{\Phi}'$ explored from each x_i with $\mathbb{T}^{(\Lambda)}[i]$

for $1 \le i \le L$ such that $\boldsymbol{\phi}'_{\Lambda,L}$ is contained in the union of $\mathbb{T}^{(\Lambda)}[1], \ldots, \mathbb{T}^{(\Lambda)}[L]$. Finally, we obtain independent copies $\mathbb{T}[1], \ldots, \mathbb{T}[L]$ of the (possibly infinite) tree $\mathbb{T}$ by continuing the Galton-Watson processes $\mathbb{T}^{(\Lambda)}[i]$ independently for depths $\ell > \Lambda$.

The remaining task is to compare $|\bar{\mathscr{L}}|$ with $\sum_{i=1}^{L} |\overline{\{\mathfrak{r}\}}_{\mathbb{T}[i]}|$. If $\boldsymbol{X}_{\Lambda,L} = -L$ and if $\mathfrak{h}_{\mathfrak{r}}(1, \mathbb{T}[i]) < \Lambda$ for all $1 \le i \le L$, then the coupling ensures that all clauses and variables of $\boldsymbol{\phi}'_{\Lambda,L}$ are contained in the disjoint union of the trees $\mathbb{T}[1], \ldots, \mathbb{T}[L]$, and thus $|\bar{\mathscr{L}}| \le \sum_{i=1}^{L} |\overline{\{\mathfrak{r}\}}_{\mathbb{T}[i]}|$. Therefore, for any $u > 0$ we have

$$\mathbb{P}\left[\mathbb{1}\left\{\boldsymbol{X}_{\Lambda,L} = -L, \max_{1\le i\le L} \mathfrak{h}_{\mathfrak{r}}(1, \mathbb{T}[i]) < \Lambda\right\}|\bar{\mathscr{L}}| > u\right] \le \mathbb{P}\left[\sum_{i=1}^{L} |\overline{\{\mathfrak{r}\}}_{\mathbb{T}[i]}| > u\right]. \tag{3.16}$$

Furthermore, since $\Lambda \ge \zeta \log\log n$ for a large $\zeta > 0$, Lemma 3.1 ensures that

$$\mathbb{P}\left[\max_{1\le i\le L} \mathfrak{h}_{\mathfrak{r}}(1, \mathbb{T}[i]) \ge \Lambda\right] \le O(n^{-2}). \tag{3.17}$$

Combining (3.16) and (3.17) completes the proof. $\square$

For later reference we make a note of the following immediate consequence of the coupling from the proof of Lemma 3.6. For two rooted Boolean formulas ϕ, ϕ' we write $\phi \cong \phi'$ if there is an isomorphism of ϕ and ϕ' that preserves the root variable. We consider the random formula $\boldsymbol{\phi}'_{\ell,1}$ rooted at x_1.

Corollary 3.7. *For every $\ell \ge 0$ and any fixed tree T we have* $\left|\mathbb{P}\left[\mathbb{T}^{(\ell)} \cong T\right] - \mathbb{P}\left[\boldsymbol{\phi}'_{\ell,1} \cong T\right]\right| = o(1)$.

From here on, we set $\Lambda = \lceil c_5 \log\log n\rceil$ for a large enough $c_5 = c_5(d,k) > 0$. We obtain the following bound on the second moment of $|\bar{\mathscr{L}}|$ on the event that the excess equals $-L$.

Corollary 3.8. *For any $0 < d < d_{\mathrm{uniq}}(k)$ and any $1 \le L \le \log^2 n$ we have* $\mathbb{E}[\mathbb{1}\{\boldsymbol{X}_{\Lambda,L} = -L\} \cdot |\bar{\mathscr{L}}|^2] = O(1)$.

Proof. Since $|\bar{\mathscr{L}}| \le 2n$ deterministically, this is an immediate consequence of Corollary 3.3 and Lemma 3.6. $\square$

As a next step we deal with the case that the excess equals $1 - L$. More precisely, with $c_6 = c_6(d,k) \gg c_5$ a large enough constant let $\Lambda^+ = \lceil c_6 \log\log n\rceil$. We are going to bound $|\bar{\mathscr{L}}|$ on the event that $\boldsymbol{X}_{\Lambda,L} = \boldsymbol{X}_{\Lambda^+,L} = 1-L$. The proof combines the bound on the probability of this event from Lemma 3.5 with a crude bound on $|\bar{\mathscr{L}}|$. To elaborate, since Lemma 3.5 shows that the event $\boldsymbol{X}_{\Lambda,L} = \boldsymbol{X}_{\Lambda^+,L} = 1-L$ has probability $\tilde{O}(n^{-1})$, we can essentially get away with simply bounding $|\bar{\mathscr{L}}|$ by the total number of variables within a $2\Lambda^+$ radius around the start variables $\mathscr{L}$. Indeed, as Lemma 3.5 shows, this number of variables is very likely polylogarithmic in n. Working out the details, we obtain the following.

Lemma 3.9. *Let $0 < d < d_{\mathrm{uniq}}(k)$ and let $1 \le L \le \log^2 n$. Then* $\mathbb{E}\left[\mathbb{1}\{\boldsymbol{X}_{\Lambda,L} = \boldsymbol{X}_{\Lambda^+,L} = 1-L\}|\bar{\mathscr{L}}|^{3/2}\right] = o(1)$.

Proof. Let $\mathscr{V}^+ = V(\boldsymbol{\phi}'_{\Lambda,L}) \setminus V(\boldsymbol{\phi}'_{\Lambda-1,L})$ and obtain $\boldsymbol{\psi}^-$ from $\boldsymbol{\Phi}'$ by deleting all variables from $V(\boldsymbol{\phi}'_{\Lambda-1,L})$ and all clauses from $F(\boldsymbol{\phi}'_{\Lambda,L})$. Further, let $\Lambda^- = \Lambda^+ - \Lambda$ and let $\boldsymbol{\psi}^+$ be the sub-formula of $\boldsymbol{\psi}^-$ comprising all clauses and variables of $\boldsymbol{\psi}^-$ with distance at most $2\Lambda^-$ from $\mathscr{V}^+$ (see Figure 3 below). If $\boldsymbol{X}_{\Lambda,L} = \boldsymbol{X}_{\Lambda^+,L} = 1-L$ then

$$|V(\boldsymbol{\phi}'_{\Lambda,L})| + |F(\boldsymbol{\phi}'_{\Lambda,L})| - |E(\boldsymbol{\phi}'_{\Lambda,L})| = L - 1, \tag{3.18}$$

$$|V(\boldsymbol{\psi}^+)| + |F(\boldsymbol{\psi}^+)| - |E(\boldsymbol{\psi}^+)| = |\mathscr{V}^+|. \tag{3.19}$$

Moreover, Lemma 3.5 shows that for suitable $c_5' = c_5'(d,k,c_5), c_6' = c_6'(d,k,c_6) > 0$ we have

$$\mathbb{P}\left[|V(\boldsymbol{\phi}'_{\Lambda,L})| + |F(\boldsymbol{\phi}'_{\Lambda,L})| > \log^{c_5'} n\right] = O(n^{-2}), \tag{3.20}$$

$$\mathbb{P}\left[|V(\boldsymbol{\phi}'_{\Lambda^+,L})| + |F(\boldsymbol{\phi}'_{\Lambda^+,L})| > \log^{c_6'} n\right] = O(n^{-2}). \tag{3.21}$$

Let $\hat{\mathscr{L}} \subseteq \bar{\mathscr{L}}$ be the set of literals l that were added to the output set $\bar{\mathscr{L}}$ by Step 7 of PULP by way of clauses $a \in F(\boldsymbol{\phi}'_{\Lambda,L})$, i.e., at distance less than 2Λ from the initial set $\mathscr{L}$. Let $\mathscr{V}^- = \{|l| : l \in \hat{\mathscr{L}}\} \cap \mathscr{V}^+$ be the set of all variables at distance 2Λ from $\mathscr{L}$ in $\boldsymbol{\Phi}'$ that underlie a literal from $\hat{\mathscr{L}}$. If $\boldsymbol{X}_{\Lambda,L} = 1-L$, the variables and clauses at distance at most 2Λ from $\mathscr{L}$ do not cause PULP to run into a contradiction, because each clause contains $k \ge 3$ literals. Therefore, there does not exist a variable x such that both x and $\neg x$ belong to $\hat{\mathscr{L}}$. Hence, because the signs of

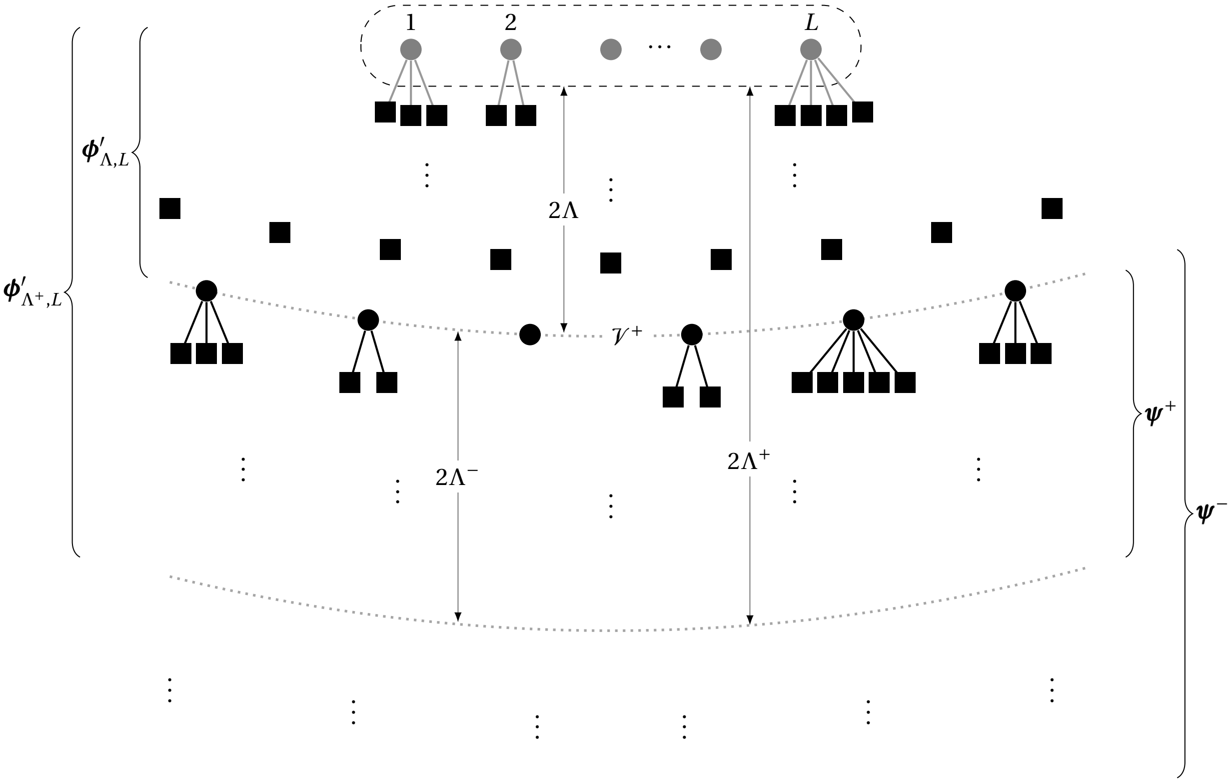


FIGURE 3. A sketch depicting the subformulas $\boldsymbol{\psi}^+, \boldsymbol{\psi}^-, \boldsymbol{\phi}'_{\Lambda,L}$, and $\boldsymbol{\phi}'_{\Lambda^+,L}$ of $\boldsymbol{\Phi}'$ constructed above.

$\boldsymbol{\Phi}'$ are uniformly random and PULP proceeds in a BFS order, we may assume without loss of generality that $\hat{\mathscr{L}}$ contains positive literals only. Thus,

$$\mathcal{V}^- = \hat{\mathscr{L}} \cap \mathcal{V}^+ \subseteq \mathcal{V}^+. \tag{3.22}$$

We now apply Lemma 3.6 to the random formula $\boldsymbol{\psi}^-$. Specifically, let $\bar{\mathscr{L}}^+$ be the output of PULP on the formula $\boldsymbol{\psi}^-$ with the start set $\mathscr{L}^+$ comprised by the positive literals of $\mathcal{V}^+$. Further, let $\mathfrak{E}$ be the event that $|V(\boldsymbol{\phi}'_{\Lambda,L})| + |F(\boldsymbol{\phi}'_{\Lambda,L})| \le \log^{c_5'} n$. Let

$$\boldsymbol{X}^+ = |E(\boldsymbol{\psi}^+)| - |V(\boldsymbol{\psi}^+)| - |F(\boldsymbol{\psi}^+)|$$

be the excess of $\boldsymbol{\psi}^+$. Since $0 < d < d_{\mathrm{uniq}}(k)$, given $\mathfrak{E}$ the formula $\boldsymbol{\psi}^-$ has the same distribution as a random k-CNF with $n^- = n - O(\log^c n)$ variables and $\boldsymbol{m}^- \sim \mathrm{Po}(d^- n^-/k)$ random clauses, with $0 < d^- = d + o(1) < d_{\mathrm{uniq}}(k)$. Hence, assuming that $c_6 = c_6(d,k) > c_5'$ is sufficiently large, Lemma 3.6 shows that

$$\mathbb{P}\left[\mathbb{1}\mathfrak{E} \cap \{\boldsymbol{X}^+ = -|\mathcal{V}^+|\} \cdot |\bar{\mathscr{L}}^+| > u\right] \le \mathbb{P}\left[\sum_{1 \le i \le \log^{c_5'} n} |\overline{\{\mathfrak{x}\}}_{\mathbb{T}[i]}| > u\right] + O(n^{-2}) \qquad (u > 0). \tag{3.23}$$

Combining (3.23) with Corollary 3.3, we conclude that for a large enough $c_7 = c_7(d,k) > c_6'$,

$$\mathbb{P}\left[\mathbb{1}\mathfrak{E} \cap \{\boldsymbol{X}^+ = -|\mathcal{V}^+|\} \cdot |\bar{\mathscr{L}}^+| > \log^{c_7} n\right] = O(n^{-2}). \tag{3.24}$$

In light of above, we see that

$$
\begin{aligned}
\mathbb{E}\left[\mathbb{1}\{\boldsymbol{X}_{\Lambda,L}=\boldsymbol{X}_{\Lambda^+,L}=1-L\}|\bar{\mathscr{L}}|^{3/2}\right] &\le \mathbb{E}\left[\mathbb{1}\mathfrak{E}\cap\{\boldsymbol{X}_{\Lambda,L}=\boldsymbol{X}_{\Lambda^+,L}=1-L\}|\bar{\mathscr{L}}|^{3/2}\right]+(2n)^{3/2}(1-\mathbb{P}[\mathfrak{E}]), && [\text{since } |\bar{\mathscr{L}}|\le 2n]\\
&\le \mathbb{E}\left[\mathbb{1}\mathfrak{E}\cap\{\boldsymbol{X}_{\Lambda,L}=\boldsymbol{X}_{\Lambda^+,L}=1-L\}(|\bar{\mathscr{L}}^+|+\log^{c_5'} n)^{3/2}\right]+o(1), && [\text{from (3.20)}]\\
&\le \mathbb{E}\left[\mathbb{1}\mathfrak{E}\cap\{\boldsymbol{X}^+=-|\mathscr{V}^+|\}(|\bar{\mathscr{L}}^+|+\log^{c_5'} n)^{3/2}\right]+o(1), && [\text{from (3.18), (3.19)}]\\
&\le \mathbb{P}\left[\mathbb{1}\mathfrak{E}\cap\{\boldsymbol{X}^+=-|\mathscr{V}^+|\}\cdot|\bar{\mathscr{L}}^+|>\log^{c_7} n\right](2n)^{3/2}\\
&\quad+\mathbb{P}\left[\{\boldsymbol{X}_{\Lambda,L}=1-L\}\right](2\log^{c_7} n)^{3/2}+o(1) && [\text{total probability}]\\
&=o(1) && [\text{from (3.24), Lemma 3.5}]
\end{aligned}
$$

completing the proof. □

Proof of Lemma 2.10. The assertion is immediate from Corollary 3.8, Lemma 3.9, Lemma 3.5 and the deterministic bound $|\bar{\mathscr{L}}|\le 2n$. □

4. Proof of Proposition 2.1

Let $\pi^{(\ell)}_{d,k}=\mathrm{BP}^{\ell}_{d,k}(\delta_{1/2})$ be the distribution obtained after ℓ iterations of $\mathrm{BP}_{d,k}(\cdot)$, with the convention $\pi^{(0)}_{d,k}=\delta_{1/2}$. We recall $(\boldsymbol{\mu}_{\pi,i,j})_{i,j\ge1}$ signify independent random variables with distribution π.

Fact 4.1. *For all $\ell\ge0$ the random variables $\boldsymbol{\mu}_{\pi^{(\ell)}_{d,k},1,1}$ and $1-\boldsymbol{\mu}_{\pi^{(\ell)}_{d,k},1,1}$ are identically distributed.*

Proof. This is an immediate consequence of the fact that the random variables $\boldsymbol{d}^+,\boldsymbol{d}^-$ from the definition (1.3)–(1.4) are identically distributed. □

While the following is a direct consequence of the fact that Belief Propagation is 'exact on trees' (see [50, Chapter 14] for precise statements), we carry out a detailed proof for the sake of completeness. Following the conventions from Section 1.2.1, we continue to denote by $\boldsymbol{\tau}^{(\ell)}$ a random satisfying assignment of the k-CNF $\mathbb{T}^{(\ell)}=\mathbb{T}^{(\ell)}_{d,k}$.

Fact 4.2. *For all $\ell\ge0$, $d>0$ we have $\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{x})=1\mid\mathbb{T}\right]\sim\pi^{(\ell)}_{d,k}$.*

Proof. We proceed by induction on ℓ. As $\pi^{(0)}_{d,k}=\delta_{1/2}$, for $\ell=0$ there is nothing to show. To go from $\ell-1$ to $\ell\ge1$, for a clause $a\in\partial_{\mathbb{T}}\mathfrak{x}$ and a variable $y\in\partial_{\mathbb{T}}a\setminus\{\mathfrak{x}\}$ let $\mathbb{T}_{y\to a}$ be the component of the forest $\mathbb{T}-a$ obtained by removing clause a that contains variable y. We consider y the root of $\mathbb{T}_{y\to a}$. Further, obtain $\mathbb{T}^{(\ell-1)}_{y\to a}$ from $\mathbb{T}_{y\to a}$ by deleting all clauses and variables at a distance greater than $2(\ell-1)$ from y. Additionally, for $s\in\{\pm1\}$ let

$$
\boldsymbol{Z}^{(\ell)}(s)=\left|\left\{\sigma\in S(\mathbb{T}^{(\ell)}):\sigma(\mathfrak{x})=s\right\}\right|,\qquad \boldsymbol{Z}^{(\ell-1)}_{y\to a}(s)=\left|\left\{\sigma\in S(\mathbb{T}^{(\ell-1)}_{y\to a}):\sigma(y)=s\right\}\right|. \tag{4.1}
$$

In words, $\boldsymbol{Z}^{(\ell)}(s)$ is the number of satisfying assignments of $\mathbb{T}^{(\ell)}$ that set the root $\mathfrak{x}$ to s, and $\boldsymbol{Z}^{(\ell-1)}_{y\to a}(s)$ is the corresponding quantity for the sub-tree $\mathbb{T}^{(\ell-1)}_{y\to a}$.

Clearly, setting $\mathfrak{x}$ to $s\in\{\pm1\}$ immediately satisfies all clauses $a\in\partial^{s}_{\mathbb{T}}\mathfrak{x}$. By contrast, once $\mathfrak{x}$ is assigned the value $+1$ each clause $a\in\partial^{-s}_{\mathbb{T}}\mathfrak{x}$ needs to be satisfied by setting some other variable $y\in\partial_{\mathbb{T}}a\setminus\{\mathfrak{x}\}$ to the value $\mathrm{sign}(y,a)$. Hence,

$$
\boldsymbol{Z}^{(\ell)}(s)=\left[\prod_{a\in\partial^+_{\mathbb{T}}\mathfrak{x}}\prod_{y\in\partial_{\mathbb{T}}a\setminus\{\mathfrak{x}\}}\sum_{t\in\{\pm1\}}\boldsymbol{Z}^{(\ell-1)}_{y\to a}(t)\right]\cdot\left[\prod_{a\in\partial^-_{\mathbb{T}}\mathfrak{x}}\left(\prod_{y\in\partial_{\mathbb{T}}a\setminus\{\mathfrak{x}\}}\sum_{t\in\{\pm1\}}\boldsymbol{Z}^{(\ell-1)}_{y\to a}(t)-\prod_{y\in\partial_{\mathbb{T}}a\setminus\{\mathfrak{x}\}}\boldsymbol{Z}^{(\ell-1)}_{y\to a}(-\mathrm{sign}(y,a))\right)\right]. \tag{4.2}
$$

Furthermore, the definition of the Galton-Watson tree $\mathbb{T}$ ensures that the sub-trees $\mathbb{T}^{(\ell-1)}_{y\to a}$ are independent copies of $\mathbb{T}^{(\ell-1)}$. Hence, by induction we have

$$
\frac{\boldsymbol{Z}^{(\ell-1)}_{y\to a}(1)}{\sum_{s\in\{\pm1\}}\boldsymbol{Z}^{(\ell-1)}_{y\to a}(s)}\sim\pi^{(\ell-1)}_{d,k}\qquad\text{for all } a\in\partial_{\mathbb{T}}\mathfrak{x},\ y\in\partial_{\mathbb{T}}a\setminus\{\mathfrak{x}\}, \tag{4.3}
$$

and the random variables $\boldsymbol{Z}^{(\ell-1)}_{y\to a}(1)/\sum_{s\in\{\pm1\}}\boldsymbol{Z}^{(\ell-1)}_{y\to a}(s)$ are mutually independent. Combining (4.2)–(4.3) with Fact 4.1, we finally obtain

$$\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r})=1\mid\mathbb{T}\right]=\frac{\boldsymbol{Z}^{(\ell)}(1)}{\sum_{s\in\{\pm1\}}\boldsymbol{Z}^{(\ell)}(s)}\sim\frac{\prod_{i=1}^{\boldsymbol{d}^-}\left[1-\prod_{j=1}^{k-1}\boldsymbol{\mu}_{\pi^{(\ell-1)}_{d,k},2i-1,j}\right]}{\prod_{i=1}^{\boldsymbol{d}^-}\left[1-\prod_{j=1}^{k-1}\boldsymbol{\mu}_{\pi^{(\ell-1)}_{d,k},2i-1,j}\right]+\prod_{i=1}^{\boldsymbol{d}^+}\left[1-\prod_{j=1}^{k-1}\boldsymbol{\mu}_{\pi^{(\ell-1)}_{d,k},2i,j}\right]}\sim\pi^{(\ell)}_{d,k},$$

thereby completing the induction. □

Combining the combinatorial interpretation of the distributions $\pi^{(\ell)}_{d,k}$ with the Gibbs uniqueness property, we proceed to show that the sequence $(\pi^{(\ell)}_{d,k})_\ell$ converges in the weak topology. To this end, it suffices to show that the sequence is Cauchy with respect to the Wasserstein W_1 metric.

Lemma 4.3. *If $d<d_{\mathrm{uniq}}(k)$ then $(\pi^{(\ell)}_{d,k})_{\ell\ge0}$ is a W_1-Cauchy sequence.*

Proof. If $d<d_{\mathrm{uniq}}(k)$ then the random tree $\mathbb{T}=\mathbb{T}_{d,k}$ enjoys the Gibbs uniqueness property; hence, (1.1) is satisfied. Consequently, given $0<\varepsilon<1$ we can choose $\ell_0=\ell_0(d,k,\varepsilon)>0$ large enough so that the event

$$\mathfrak{U}_{\varepsilon,\ell}=\left\{\max_{\tau\in S(\mathbb{T}^{(\ell)})}\left|\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r})=1\mid\mathbb{T}\right]-\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r})=1\mid\mathbb{T},\forall x\in\partial^{2\ell}\mathfrak{r}:\boldsymbol{\tau}^{(\ell)}(x)=\tau(x)\right]\right|>\varepsilon\right\}$$

has probability

$$\mathbb{P}\left[\mathfrak{U}_{\varepsilon,\ell}\right]<\varepsilon\ ,\qquad\qquad \text{for all } \ell\ge\ell_0. \tag{4.4}$$

Now suppose that $\ell_0\le\ell<\ell'$. Let $\boldsymbol{\tau}^{(\ell)},\boldsymbol{\tau}^{(\ell')}$ be independent uniformly random satisfying assignments of $\mathbb{T}^{(\ell)}$ and $\mathbb{T}^{(\ell')}$, respectively. We claim that

$$\mathbb{P}\left[\left|\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r})=1\mid\mathbb{T}\right]-\mathbb{P}\left[\boldsymbol{\tau}^{(\ell')}(\mathfrak{r})=1\mid\mathbb{T}\right]\right|>\varepsilon\right]<\varepsilon. \tag{4.5}$$

To see this, let $\boldsymbol{\tau}^{(\ell,\ell')}=(\boldsymbol{\tau}^{(\ell')}(x))_{x\in\partial^{2\ell}_{\mathbb{T}}\mathfrak{r}}$ comprise the truth values that $\boldsymbol{\tau}^{(\ell')}$ assigns to the variables at distance exactly 2ℓ from $\mathfrak{r}$. Then

$$\begin{aligned}\mathbb{P}\left[\boldsymbol{\tau}^{(\ell')}(\mathfrak{r})=1\mid\mathbb{T}\right]&=\mathbb{E}\left[\mathbb{P}\left[\boldsymbol{\tau}^{(\ell')}(\mathfrak{r})=1\mid\mathbb{T},\boldsymbol{\tau}^{(\ell,\ell')}\right]\mid\mathbb{T}\right]\\&=\mathbb{E}\left[\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r})=1\mid\mathbb{T},\boldsymbol{\tau}^{(\ell,\ell')},\forall x\in\partial^{2\ell}\mathfrak{r}:\boldsymbol{\tau}^{(\ell)}(x)=\boldsymbol{\tau}^{(\ell,\ell')}_x\right]\mid\mathbb{T}\right].\end{aligned}$$

Hence, for every $T\in\mathfrak{U}_{\varepsilon,\ell}$ we have

$$\left|\mathbb{P}\left[\boldsymbol{\tau}^{(\ell')}(\mathfrak{r})=1\mid\mathbb{T}=T\right]-\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r})=1\mid\mathbb{T}=T\right]\right|\le\varepsilon. \tag{4.6}$$

Thus, (4.5) follows from (4.4) and (4.6).

Finally, since Fact 4.2 demonstrates that $\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r})=1\mid\mathbb{T}=T\right]\sim\pi^{(\ell)}_{d,k}$ and $\mathbb{P}\left[\boldsymbol{\tau}^{(\ell')}(\mathfrak{r})=1\mid\mathbb{T}=T\right]\sim\pi^{(\ell')}_{d,k}$, (4.5) shows that

$$W_1\left(\pi^{(\ell)}_{d,k},\pi^{(\ell')}_{d,k}\right)<2\varepsilon\qquad\qquad\text{for all } \ell_0\le\ell<\ell'.$$

Hence, the sequence $(\pi^{(\ell)}_{d,k})_\ell$ is Cauchy. □

We are left to bound the lower tail of the limiting distribution $\pi_{d,k}=\lim_{\ell\to\infty}\pi^{(\ell)}_{d,k}$.

Lemma 4.4. *If $d<d_{\mathrm{uniq}}(k)$ then $\mathbb{E}\log^2\boldsymbol{\mu}_{\pi_{d,k},1,1}<\infty$.*

Proof. We are going to bound $\mathbb{E}\log^2\boldsymbol{\mu}_{\pi^{(\ell)}_{d,k},1,1}$ and subsequently invoke the monotone convergence theorem to complete the proof. First, we note that for all $\ell\ge0$ we have

$$\begin{aligned}\mathbb{E}\log^2\boldsymbol{\mu}_{\pi^{(\ell)}_{d,k},1,1}&=\mathbb{E}\log^2\frac{Z(\mathbb{T}^{(\ell)},\{\mathfrak{r}\})}{Z(\mathbb{T}^{(\ell)})}&&\text{[by Fact 4.2]}\\&\le\mathbb{E}\left[|\overline{\{\mathfrak{r}\}}_{\mathbb{T}}|^2\right]&&\text{[by Lemma 2.8]}\\&<\infty&&\text{[by Corollary 3.4].}\end{aligned}$$

Since $\pi_{d,k}$ is the weak limit of $(\pi^{(\ell)}_{d,k})_\ell$, we conclude that for any $N \in \mathbb{N}$,

$$\mathbb{E}\left[N \wedge \log^2 \boldsymbol{\mu}_{\pi_{d,k},1,1}\right] = \lim_{\ell\to\infty} \mathbb{E}\left[N \wedge \log^2 \boldsymbol{\mu}_{\pi^{(\ell)}_{d,k},1,1}\right] \le \mathbb{E}\left[|\overline{\{\mathfrak{r}\}}_{\mathbb{T}}|^2\right] < \infty. \tag{4.7}$$

Finally, applying the monotone convergence theorem to the limit $N \to \infty$, we see that the uniform bound (4.7) implies the assertion. □

Proof of Proposition 2.1. In light of Fact 4.1 and Lemmas 4.3 and 4.4, it only remains to show that

$$\mathbb{E}\left|\log\left(\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\pi_{d,k},2i} + \prod_{i=1}^{\boldsymbol{d}^+} \boldsymbol{\mu}_{\pi_{d,k},2i-1}\right)\right| < \infty \qquad \text{and} \qquad \mathbb{E}\left|\log\left(1 - \prod_{j=1}^{k} \boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right| < \infty \ .$$

Recall the definition of $\boldsymbol{\mu}_{\pi_{d,k},i}$ from (1.4). Using Fact 4.1 and Lemma 4.4, we obtain

$$\begin{aligned}\mathbb{E}\left|\log\left(\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\pi_{d,k},2i} + \prod_{i=1}^{\boldsymbol{d}^+} \boldsymbol{\mu}_{\pi_{d,k},2i-1}\right)\right| &\le \log(2) + \mathbb{E}\left|\log \prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\pi_{d,k},2i}\right| \le \log(2) + \frac{d}{2}\mathbb{E}\left|\log \boldsymbol{\mu}_{\pi_{d,k},1}\right| \\ &\le \log(2) + \frac{d}{2}\sqrt{\mathbb{E}\left|\log^2 \boldsymbol{\mu}_{\pi_{d,k},1,1}\right|} < \infty \ ,\end{aligned}$$

yielding the first inequality. Similarly, invoking Fact 4.1 and Lemma 4.4 for the second l.h.s. above gives

$$\mathbb{E}\left|\log\left(1 - \prod_{j=1}^{k} \boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right| \le \mathbb{E}\left|\log\left(1 - \boldsymbol{\mu}_{\pi_{d,k},1,1}\right)\right| = \mathbb{E}\left|\log\left(\boldsymbol{\mu}_{\pi_{d,k},1,1}\right)\right| \le \sqrt{\mathbb{E}\left|\log^2 \boldsymbol{\mu}_{\pi_{d,k},1,1}\right|} < \infty \ ,$$

thereby completing the proof. □

5. Proof of Corollary 2.2

In order to turn the estimate of the expectation of $\log 1 \vee Z(\boldsymbol{\Phi})$ provided by Proposition 2.3 into a 'with high probability' statement, we harness a 'soft' version of the k-SAT problem where violated clauses are discouraged but not strictly forbidden. To be precise, for a k-CNF Φ and a real $\beta > 0$ define

$$Z_\beta(\Phi) = \sum_{\sigma\in\{\pm1\}^{V(\Phi)}} \prod_{a\in F(\Phi)} \exp(-\beta \mathbb{1}\{\sigma \not\models a\}). \tag{5.1}$$

Thus, each satisfying assignment contributes one to the sum on the r.h.s. of (5.1), while the contribution of assignments that violate a number M of clauses equals $\exp(-\beta M)$. The value $Z_\beta(\boldsymbol{\Phi})$, called the *partition function of the random k-SAT model at inverse temperature β*, has received a considerable amount of attention in the mathematical physics literature (see, e.g., [57]). Crucially, by means of an interpolation argument [35, 41] it is possible to prove the following.

Theorem 5.1 ([60, Theorem 1])**.** *For any $k \ge 3$, any $\beta > 0$ and any probability measure π on $[0,1]$ we have*

$$\frac{1}{n}\mathbb{E}\left[\log Z_\beta(\boldsymbol{\Phi})\right] \le \mathbb{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\beta,\pi,2i} + \prod_{i=1}^{\boldsymbol{d}^+} \boldsymbol{\mu}_{\beta,\pi,2i-1}\right) - \frac{d(k-1)}{k}\log\left(1 - \left(1-e^{-\beta}\right)\prod_{j=1}^{k} \boldsymbol{\mu}_{\pi,1,j}\right)\right], \qquad \textit{where} \tag{5.2}$$

$$\boldsymbol{\mu}_{\beta,\pi,i} = 1 - (1-\exp(-\beta)) \prod_{j=1}^{k-1} \boldsymbol{\mu}_{\pi,i,j} \qquad \textit{for } i \ge 1.$$

We emphasise that the bound (5.2) holds for any $n \ge k$ *without* an error term. We also notice that by the monotone convergence theorem for the measure $\pi = \pi_{d,k}$ from Theorem 1.1 we have

$$\begin{aligned}&\lim_{\beta\to\infty} \mathbb{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\beta,\pi_{d,k},2i} + \prod_{i=1}^{\boldsymbol{d}^+} \boldsymbol{\mu}_{\beta,\pi_{d,k},2i-1}\right) - \frac{d(k-1)}{k}\log\left(1 - \left(1-e^{-\beta}\right)\prod_{j=1}^{k} \boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right] \\ &= \mathbb{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}^-} \boldsymbol{\mu}_{\pi_{d,k},2i} + \prod_{i=1}^{\boldsymbol{d}^+} \boldsymbol{\mu}_{\pi_{d,k},2i-1}\right) - \frac{d(k-1)}{k}\log\left(1 - \prod_{j=1}^{k} \boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right] = \mathfrak{B}_{d,k}(\pi_{d,k}) \ .\end{aligned} \tag{5.3}$$

The reason why we proceed by way of the 'soft' model with $\beta < \infty$ is that for this model a routine application of Azuma-Hoeffding implies the following concentration bound.

Lemma 5.2. *For any fixed $\beta > 0$ we have* $\mathbb{P}\left[\left|\log Z_\beta(\boldsymbol{\Phi}) - \mathbb{E}\log Z_\beta(\boldsymbol{\Phi})\right| > \sqrt{n}\log n\right] = o(1/n)$.

Proof. The clauses of the random formula $\mathbf{\Phi}$ are drawn independently, and adding or removing a single clause can alter the value of $\log Z_\beta(\cdot)$ by no more than $\pm\beta$. □

Proof of Corollary 2.2. We proceed with a proof by contradiction. In particular, towards a contradiction, assume there exists an $\varepsilon > 0$ such that for infinitely many $n \geq 1$ we have

$$\mathbb{P}\left[\frac{1}{n}\log Z(\mathbf{\Phi}) > \mathfrak{B}_{d,k}(\pi_{d,k}) + \varepsilon\right] > \varepsilon \ . \tag{5.4}$$

Moreover, by (5.3) we can find a $\beta_0 > 0$ such that for every $\beta \geq \beta_0$ we have

$$\left|\mathbb{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}^-}\boldsymbol{\mu}_{\beta,\pi_{d,k},2i} + \prod_{i=1}^{\boldsymbol{d}^+}\boldsymbol{\mu}_{\beta,\pi_{d,k},2i-1}\right) - \frac{d(k-1)}{k}\log\left(1-\left(1-e^{-\beta}\right)\prod_{j=1}^{k}\boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right] - \mathfrak{B}_{d,k}(\pi_{d,k})\right| < \varepsilon/3 \ . \tag{5.5}$$

Invoking Lemma 5.2 for $\beta = \beta_0$ and sufficiently large n gives

$$\mathbb{P}\left[\frac{1}{n}\log Z_{\beta_0}(\mathbf{\Phi}) > \frac{1}{n}\mathbb{E}\log Z_{\beta_0}(\mathbf{\Phi}) + \varepsilon/3\right] \leq \varepsilon/3 \ . \tag{5.6}$$

The definition (5.1) of the partition function ensures that $Z_\beta(\mathbf{\Phi}) \geq Z(\mathbf{\Phi})$ for all $\beta > 0$. Therefore, combining (5.4)–(5.6), and Theorem 5.1 we see that for large enough n the following holds with probability at least $1 - \frac{2}{3}\varepsilon$:

$$\frac{1}{n}\log Z(\mathbf{\Phi}) \leq \frac{1}{n}\log Z_{\beta_0}(\mathbf{\Phi}) \leq \frac{1}{n}\mathbb{E}\log Z_{\beta_0}(\mathbf{\Phi}) + \frac{\varepsilon}{3} \leq \mathfrak{B}_{d,k}(\pi_{d,k}) + \frac{2}{3}\varepsilon \ ,$$

contradicting our assumption, and thus completing the proof. □

6. Proof of Proposition 2.3

In this section we prove Propositions 2.5 and 2.6, which in light of Fact 2.4, imply Proposition 2.3. Both proofs follow a similar structure and make use of Propositions 2.7 and 2.11, which we therefore prove first.

6.1. **Proof of Proposition 2.7.** We show that both terms of (2.5) have finite expectation. Let us begin with the first one.

Lemma 6.1. *If $d < d_{\mathrm{uniq}}(k)$ then $\mathbb{E}\left[\left|\log\frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right] = O(1)$.*

Proof. Since $\mathbf{\Phi}''$ is obtained from $\mathbf{\Phi}'$ by adding clauses, we have

$$0 \leq Z(\mathbf{\Phi}'') \leq Z(\mathbf{\Phi}') \leq 2^n. \tag{6.1}$$

Hence,

$$\log\frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1} = 0 \qquad\qquad \text{if } Z(\mathbf{\Phi}') = 0. \tag{6.2}$$

Therefore, we may assume from now on that $\mathbf{\Phi}'$ is satisfiable.

The number $\mathbf{\Delta}'' \sim \mathrm{Po}(d(k-1)/k)$ of new clauses is a Poisson variable with bounded mean. Therefore, Bennett's inequality shows that $\mathbb{P}\left[\mathbf{\Delta}'' > \log n\right] = O(n^{-2})$. Since (6.1) shows that $|\log((Z(\mathbf{\Phi}'')\vee 1)/(Z(\mathbf{\Phi}')\vee 1))|^{3/2} \leq n^{3/2}$, we conclude that

$$\mathbb{E}\left[\mathbb{1}\left\{\mathbf{\Delta}'' > \log n\right\}\cdot\left|\log\frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right] = o(1). \tag{6.3}$$

Further, let $c_1,\ldots,c_{\mathbf{\Delta}''}$ be the new clauses added by **CPL2**. Let $\boldsymbol{x}_{1,1},\ldots,\boldsymbol{x}_{1,k},\ldots,\boldsymbol{x}_{\mathbf{\Delta}'',1},\ldots,\boldsymbol{x}_{\mathbf{\Delta}'',k}$ be their constituent variables and let $\mathscr{X} = \{\boldsymbol{x}_{1,1},\ldots,\boldsymbol{x}_{1,k},\ldots,\boldsymbol{x}_{\mathbf{\Delta}'',1},\ldots,\boldsymbol{x}_{\mathbf{\Delta}'',k}\}$. Since the clauses $c_1,\ldots,c_{\mathbf{\Delta}''}$ are chosen uniformly and independently, a routine balls-into-bins consideration shows that

$$\mathbb{P}\left[|\mathscr{X}| \leq k(\mathbf{\Delta}''-1) \mid \mathbf{\Delta}'' \leq \log n\right] = \tilde{O}(n^{-2}). \tag{6.4}$$

Now, consider the 'good' event

$$\mathfrak{G} = \left\{Z(\mathbf{\Phi}') > 0, \mathbf{\Delta}'' \leq \log n, |\mathscr{X}| > k(\mathbf{\Delta}''-1)\right\}.$$

Combining (6.1)–(6.4), we see that

$$\mathbb{E}\left[(1-\mathbb{1}\mathfrak{G})\cdot\left|\log\frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right]=o(1). \tag{6.5}$$

Hence, we are left to bound $\mathbb{E}[\mathbb{1}\mathfrak{G}\cdot|\log((Z(\mathbf{\Phi}'')\vee 1)/(Z(\mathbf{\Phi}')\vee 1))|^{3/2}]$. If $\mathfrak{G}$ occurs and thus $|\mathscr{X}|>k(\mathbf{\Delta}''-1)$, then there exists a set of literals $\mathscr{L}\subseteq\left\{\boldsymbol{x}_{1,1},\neg\boldsymbol{x}_{1,1},\ldots,\boldsymbol{x}_{1,k},\neg\boldsymbol{x}_{1,k},\ldots,\boldsymbol{x}_{\mathbf{\Delta}'',1},\neg\boldsymbol{x}_{\mathbf{\Delta}'',1},\ldots,\boldsymbol{x}_{\mathbf{\Delta}'',k},\neg\boldsymbol{x}_{\mathbf{\Delta}'',k}\right\}$ such that

- every clause c_i contains a literal from $\mathscr{L}$ ($1\le i\le\mathbf{\Delta}''$), and
- there does not exist $x\in\mathscr{X}$ such that $x\in\mathscr{L}$ and $\neg x\in\mathscr{L}$.

Moreover, on $\mathfrak{G}$ we have $|\mathscr{L}|\le|\mathscr{X}|\le k\log n$. Let $\bar{\mathscr{L}}=\bar{\mathscr{L}}_{\mathbf{\Phi}'}$ be the output of PULP on $(\mathbf{\Phi}',\mathscr{L})$. Then Lemma 2.8 shows that

$$\mathbb{E}\left[\mathbb{1}\mathfrak{G}\cdot\left|\log\frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right]\le\mathbb{E}\left[\mathbb{1}\mathfrak{G}\cdot\left|\bar{\mathscr{L}}\right|^{3/2}\right]. \tag{6.6}$$

Furthermore, since by **CPL2** the new clauses $c_1,\ldots,c_{\mathbf{\Delta}''}$ are chosen independently of the formula $\mathbf{\Phi}'$, Lemma 2.10 implies that there exists $C=C(d,k)>0$ such that

$$\mathbb{E}\left[\mathbb{1}\mathfrak{G}\cdot\left|\bar{\mathscr{L}}\right|^{3/2}\mid\mathbf{\Delta}''\right]\le C\cdot\left(\mathbf{\Delta}''\right)^{3/2}. \tag{6.7}$$

Combining (6.6)–(6.7) and recalling that $\mathbf{\Delta}''\sim\mathrm{Po}(d(k-1)/k)$, we obtain

$$\mathbb{E}\left[\mathbb{1}\mathfrak{G}\cdot\left|\log\frac{Z(\mathbf{\Phi}'')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right]=O(1). \tag{6.8}$$

Finally, the assertion follows from (6.5) and (6.8). □

We move on to the second term of (2.5).

Lemma 6.2. *If $d<d_{\mathrm{uniq}}(k)$ then* $\mathbb{E}\left[\left|\log\frac{Z(\mathbf{\Phi}''')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right]=O(1)$.

Proof. We proceed similarly as in the proof of Lemma 6.1. The construction in **CPL3** ensures that $\mathbf{\Phi}'''$ contains one additional variable x_{n+1} and $\mathbf{\Delta}'''\sim\mathrm{Po}(d)$ new clauses $b_1,\ldots,b_{\mathbf{\Delta}'''}$ that each contain x_{n+1} and $k-1$ other variables. Let $\boldsymbol{x}_{1,1},\ldots,\boldsymbol{x}_{1,k-1},\ldots,\boldsymbol{x}_{\mathbf{\Delta}''',1},\ldots,\boldsymbol{x}_{\mathbf{\Delta}''',k-1}\in\{x_1,\ldots,x_n\}$ be the variables among $x_1,\ldots,x_n$ that appear in $b_1,\ldots,b_{\mathbf{\Delta}'''}$ and let $\mathscr{X}=\{\boldsymbol{x}_{1,1},\ldots,\boldsymbol{x}_{\mathbf{\Delta}''',k-1}\}$. Then

$$0\le Z(\mathbf{\Phi}'')\le 2Z(\mathbf{\Phi}')\le 2^{n+1}. \tag{6.9}$$

Hence, if $\mathbf{\Phi}'$ is unsatisfiable, then so is $\mathbf{\Phi}'''$ and thus

$$\log\frac{Z(\mathbf{\Phi}''')\vee 1}{Z(\mathbf{\Phi}')\vee 1}=0 \qquad\qquad \text{if } Z(\mathbf{\Phi}')=0. \tag{6.10}$$

Furthermore, since $\mathbf{\Delta}'''\sim\mathrm{Po}(d)$, Bennett's inequality shows that $\mathbb{P}\left[\mathbf{\Delta}'''>\log n\right]=O(n^{-2})$. Therefore, (6.9) shows that

$$\mathbb{E}\left[\mathbb{1}\left\{\mathbf{\Delta}'''>\log n\right\}\cdot\left|\log\frac{Z(\mathbf{\Phi}''')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right]=o(1). \tag{6.11}$$

Moreover, since the $k-1$ variables among $x_1,\ldots,x_n$ that appear in the clauses $b_1,\ldots,b_{\Delta'''}$ are chosen uniformly and independently, a simple balls-into-bins argument shows that

$$\mathbb{P}\left[|\mathscr{X}|\le(k-1)(\mathbf{\Delta}''-1)\mid\mathbf{\Delta}''\le\log n\right]=\tilde{O}(n^{-2}). \tag{6.12}$$

Hence, consider the event

$$\mathfrak{G}=\left\{Z(\mathbf{\Phi}')>0,\,\mathbf{\Delta}'''\le\log n,\,|\mathscr{X}|>(k-1)(\mathbf{\Delta}''-1)\right\}.$$

Combining (6.9)–(6.12), we obtain

$$\mathbb{E}\left[(1-\mathbb{1}\mathfrak{G})\cdot\left|\log\frac{Z(\mathbf{\Phi}''')\vee 1}{Z(\mathbf{\Phi}')\vee 1}\right|^{3/2}\right]=o(1). \tag{6.13}$$

Furthermore, if the event $\mathfrak{G}$ occurs, then there exists a set $\mathscr{L} \subseteq \{x, \neg x : x \in \mathscr{X}\}$ of literals such that each clause b_i, $1 \le i \le \mathbf{\Delta}'''$, contains a literal $l \in \mathscr{L}$ and such that $\{x, \neg x\} \not\subseteq \mathscr{L}$ for all $x \in \mathscr{X}$. Hence, with $\bar{\mathscr{L}} = \bar{\mathscr{L}}_{\mathbf{\Phi}'}$ the output of PULP on $(\mathbf{\Phi}', \mathscr{L})$, Lemma 2.8 shows that

$$\mathbb{E}\left[\mathbb{1}\mathfrak{G} \cdot \left|\log \frac{Z(\mathbf{\Phi}''') \vee 1}{Z(\mathbf{\Phi}') \vee 1}\right|^{3/2}\right] \le \mathbb{E}\left[\mathbb{1}\mathfrak{G} \cdot |\bar{\mathscr{L}}|^{3/2}\right]. \tag{6.14}$$

Furthermore, since the clauses $b_1, \dots, b_{\mathbf{\Delta}'''}$ are drawn independently of $\mathbf{\Phi}'''$, Lemma 2.10 shows that there exists $C = C(d,k) > 0$ such that

$$\mathbb{E}\left[\mathbb{1}\mathfrak{G} \cdot |\bar{\mathscr{L}}|^{3/2} \mid \mathbf{\Delta}'''\right] \le C \cdot (\mathbf{\Delta}''')^{3/2}. \tag{6.15}$$

Finally, since $\mathbf{\Delta}''' \sim \mathrm{Po}(d)$, the assertion follows from (6.13), (6.14) and (6.15). □

Proof of Proposition 2.7. The proposition follows immediately from Lemmas 6.1–6.2. □

6.2. **Proof of Proposition 2.11.** Let $\pi^{(\ell)}_{d,k} = \mathrm{BP}^{\ell}_{d,k}(\delta_{1/2})$ be the result of an ℓ-fold application of the operator $\mathrm{BP}_{d,k}$ from (1.3) to the point mass at $1/2$. Also recall from (2.7) that $\boldsymbol{\pi}'_n$ denotes the empirical distribution of the marginals $(\mathbb{P}[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i) = 1 \mid \mathbf{\Phi}'])_{1 \le i \le n}$.

Lemma 6.3. *Suppose that $d < d_{\mathrm{uniq}}(k)$. For any $\varepsilon > 0$ there exists $\ell_0 = \ell_0(d,k,\varepsilon) > 0$ such that for all $\ell \ge \ell_0$ we have*

$$\mathbb{E}[W_1(\boldsymbol{\pi}'_n, \pi^{(\ell)}_{d,k}) \mid Z(\mathbf{\Phi}') > 0] < \varepsilon + o(1).$$

Proof. Assume that $\ell \ge \ell_0$ for a large enough $\ell_0 = \ell_0(d,k,\varepsilon) > 0$. Since $d < d_{\mathrm{uniq}}(k)$ and since $\mathbb{T} = \mathbb{T}_{d,k}$ is a Galton-Watson tree in which every variable node has $\mathrm{Po}(d)$ clause nodes as offspring and the offspring of every clause node consists of $k-1$ variable nodes, there exists a set $\mathscr{T}_\ell$ of trees, with $|\mathscr{T}_\ell| = O(1)$, such that the following hold:

- **T0:** for every $T \in \mathscr{T}_\ell$ we have $\mathbb{P}\left[\mathbb{T}^{(\ell)} = T\right] > 0$.
- **T1:** $\mathbb{P}\left[\mathbb{T}^{(\ell)} \in \mathscr{T}_\ell\right] > 1 - \varepsilon$.
- **T2:** given $\mathbb{T}^{(\ell)} \in \mathscr{T}_\ell$ we have

$$\max_{\tau \in S(\mathbb{T}^{(\ell)})} \left|\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \mathbb{T}^{(\ell)}\right] - \mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \mathbb{T}^{(\ell)}, \forall x \in \partial^{2\ell}\mathfrak{r} : \boldsymbol{\tau}^{(\ell)}(x) = \tau(x)\right]\right| < \varepsilon.$$

For a variable node x_i of $\mathbf{\Phi}'$ obtain $\boldsymbol{\phi}'_\ell(x_i)$ from $\mathbf{\Phi}'$ by deleting all variables and clauses at distance greater than 2ℓ from x_i. We consider x_i being the root of $\boldsymbol{\phi}'_\ell(x_i)$. Moreover, for a tree $T \in \mathscr{T}_\ell$ let $\mathscr{V}_T$ be the set of variable nodes x_i, $1 \le i \le n$, such that $\boldsymbol{\phi}'_\ell(x_i) \cong T$; thus, there is an isomorphism of the CNFs T and $\boldsymbol{\phi}'_\ell(x_i)$ that maps the root $\mathfrak{r}$ of T to x_i. Consider the event

$$\mathfrak{T}_\ell = \left\{\sum_{T \in \mathscr{T}_\ell} \left|\mathbb{P}\left[\mathbb{T}^{(\ell)} \cong T\right] - |\mathscr{V}_T|/n\right| < \varepsilon\right\}. \tag{6.16}$$

Then Corollary 3.7 implies that

$$\mathbb{P}[\mathfrak{T}_\ell] = 1 - o(1) \qquad \text{for every } \ell \ge 0. \tag{6.17}$$

We now claim that

$$\left|\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \mathbb{T}^{(\ell)} = T\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i) = 1 \mid \mathbf{\Phi}'\right]\right| < \varepsilon \qquad \text{for all } T \in \mathscr{T}_\ell,\ x_i \in \mathscr{V}_T. \tag{6.18}$$

To see this, let $S_\ell(\mathbf{\Phi}', x_i)$ be the set of all assignments $\sigma \in \{\pm 1\}^{\partial^{2\ell} x_i}$ of the variables at distance 2ℓ from x_i in $\mathbf{\Phi}'$ such that there exists a satisfying assignment $\sigma' \in S(\mathbf{\Phi}')$ with $\sigma'(y) = \sigma(y)$ for all $y \in \partial^{2\ell} x_i$. Then the law of total probability shows that

$$\mathbb{P}\left[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i) = 1 \mid \mathbf{\Phi}'\right] = \sum_{\sigma \in S_\ell(\mathbf{\Phi}', x_i)} \mathbb{P}\left[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i) = 1 \mid \mathbf{\Phi}', \forall y \in \partial^{2\ell} x_i : \boldsymbol{\sigma}_{\mathbf{\Phi}'}(y) = \sigma(y)\right] \mathbb{P}\left[\forall y \in \partial^{2\ell} x_i : \boldsymbol{\sigma}_{\mathbf{\Phi}'}(y) = \sigma(y) \mid \mathbf{\Phi}'\right]. \tag{6.19}$$

Further, since for $T \in \mathscr{T}_\ell$ and $x_i \in \mathscr{V}_T$ we have $\boldsymbol{\phi}'_\ell(x_i) \cong T$, condition **T2** implies that

$$\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\mathbf{\Phi}'}(x_i) = 1 \mid \mathbf{\Phi}', \forall y \in \partial^{2\ell} x_i : \boldsymbol{\sigma}_{\mathbf{\Phi}'}(y) = \sigma(y)\right] - \mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \mathbb{T}^{(\ell)} = T\right]\right| < \varepsilon. \tag{6.20}$$

Combining (6.19) and (6.20), we obtain (6.18).

To complete the proof, we recall from Fact 4.2 that $\pi^{(\ell)}_{d,k}$ is precisely the distribution of $\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \mathbb{T}^{(\ell)}\right]$. Therefore, coupling the formulas $\mathbb{T}^{(\ell)}, \boldsymbol{\Phi}'$ on the event $\mathfrak{T}_\ell$ we have

$$\begin{aligned} W_1(\boldsymbol{\pi}'_n, \pi^{(\ell)}_{d,k}) &\le \mathbb{P}\left[\mathbb{T}^{(\ell)} \notin \mathcal{T}_\ell\right] + \frac{1}{n}\sum_{T\in\mathcal{T}_\ell}\sum_{x\in\mathcal{V}_T}\left|\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \mathbb{T}^{(\ell)} = T\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x) = 1 \mid \boldsymbol{\Phi}'\right]\right| + \varepsilon \quad &&\text{[by (6.16)]}\\ &\le 3\varepsilon &&\text{[by } \mathbf{T1} \text{ and (6.18)].} \end{aligned}$$

Combining this bound with (6.17) completes the proof. □

Proof of Proposition 2.11. The first assertion follows from Proposition 2.1, Lemma 6.3 and the fact that, since $0 < d < d_{\mathrm{uniq}}(k) < d_{\mathrm{sat}}(k)$, we have that $\mathbb{P}\left[Z(\boldsymbol{\Phi}') > 0\right] = 1 - o(1)$.

The second follows a routine argument, which we present below for the case $\ell = 2$ and it is standard to extend to any finite ℓ (see [65, Proposition 2.5]). Let $t = \Theta(\log\log n)$ and recall the definitions of $\boldsymbol{\phi}'_t(x_i)$, $\mathcal{T}_t$ and $S_t(\boldsymbol{\Phi}', x_i)$ from the proof of Lemma 6.3. Consider the event $\mathfrak{D} = \{\boldsymbol{\phi}'_t(x_1), \boldsymbol{\phi}'_t(x_2) \text{ are disjoint tree formulas}\}$.

From Lemma 3.5, we have that $\mathbb{P}[\mathfrak{D}] = 1 - o(1)$. On the event $\mathfrak{D}$, Lemma 6.3 implies that for every $\sigma_1, \sigma_2 \in \{\pm 1\}$, and $\tau_1 \in S_t(\boldsymbol{\Phi}', x_1), \tau_2 \in S_t(\boldsymbol{\Phi}', x_2)$ we have

$$|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_i) = \sigma_i \mid \boldsymbol{\Phi}', \forall y \in \partial^{2t}x_i : \boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(y) = \tau_i(y)\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_i) = \sigma_i \mid \boldsymbol{\Phi}'\right]| = o(1) \;, \qquad \text{for } i = 1,2 \;. \tag{6.21}$$

Therefore, from the law of total probability and the triangle inequality we see that for every $\sigma_1, \sigma_2 \in \{\pm 1\}$

$$\begin{aligned} &\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_1) = \sigma_1, \boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}'\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_1) = \sigma_1 \mid \boldsymbol{\Phi}'\right]\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}'\right]\right|\\ &\le \Big|\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_1) = \sigma_1 \mid \boldsymbol{\Phi}', \boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2\right] - \mathbb{E}_{\tau_1,\tau_2}\left[\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_1) = \sigma_1, \boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}', \tau_1, \tau_2\right]\right]\right|\\ &- \left|\mathbb{E}_{\tau_1,\tau_2}\left[\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_1) = \sigma_1 \mid \boldsymbol{\Phi}', \tau_1\right]\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}', \tau_2\right]\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}'\right]\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}'\right]\right|\Big|\\ &= \Big|\left|\mathbb{E}_{\tau_1,\tau_2}\left[\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_1) = \sigma_1 \mid \boldsymbol{\Phi}', \tau_1\right]\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}', \tau_2\right]\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}'\right]\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}'\right]\right|\Big|\\ &\le \mathbb{E}_{\tau_1}\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_1) = \sigma_1 \mid \boldsymbol{\Phi}', \tau_1\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_1) = \sigma_1 \mid \boldsymbol{\Phi}'\right]\right| + \mathbb{E}_{\tau_2}\left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}', \tau_2\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{\Phi}'}(x_2) = \sigma_2 \mid \boldsymbol{\Phi}'\right]\right|\\ &= o(1). \qquad\qquad \text{[by (6.21)]} \end{aligned}$$

Summing over the four sign combinations of σ_1, σ_2 gives the desired result. □

6.3. **Proof of Proposition 2.5.**

As in the proof of Lemma 6.1 let $c_1, \ldots, c_{\boldsymbol{\Delta}''}$ be the new clauses added by **CPL2** and let $\boldsymbol{x}_{1,1}, \ldots, \boldsymbol{x}_{1,k}, \ldots, \boldsymbol{x}_{\boldsymbol{\Delta}'',1}, \ldots, \boldsymbol{x}_{\boldsymbol{\Delta}'',k}$ be their constituent variables. Let $\mathcal{X} = \{\boldsymbol{x}_{1,1}, \ldots, \boldsymbol{x}_{1,k}, \ldots, \boldsymbol{x}_{\boldsymbol{\Delta}'',1}, \ldots, \boldsymbol{x}_{\boldsymbol{\Delta}'',k}\}$. For $\varepsilon > 0$ and $z \in \mathbb{R}$ define $\lambda_\varepsilon(z) = \log(z \vee \varepsilon)$. Finally, let $(\boldsymbol{s}_i)_{i\ge 0}$ be a sequence of uniformly random ± 1-valued random variables, mutually independent and independent of all other randomness.

Lemma 6.4. *Assume that $d < d_{\mathrm{uniq}}(k)$. There exists $B = B(d,k) > 0$ such that for all $0 < \varepsilon < 1$ we have*

$$\limsup_{n\to\infty}\mathbb{E}\left[\left(\sum_{i=1}^{\boldsymbol{\Delta}''}\lambda_\varepsilon\left(1 - \prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j}) \ne \mathrm{sign}(\boldsymbol{x}_{i,j}, c_i) \mid \boldsymbol{\Phi}'\right]\right)\right)^2 \mid Z(\boldsymbol{\Phi}') > 0\right] \le B.$$

Proof. Given $Z(\boldsymbol{\Phi}') > 0$ we have

$$0 \ge \lambda_\varepsilon\left(1 - \prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{1,j}) \ne \mathrm{sign}(\boldsymbol{x}_{1,j}, c_1) \mid \boldsymbol{\Phi}'\right]\right) \ge \lambda_\varepsilon\left(1 - \mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{1,1}) \ne \mathrm{sign}(\boldsymbol{x}_{1,1}, c_1) \mid \boldsymbol{\Phi}'\right]\right). \tag{6.22}$$

Recalling that $\boldsymbol{\Delta}'' \sim \mathrm{Po}(d(k-1)/k)$, we combine (6.22) with Cauchy-Schwarz to obtain $B' = B'(d,k) > 0$ such that

$$\begin{aligned} &\mathbb{E}\left[\left(\sum_{i=1}^{\boldsymbol{\Delta}''}\lambda_\varepsilon\left(1 - \prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j}) \ne \mathrm{sign}(\boldsymbol{x}_{i,j}, c_i) \mid \boldsymbol{\Phi}'\right]\right)\right)^2 \mid Z(\boldsymbol{\Phi}') > 0\right]\\ &\le B' \cdot \mathbb{E}\left[\lambda_\varepsilon\left(1 - \mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{1,1}) \ne \mathrm{sign}(\boldsymbol{x}_{1,1}, c_1) \mid \boldsymbol{\Phi}'\right]\right)^2 \mid Z(\boldsymbol{\Phi}') > 0\right]. \end{aligned} \tag{6.23}$$

Further, since the function λ_ε is bounded and continuous for every $\varepsilon > 0$ and since $\mathrm{sign}(\boldsymbol{x}_{1,1}, c_1)$ is chosen independently of $\boldsymbol{\Phi}'$, Proposition 2.11 shows that for any $\varepsilon > 0$,

$$\begin{aligned}\mathbb{E}\left[\lambda_\varepsilon\left(1-\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{1,1}) \neq \mathrm{sign}(\boldsymbol{x}_{1,1}, c_1) \mid \boldsymbol{\Phi}'\right]\right)^2 \mid Z(\boldsymbol{\Phi}') > 0\right] &= \mathbb{E}\left[\lambda_\varepsilon\left(\boldsymbol{\mu}_{\pi_{d,k},1,1}\right)^2\right] + o(1) \\ &\leq \mathbb{E}\left[\log\left(\boldsymbol{\mu}_{\pi_{d,k},1,1}\right)^2\right] + o(1). \end{aligned} \tag{6.24}$$

Since Proposition 2.1 shows that $\mathbb{E}\left[\log\left(\boldsymbol{\mu}_{\pi_{d,k},1,1}\right)^2\right] = O(1)$, the assertion follows from (6.23) and (6.24). $\square$

Lemma 6.5. *Assume that $d < d_{\mathrm{uniq}}(k)$. For any $\delta > 0$ there exists $\varepsilon_0 > 0$ such that for all $\varepsilon_0 > \varepsilon > 0$ we have*

$$\limsup_{n\to\infty}\left|\mathbb{E}\left[\log\frac{Z(\boldsymbol{\Phi}'')\vee 1}{Z(\boldsymbol{\Phi}')\vee 1}\right] - \frac{d(k-1)}{k}\mathbb{E}\left[\lambda_\varepsilon\left(1-\prod_{j=1}^k \mathbb{P}\left[\boldsymbol{\sigma}(x_j) = \boldsymbol{s}_j \mid \boldsymbol{\Phi}'\right]\right) \mid Z(\boldsymbol{\Phi}') > 0\right]\right| < \delta.$$

Proof. We choose small enough $\xi = \xi(d,k,\delta) > \zeta(\xi) > \eta = \eta(\zeta) > \varepsilon_0 = \varepsilon_0(\eta) > 0$, let $0 < \varepsilon < \varepsilon_0$ and assume that $n \geq n_0(\varepsilon)$ is large enough. Also let $\gamma = \gamma(n) = o(1)$ be a sequence that tends to zero sufficiently slowly. Additionally, let $\mathfrak{E}$ be the event that all of the following conditions occur.

E1: $Z(\boldsymbol{\Phi}') > 0$.
E2: $\boldsymbol{\Delta}'' \leq \zeta^{-1}$.
E3: $|\mathscr{X}| = k\boldsymbol{\Delta}''$.
E4: $\max_{x\in\mathscr{X}, s\in\{\pm1\}} \mathbb{P}[\boldsymbol{\sigma}(x) = s \mid \boldsymbol{\Phi}'] \leq 1-\eta$.
E5: $\sum_{\tau\in\{\pm1\}^{\mathscr{X}}} \left|\mathbb{P}[\forall x \in \mathscr{X} : \boldsymbol{\sigma}(x) = \tau(x) \mid \boldsymbol{\Phi}'] - \prod_{x\in\mathscr{X}} \mathbb{P}[\boldsymbol{\sigma}(x) = \tau(x) \mid \boldsymbol{\Phi}']\right| < \gamma$.

We claim that

$$\mathbb{P}[\mathfrak{E}] \geq 1 - 2\xi + o(1). \tag{6.25}$$

Indeed, since $0 < d < d_{\mathrm{uniq}}(k) < d_{\mathrm{sat}}(k)$, we have that $\mathbb{P}\left[Z(\boldsymbol{\Phi}') > 0\right] = 1 - o(1)$. Moreover, since $\boldsymbol{\Delta}'' \sim \mathrm{Po}(d(k-1)/k)$, Markov's inequality shows that $\mathbb{P}\left[\boldsymbol{\Delta}'' > \zeta^{-1}\right] \leq \zeta d < \xi$. Further, since the new clauses $c_1, \ldots, c_{\boldsymbol{\Delta}''}$ are chosen independently, we have $\mathbb{P}\left[|\mathscr{X}| = k\boldsymbol{\Delta}'' \mid \boldsymbol{\Delta}'' \leq \zeta^{-1}\right] = 1 - O(1/n)$.

Moreover, per Proposition 2.11 we see that the joint distribution on the assignments over $\mathscr{X}$ must be approximately the product measure. The tails of the limiting distribution of the latter are controlled by (2.1). Therefore, for small enough η we should have

$$\mathbb{P}\left[\max_{x\in\mathscr{X}, s\in\{\pm1\}} \mathbb{P}[\boldsymbol{\sigma}(x) = s \mid \boldsymbol{\Phi}'] \leq 1-\eta \mid \boldsymbol{\Delta}'' \leq \zeta^{-1}, Z(\boldsymbol{\Phi}') > 0\right] \geq 1 - \xi\ .$$

Similarly, Proposition 2.11 shows together with Markov's inequality that

$$\mathbb{P}\left[\textbf{E5}\text{ occurs} \mid \boldsymbol{\Delta}'' \leq \zeta^{-1}, Z(\boldsymbol{\Phi}') > 0\right] = 1 - o(1)\ ,$$

provided that $\gamma \to \infty$ sufficiently slowly. Thus, we obtain (6.25).

Furthermore, (6.25) implies together with Proposition 2.7 and Hölder's inequality that

$$\mathbb{E}\left|(1-\mathbb{1}\mathfrak{E})\cdot\log\frac{Z(\boldsymbol{\Phi}'')}{Z(\boldsymbol{\Phi}')}\right| \leq \delta/3 + o(1), \tag{6.26}$$

provided that $\xi = \xi(d,k,\delta) > 0$ is small enough. Analogously, (6.25), Lemma 6.4 and Cauchy-Schwarz yield

$$\mathbb{E}\left|(1-\mathbb{1}\mathfrak{E})\lambda_\varepsilon\left(1-\prod_{j=1}^k \mathbb{P}\left[\boldsymbol{\sigma}(x_j) = \boldsymbol{s}_j \mid \boldsymbol{\Phi}'\right]\right)\right| \leq \delta/3 + o(1). \tag{6.27}$$

Thus, we confine ourselves to the event $\mathfrak{E}$, on which we have $Z(\boldsymbol{\Phi}'), Z(\boldsymbol{\Phi}'') > 0$ due to **E1**, **E3**, **E4** and **E5**. Hence,

$$\begin{aligned}\log\frac{Z(\boldsymbol{\Phi}'')\vee 1}{Z(\boldsymbol{\Phi}')\vee 1} &= \log\frac{Z(\boldsymbol{\Phi}'')}{Z(\boldsymbol{\Phi}')} = \log\sum_{\tau\in\{\pm1\}^{\mathscr{X}}} \mathbb{1}\{\tau \models c_1, \ldots, c_{\boldsymbol{\Delta}''}\}\mathbb{P}[\forall x \in \mathscr{X} : \boldsymbol{\sigma}(x) = \tau(x) \mid \boldsymbol{\Phi}'] \\ &= \log\sum_{\tau\in\{\pm1\}^{\mathscr{X}}} \mathbb{1}\{\tau \models c_1, \ldots, c_{\boldsymbol{\Delta}''}\}\prod_{x\in\mathscr{X}} \mathbb{P}[\boldsymbol{\sigma}(x) = \tau(x) \mid \boldsymbol{\Phi}'] + o(1) && [\text{by } \textbf{E4, E5}] \\ &= \sum_{i=1}^{\boldsymbol{\Delta}''}\log\left[1-\prod_{j=1}^k \mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j}) \neq \mathrm{sign}(\boldsymbol{x}_{i,j}, c_i) \mid \boldsymbol{\Phi}'\right]\right] + o(1) && [\text{by } \textbf{E3}]. \end{aligned} \tag{6.28}$$

Further, **E4** ensures that for any $1 \le i \le \Delta''$,

$$\left|\log\left[1-\prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j}) \neq \mathrm{sign}(\boldsymbol{x}_{i,j},c_i) \mid \boldsymbol{\Phi}'\right]\right] - \lambda_\varepsilon\left[1-\prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j}) \neq \mathrm{sign}(\boldsymbol{x}_{i,j},c_i) \mid \boldsymbol{\Phi}'\right]\right]\right| < \xi. \tag{6.29}$$

Thus, combining (6.28) and (6.29), we obtain

$$\mathbb{E}\left|\mathbb{1}\mathfrak{E}\left(\log\frac{Z(\boldsymbol{\Phi}'')\vee 1}{Z(\boldsymbol{\Phi}')\vee 1} - \sum_{i=1}^{\Delta''}\lambda_\varepsilon\left(1-\prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j}) \neq \mathrm{sign}(\boldsymbol{x}_{i,j},c_i) \mid \boldsymbol{\Phi}'\right]\right)\right)\right| < \delta/3 + o(1). \tag{6.30}$$

Further, combining (6.26) and (6.30) with Lemma 6.4, we obtain

$$\left|\mathbb{E}\left[\log\frac{Z(\boldsymbol{\Phi}'')\vee 1}{Z(\boldsymbol{\Phi}')\vee 1}\right] - \mathbb{E}\left[\sum_{i=1}^{\Delta''}\lambda_\varepsilon\left(1-\prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j}) \neq \mathrm{sign}(\boldsymbol{x}_{i,j},c_i) \mid \boldsymbol{\Phi}'\right]\right) \mid Z(\boldsymbol{\Phi}')>0\right]\right| < \delta + o(1). \tag{6.31}$$

Finally, since the clauses $c_1,\ldots,c_{\Delta''}$ are drawn uniformly and independently and since the distribution of $\boldsymbol{\Phi}'$ is invariant under permutation of the variable nodes, we find

$$\begin{aligned}&\mathbb{E}\left[\sum_{i=1}^{\Delta''}\lambda_\varepsilon\left(1-\prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j}) \neq \mathrm{sign}(\boldsymbol{x}_{i,j},c_i) \mid \boldsymbol{\Phi}'\right] \mid Z(\boldsymbol{\Phi}')>0\right)\right]\\ &\quad= \frac{d(k-1)}{k}\mathbb{E}\left[\lambda_\varepsilon\left(1-\prod_{j=1}^{k}\mathbb{P}\left[\boldsymbol{\sigma}(x_j)=\boldsymbol{s}_j\right] \mid \boldsymbol{\Phi}'\right) \mid Z(\boldsymbol{\Phi}')>0\right].\end{aligned} \tag{6.32}$$

Combining (6.31) and (6.32) completes the proof. □

Proof of Proposition 2.5. Proposition 2.11 shows together with Lemma 6.5 that

$$\mathbb{E}\left[\log\frac{Z(\boldsymbol{\Phi}'')\vee 1}{Z(\boldsymbol{\Phi}')\vee 1}\right] = \frac{d(k-1)}{k}\mathbb{E}\left[\lambda_\varepsilon\left(1-\prod_{j=1}^{k}\boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right] + o_\varepsilon(1), \tag{6.33}$$

with $o_\varepsilon(1)$ hiding a term that vanishes in the limit $\varepsilon \to 0$. Furthermore, in light of (2.1) the monotone convergence theorem yields

$$\mathbb{E}\left[\log\left(1-\prod_{j=1}^{k}\boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right] = \lim_{\varepsilon\to 0}\mathbb{E}\left[\lambda_\varepsilon\left(1-\prod_{j=1}^{k}\boldsymbol{\mu}_{\pi_{d,k},1,j}\right)\right]. \tag{6.34}$$

The assertion follows from (6.33) and (6.34). □

6.4. **Proof of Proposition 2.6.** We adapt the steps from Section 6.3 to the coupling of $\boldsymbol{\Phi}'$, $\boldsymbol{\Phi}'''$. Recall that the latter is obtained by adding to $\boldsymbol{\Phi}'$ a single variable x_{n+1} along with $\boldsymbol{\Delta}'''$ clauses $b_1,\ldots,b_{\boldsymbol{\Delta}'''}$ that each contain x_{n+1} and $k-1$ other variables. Thus, let $\boldsymbol{x}_{1,1},\ldots,\boldsymbol{x}_{1,k-1},\ldots,\boldsymbol{x}_{\boldsymbol{\Delta}''',1},\ldots,\boldsymbol{x}_{\boldsymbol{\Delta}''',k-1} \in \{x_1,\ldots,x_n\}$ be the variables other than x_{n+1} that appear in $b_1,\ldots,b_{\boldsymbol{\Delta}'''}$ and let $\mathscr{X} = \{\boldsymbol{x}_{1,1},\ldots,\boldsymbol{x}_{\boldsymbol{\Delta}''',k-1}\}$ be the set comprising all these variables.

Lemma 6.6. *Assume that* $0 < d < d_{\mathrm{uniq}}(k)$. *There exists* $B = B(d,k) > 0$ *such that for all* $0 < \varepsilon < 1$ *we have*

$$\limsup_{n\to\infty}\mathbb{E}\left[\lambda_\varepsilon\left(\sum_{s\in\{\pm 1\}}\prod_{i=1}^{\boldsymbol{\Delta}'''}\left(1-\mathbb{1}\{\mathrm{sign}(x_{n+1},b_i)\neq s\}\prod_{j=1}^{k-1}\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\boldsymbol{\Phi}']\right)\right)^2 \mid Z(\boldsymbol{\Phi}')>0\right] \le B.$$

Proof. Given that $\boldsymbol{\Phi}'$ is satisfiable, and noticing that λ_ε is increasing, and $\varepsilon \in (0,1)$, we see that

$$\begin{aligned}0 \wedge \lambda_\varepsilon&\left(\sum_{s\in\{\pm 1\}}\prod_{i=1}^{\boldsymbol{\Delta}'''}\left(1-\mathbb{1}\{\mathrm{sign}(x_{n+1},b_i)\neq s\}\prod_{j=1}^{k-1}\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\boldsymbol{\Phi}']\right)\right)\\
&\ge \lambda_\varepsilon\left(\sum_{s\in\{\pm 1\}}\prod_{i=1}^{\boldsymbol{\Delta}'''}1-\prod_{j=1}^{k-1}\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\boldsymbol{\Phi}']\right)\\
&\ge \lambda_\varepsilon\left(\prod_{i=1}^{\boldsymbol{\Delta}'''}1-\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,1})\neq\mathrm{sign}(\boldsymbol{x}_{i,1},b_i)\mid\boldsymbol{\Phi}']\right)\\
&= \lambda_\varepsilon\left(\prod_{i=1}^{\boldsymbol{\Delta}'''}\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,1})=\mathrm{sign}(\boldsymbol{x}_{i,1},b_i)\mid\boldsymbol{\Phi}']\right) \ge \sum_{i=1}^{\boldsymbol{\Delta}'''}\lambda_\varepsilon(\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,1})=\mathrm{sign}(\boldsymbol{x}_{i,1},b_i)\mid\boldsymbol{\Phi}']).\end{aligned} \tag{6.35}$$

We also notice that

$$0\vee\lambda_\varepsilon\left(\sum_{s\in\{\pm1\}}\prod_{i=1}^{\mathbf{\Delta}'''}\left(1-\mathbb{1}\{\mathrm{sign}(x_{n+1},b_i)\neq s\}\prod_{j=1}^{k-1}\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\mathbf{\Phi}']\right)\right)<1\ . \tag{6.36}$$

In light of the above, we now bound

$$\begin{aligned}
&\mathbb{E}\left[\lambda_\varepsilon\left(\sum_{s\in\{\pm1\}}\prod_{i=1}^{\mathbf{\Delta}'''}\left(1-\mathbb{1}\{\mathrm{sign}(x_{n+1},b_i)\neq s\}\prod_{j=1}^{k-1}\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\mathbf{\Phi}']\right)\right)^2\mid Z(\mathbf{\Phi}')>0\right]\\
&\le\mathbb{E}\left[1+\left(\sum_{i=1}^{\mathbf{\Delta}'''}\lambda_\varepsilon(\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,1})=\mathrm{sign}(\boldsymbol{x}_{i,1},b_i)\mid\mathbf{\Phi}']\right)^2\mid Z(\mathbf{\Phi}')>0\right] && \text{[from (6.35),(6.36)]}\\
&\le d(d+1)\mathbb{E}\left[1+\left(\lambda_\varepsilon(\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{1,1})=\mathrm{sign}(\boldsymbol{x}_{1,1},b_1)\mid\mathbf{\Phi}'])\right)^2\mid Z(\mathbf{\Phi}')>0\right] && [\mathbf{\Delta}'''\sim\mathrm{Po}(d)]\\
&\le d(d+1)\left(1+\mathbb{E}\left[\lambda_\varepsilon(\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{1,1})=\mathrm{sign}(\boldsymbol{x}_{1,1},b_1)\mid\mathbf{\Phi}'])^2\mid Z(\mathbf{\Phi}')>0\right]\right).
\end{aligned}\tag{6.37}$$

Further, Proposition 2.11 implies that for any $\varepsilon>0$,

$$\mathbb{E}\left[\lambda_\varepsilon(\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{1,1})=\mathrm{sign}(\boldsymbol{x}_{1,1},b_1)\mid\mathbf{\Phi}'])^2\mid Z(\mathbf{\Phi}')>0\right]=\mathbb{E}\left[\lambda_\varepsilon(\boldsymbol{\mu}_{\pi_{d,k},1,1})^2\right]+o(1)\le\mathbb{E}\left[\log^2\boldsymbol{\mu}_{\pi_{d,k},1,1}\right]+o(1). \tag{6.38}$$

Finally, the assertion follows from (6.37) and (6.38). $\square$

Lemma 6.7. *Assume that* $0<d<d_{\mathrm{uniq}}(k)$. *For any* $\delta>0$ *there exists* $\varepsilon_0>0$ *such that for all* $\varepsilon_0>\varepsilon>0$ *we have*

$$\begin{aligned}
\limsup_{n\to\infty}&\left|\mathbb{E}\left[\log\frac{Z(\mathbf{\Phi}''')\vee1}{Z(\mathbf{\Phi}')\vee1}\right]\right.\\
&\left.-\mathbb{E}\left[\lambda_\varepsilon\left(\sum_{s\in\{\pm1\}}\left(\prod_{i=1}^{\mathbf{\Delta}'''}1-\mathbb{1}\{\mathrm{sign}(x_{n+1},b_i)\neq s\}\prod_{j=1}^{k-1}\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\mathbf{\Phi}']\right)\right)^2\mid Z(\mathbf{\Phi}')>0\right]\right|<\delta.
\end{aligned}$$

Proof. Choose small enough $\xi=\xi(d,k,\delta)>\zeta(\xi)>\eta=\eta(\zeta)>\varepsilon_0=\varepsilon_0(\eta)>0$, let $0<\varepsilon<\varepsilon_0$, suppose that $n>n_0(\varepsilon)$ is sufficiently large and let $0<\gamma=\gamma(n)=o(1)$ be a sequence that converges to zero slowly. Let $\mathfrak{E}$ be the event that the following conditions occur.

E1: $Z(\mathbf{\Phi}')>0$.
E2: $\mathbf{\Delta}'''\le\zeta^{-1}$.
E3: $|\mathscr{X}|=(k-1)\mathbf{\Delta}'''$.
E4: $\max_{x\in\mathscr{X},s\in\{\pm1\}}\mathbb{P}[\boldsymbol{\sigma}(x)=s\mid\mathbf{\Phi}']\le1-\eta$.
E5: $\sum_{\tau\in\{\pm1\}^{\mathscr{X}}}\left|\mathbb{P}[\forall x\in\mathscr{X}:\boldsymbol{\sigma}(x)=\tau(x)\mid\mathbf{\Phi}']-\prod_{x\in\mathscr{X}}\mathbb{P}[\boldsymbol{\sigma}(x)=\tau(x)\mid\mathbf{\Phi}']\right|<\gamma$.

As in the proof of Lemma 6.5 we find that

$$\mathbb{P}[\mathfrak{E}]\ge1-2\xi+o(1). \tag{6.39}$$

Let

$$\boldsymbol{L}_\varepsilon=\lambda_\varepsilon\left(\sum_{s\in\{\pm1\}}\prod_{i=1}^{\mathbf{\Delta}'''}\left(1-\mathbb{1}\{\mathrm{sign}(x_{n+1},b_i)\neq s\}\prod_{j=1}^{k-1}\mathbb{P}[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\mathbf{\Phi}']\right)\right)$$

for brevity. Combining Proposition 2.7, Lemma 6.6 and (6.39) and using Hölder's inequality, we obtain

$$\mathbb{E}\left|(1-\mathbb{1}\mathfrak{E})\log\frac{Z(\mathbf{\Phi}'')}{Z(\mathbf{\Phi}')}\right|\le\delta/3+o(1),\qquad\qquad\mathbb{E}\left|(1-\mathbb{1}\mathfrak{E})\boldsymbol{L}_\varepsilon\mid Z(\mathbf{\Phi}')>0\right|\le\delta/3+o(1). \tag{6.40}$$

Hence, we are left to compare $\mathbb{E}\left|\mathbb{1}\mathfrak{E}\cdot\log\frac{Z(\mathbf{\Phi}'')}{Z(\mathbf{\Phi}')}\right|$ and $\mathbb{E}\left|\mathbb{1}\mathfrak{E}\cdot\boldsymbol{L}_\varepsilon \mid Z(\mathbf{\Phi}')>0\right|$. On the event $\mathfrak{E}$ we have $Z(\mathbf{\Phi}'), Z(\mathbf{\Phi}''')>0$. Consequently,

$$\begin{aligned}\log\frac{Z(\mathbf{\Phi}''')\vee 1}{Z(\mathbf{\Phi}')\vee 1} &= \log\frac{Z(\mathbf{\Phi}''')}{Z(\mathbf{\Phi}')} = \log\sum_{\tau\in\{\pm1\}^{\mathscr{X}\cup\{x_{n+1}\}}} \mathbb{1}\{\tau\models b_1,\dots,b_{\Delta'''}\}\,\mathbb{P}[\forall x\in\mathscr{X}:\boldsymbol{\sigma}(x)=\tau(x)\mid \mathbf{\Phi}'] \\ &= \log\sum_{\tau\in\{\pm1\}^{\mathscr{X}\cup\{x_{n+1}\}}} \mathbb{1}\{\tau\models b_1,\dots,b_{\Delta'''}\}\prod_{x\in\mathscr{X}}\mathbb{P}[\boldsymbol{\sigma}(x)=\tau(x)\mid\mathbf{\Phi}']+o(1) && \text{[by \textbf{E4, E5}]} \\ &= \log\left[\sum_{s\in\{\pm1\}}\prod_{i=1}^{\Delta'''}\left(1-\mathbb{1}\{\mathrm{sign}(x_{n+1},b_i)\neq s\}\prod_{j=1}^{k-1}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\mathbf{\Phi}'\right]\right)\right]+o(1) && \text{[by \textbf{E3}]}. \end{aligned} \tag{6.41}$$

Now, **E4** guarantees that

$$\log\left[\sum_{s\in\{\pm1\}}\prod_{i=1}^{\Delta'''}\left(1-\mathbb{1}\{\mathrm{sign}(x_{n+1},b_i)\neq s\}\prod_{j=1}^{k-1}\mathbb{P}\left[\boldsymbol{\sigma}(\boldsymbol{x}_{i,j})\neq\mathrm{sign}(\boldsymbol{x}_{i,j},b_i)\mid\mathbf{\Phi}'\right]\right)\right]=\boldsymbol{L}_\varepsilon. \tag{6.42}$$

Therefore, we combine (6.41) and (6.42) to obtain

$$\mathbb{E}\left|\mathbb{1}\mathfrak{E}\left(\log\frac{Z(\mathbf{\Phi}''')\vee 1}{Z(\mathbf{\Phi}')\vee 1}-\boldsymbol{L}_\varepsilon\right)\right|<\delta/3+o(1). \tag{6.43}$$

Finally, the assertion follows from (6.40) and (6.43). □

Proof of Proposition 2.6. Following similar steps as in the proof of Proposition 2.5, we see that the assertion follows from Lemma 6.7, Proposition 2.1, Proposition 2.11, and the dominated convergence theorem. □

Proof of Proposition 2.3. Immediate from Fact 2.4, Proposition 2.5 and Proposition 2.6. □

7. Proof of Proposition 2.15

7.1. **Proof of Lemma 2.12.** The proof is by induction on the height of the tree. The following claim summarises the main step of the induction.

Claim 7.1. *For all $\ell\geq 0$, all variables x of $\mathbb{T}^{(\ell)}$ and all satisfying assignments $\tau\in S(\mathbb{T}^{(\ell)})$ we have*

$$\frac{Z(\mathbb{T}_x^{(\ell)},\tau,\boldsymbol{\tau}^+(x))}{Z(\mathbb{T}_x^{(\ell)},\tau)}\leq\frac{Z(\mathbb{T}_x^{(\ell)},\boldsymbol{\tau}^+,\boldsymbol{\tau}^+(x))}{Z(\mathbb{T}_x^{(\ell)},\boldsymbol{\tau}^+)}. \tag{7.1}$$

Proof. For boundary variables $x\in\partial^{2\ell}\mathfrak{x}$ there is nothing to show because the r.h.s. of (7.1) equals one. Hence, consider a variable $x\in\partial^{2q}\mathfrak{x}$ for some $q<\ell$. If $Z(\mathbb{T}_x^{(\ell)},\tau,\boldsymbol{\tau}^+(x))=0$, then (7.1) is trivially satisfied. Hence, assume that $Z(\mathbb{T}_x^{(\ell)},\tau,\boldsymbol{\tau}^+(x))>0$.

Let $a_1^+,\dots,a_g^+$ be the children (clauses) of x with $\mathrm{sign}(x,a_i^+)=\boldsymbol{\tau}^+(x)$. Also let $y_{11},\dots,y_{1(k-1)},\dots,y_{g1},\dots,y_{g(k-1)}$ be the children (variables) of $a_1^+,\dots,a_g^+$. Similarly, let $a_1^-,\dots,a_h^-$ be the children of x with $\mathrm{sign}(x,a_i^-)=-\boldsymbol{\tau}^+(x)$ and let $z_{11},\dots,z_{1(k-1)},\dots,z_{h1},\dots,z_{h(k-1)}$ be their children. We claim that for all $\tau\in S(\mathbb{T}^{(\ell)})$,

$$Z(\mathbb{T}_x^{(\ell)},\tau,\boldsymbol{\tau}^+(x))=\left(\prod_{i=1}^{g}\prod_{t=1}^{k-1}Z(\mathbb{T}_{y_{it}}^{(\ell)},\tau)\right)\cdot\prod_{j=1}^{h}\left(\prod_{t=1}^{k-1}Z(\mathbb{T}_{z_{jt}}^{(\ell)},\tau)-\prod_{t=1}^{k-1}Z(\mathbb{T}_{z_{jt}}^{(\ell)},\tau,-\boldsymbol{\tau}^+(z_{jt}))\right), \tag{7.2}$$

$$Z(\mathbb{T}_x^{(\ell)},\tau,-\boldsymbol{\tau}^+(x))=\prod_{i=1}^{g}\left(\prod_{t=1}^{k-1}Z(\mathbb{T}_{y_{it}}^{(\ell)},\tau)-\prod_{t=1}^{k-1}Z(\mathbb{T}_{y_{it}}^{(\ell)},\tau,\boldsymbol{\tau}^+(y_{it}))\right)\cdot\left(\prod_{j=1}^{h}\prod_{t=1}^{k-1}Z(\mathbb{T}_{z_{jt}}^{(\ell)},\tau)\right). \tag{7.3}$$

For setting x to $\boldsymbol{\tau}^+(x)$ satisfies $a_1^+,\dots,a_g^+$; hence, arbitrary satisfying assignments of the sub-trees $\mathbb{T}_{y_{it}}^{(\ell)}$ can be combined, which explains the first product in (7.2). By contrast, upon assigning x the value $\boldsymbol{\tau}^+(x)$ we need to ensure that each of the clauses $a_1^-,\dots,a_g^-$ are satisfied by at least one variable other than x. This explains the second factor of (7.2). A similar argument yields (7.3). Dividing (7.3) by (7.2) and invoking the induction hypothesis (for $q+1$), we obtain

$$\frac{Z(\mathbb{T}_x^{(\ell)},\tau,-\boldsymbol{\tau}^+(x))}{Z(\mathbb{T}_x^{(\ell)},\tau,\boldsymbol{\tau}^+(x))}=\prod_{i=1}^{g}\left(1-\prod_{t=1}^{k-1}\frac{Z(\mathbb{T}_{y_{it}}^{(\ell)},\tau,\boldsymbol{\tau}^+(y_{it}))}{Z(\mathbb{T}_{y_{it}}^{(\ell)},\tau)}\right)\cdot\prod_{j=1}^{h}\left(1-\prod_{t=1}^{k-1}\frac{Z(\mathbb{T}_{z_{jt}}^{(\ell)},\tau,-\boldsymbol{\tau}^+(z_i))}{Z(\mathbb{T}_{z_{jt}}^{(\ell)},\tau)}\right)^{-1}$$

$$\geq\prod_{i=1}^{g}\left(1-\prod_{t=1}^{k-1}\frac{Z(\mathbb{T}_{y_{it}}^{(\ell)},\boldsymbol{\tau}^+,\boldsymbol{\tau}^+(y_{it}))}{Z(\mathbb{T}_{y_{it}}^{(\ell)},\boldsymbol{\tau}^+)}\right)\cdot\prod_{j=1}^{h}\left(1-\prod_{t=1}^{k-1}\frac{Z(\mathbb{T}_{z_{jt}}^{(\ell)},\boldsymbol{\tau}^+,-\boldsymbol{\tau}^+(z_i))}{Z(\mathbb{T}_{z_{jt}}^{(\ell)},\boldsymbol{\tau}^+)}\right)^{-1}=\frac{Z(\mathbb{T}_x^{(\ell)},\boldsymbol{\tau}^+,-\boldsymbol{\tau}^+(x))}{Z(\mathbb{T}_x^{(\ell)},\boldsymbol{\tau}^+,\boldsymbol{\tau}^+(x))}\ ,$$

completing the induction. □

Proof of Lemma 2.12. Applying Claim 7.1 to $x=\mathfrak{x}$ completes the proof of Lemma 2.12. □

7.2. **Proof of Lemma 2.14.** We employ the PULP algorithm introduced in Section 2.3 and its analysis on the random tree from Section 3.2. Recall that given an initial set of literals $\mathscr{L}$, PULP returns a superset $\bar{\mathscr{L}}$ with the property that the partial assignment obtained from setting all literals of $\bar{\mathscr{L}}$ to true, leaves no clause with only unsatisfying literals. Let us write $\bar{\mathscr{L}}=\bar{\mathscr{L}}_{x,s}$ for the set returned by PULP algorithm, initialized with the literal set $\mathscr{L}=\{s\cdot x\}$.

Claim 7.2. *Let $0\leq t<\ell$ and assume that $x\in\partial_{\mathbb{T}}^{2t}\mathfrak{x}$, $s\in\{\pm1\}$, satisfy $|\bar{\mathscr{L}}_{x,s}|<\ell-t$. Then for all $\tau\in S(\mathbb{T}^{(\ell)})$*

$$Z(\mathbb{T}_x^{(\ell)},\tau)\leq 2^{|\bar{\mathscr{L}}_{x,s}|}\cdot Z(\mathbb{T}_x^{(\ell)},\tau,s)\ . \tag{7.4}$$

Proof. Notice that under our assumption on the size of $\bar{\mathscr{L}}_{x,s}$, the assignment τ does not clash with the one imposed by PULP. The assertion therefore follows immediately from the same argument as in the proof of Lemma 2.8. □

Claim 7.3. *We have* $\lim_{t\to\infty}\mathbb{P}\left[|\partial_{\mathbb{T}}^{2t}\mathfrak{x}|>(200d\cdot(k-1))^t\right]=0$.

Proof. This is an immediate consequence of Lemma 3.2. □

Proof of Lemma 2.14. Assume that $\ell>ct^c$ for a large enough $c=c(d,k)>0$ and that $t>t_0=t_0(d,k)$ is sufficiently large. Then Corollary 3.3 shows that

$$\mathbb{P}\left[|\bar{\mathscr{L}}_{\mathfrak{x},\pm1}|\geq t^c\right]\leq\exp(-t^2)\ . \tag{7.5}$$

Combining Claim 7.3 with (7.5) and using the union bound, we obtain a sequence $\varepsilon_t\to0$ such that

$$\mathbb{P}\left[\forall x\in\partial_{\mathbb{T}}^{2t}\mathfrak{x}:|\bar{\mathscr{L}}_{x,\pm1}|<t^c\right]\geq1-\varepsilon_t\ . \tag{7.6}$$

If $x\in\partial_{\mathbb{T}}^{2t}\mathfrak{x}$ satisfies $|\bar{\mathscr{L}}_{x,\pm1}|<t^c$ and $\ell>ct^c$, then Claim 7.2 yields that for all $x\in\partial^{2t}\mathfrak{x}$

$$\left|\boldsymbol{\eta}_x^{(\ell)}\right|\leq\log\frac{Z(\mathbb{T}_x^{(\ell)},\boldsymbol{\sigma}^+)}{Z(\mathbb{T}_x^{(\ell)},\boldsymbol{\sigma}^+,+1)}+\log\frac{Z(\mathbb{T}_x^{(\ell)},\boldsymbol{\sigma}^+)}{Z(\mathbb{T}_x^{(\ell)},\boldsymbol{\sigma}^+,-1)}\leq|\bar{\mathscr{L}}_{x,+1}|+|\bar{\mathscr{L}}_{x,-1}|\leq2t^c\ . \tag{7.7}$$

The result now follows from (7.6) and (7.7). □

7.3. **Proof of Proposition 2.15.** We focus on the operator $\mathrm{LL}_{d,k}^{\star}$ introduced in Section 2.5. Let $\rho=(\rho_\bullet,\rho_\oplus,\rho_\ominus)$, and $\rho'=(\rho'_\bullet,\rho'_\oplus,\rho'_\ominus)$ be two arbitrary triplets in $\mathscr{P}(-\infty,\infty]\times\mathscr{P}(0,+\infty]\times\mathscr{P}(-\infty,0]$, and write $\hat\rho=(\hat\rho_\bullet\hat\rho_\oplus,\hat\rho_\ominus)$ and $\hat\rho'=(\hat\rho'_\bullet\hat\rho'_\oplus,\hat\rho'_\ominus)$ for the images $\mathrm{LL}_{d,k}^{\star}(\rho)$ and $\mathrm{LL}_{d,k}^{\star}(\rho')$, respectively. We wish to bound $\mathrm{dist}_d(\hat\rho,\hat\rho')$ in terms of $\mathrm{dist}_d(\rho,\rho')$.

To this end, we begin with bounding the W_1-distance separately for each of the coordinates $(\hat\rho_\oplus,\hat\rho'_\oplus),(\hat\rho_\ominus,\hat\rho'_\ominus)$ and $(\hat\rho_\bullet,\hat\rho'_\bullet)$. Observe that it is sufficient to consider only $W_1(\hat\rho_\oplus,\hat\rho'_\oplus)$ and $W_1(\hat\rho_\ominus,\hat\rho'_\ominus)$, as the triangle inequality implies that $W_1(\hat\rho_\bullet,\hat\rho'_\bullet)\leq W_1(\hat\rho_\oplus,\hat\rho'_\oplus)+W_1(\hat\rho_\ominus,\hat\rho'_\ominus)$.

To spell out our bounds, we need to introduce some additional notation. Recall that for $i,j\geq1$ the random variables $\boldsymbol{\eta}_{\bullet,i,j}$, $\boldsymbol{\eta}_{\oplus,i,j}$, $\boldsymbol{\eta}_{\ominus,i,j}$ follow the law of $\rho_\bullet$, $\rho_\oplus$, $\rho_\ominus$, respectively. Similarly, let $\boldsymbol{\eta}'_{\bullet,i,j}$, $\boldsymbol{\eta}'_{\oplus,i,j}$, $\boldsymbol{\eta}'_{\ominus,i,j}$ be random variables with law $\rho'_\bullet,\rho'_\oplus$ and $\rho'_\ominus$, respectively. We denote with $\boldsymbol{\eta}^{\wedge}_{\bullet,i,j}$ the random variable $\boldsymbol{\eta}_{\bullet,i,j}\wedge\boldsymbol{\eta}'_{\bullet,i,j}$, and with $\boldsymbol{\eta}^{\vee}_{\bullet,i,j}$ the random variable $\boldsymbol{\eta}_{\bullet,i,j}\vee\boldsymbol{\eta}'_{\bullet,i,j}$. Similarly, we write $\boldsymbol{\eta}^{\wedge}_{\oplus,i,j}=\boldsymbol{\eta}_{\oplus,i,j}\wedge\boldsymbol{\eta}'_{\oplus,i,j}$ and $\boldsymbol{\eta}^{\vee}_{\oplus,i,j}=\boldsymbol{\eta}_{\oplus,i,j}\vee\boldsymbol{\eta}'_{\oplus,i,j}$, and also write $\boldsymbol{\eta}^{\wedge}_{\ominus,i,j}=\boldsymbol{\eta}_{\ominus,i,j}\wedge\boldsymbol{\eta}'_{\ominus,i,j}$ and $\boldsymbol{\eta}^{\vee}_{\ominus,i,j}=\boldsymbol{\eta}_{\ominus,i,j},\boldsymbol{\eta}'_{\ominus,i,j}$.

Moreover, for a sign $\varepsilon \in \{\pm 1\}$ and a vector $r = (r_\bullet, r_\oplus, r_\ominus, r_\circ)$ of non-negative integers with $r_\bullet + r_\oplus + r_\ominus + r_\circ = k-1$ and $1 \le i \le r_\bullet$, $1 \le j \le r_\oplus$, $1 \le \ell \le r_\ominus$, we let

$$\mathscr{D}_i^\bullet(z,r;\varepsilon) = \left|\frac{\partial}{\partial z}\log\left(1-\frac{1}{2^{r_\circ}}\Gamma\left(\varepsilon(\boldsymbol{\eta}_{\bullet,1,1},\dots,\boldsymbol{\eta}_{\bullet,1,i-1},z,\boldsymbol{\eta}'_{\bullet,1,i+1},\dots\boldsymbol{\eta}'_{\bullet,1,r_\bullet})\right)\Gamma\left(\varepsilon(\boldsymbol{\eta}'_{\oplus,1,1},\dots,\boldsymbol{\eta}'_{\oplus,1,r_\oplus})\right)\Gamma\left(\varepsilon(\boldsymbol{\eta}'_{\ominus,1,1},\dots,\boldsymbol{\eta}'_{\ominus,1,r_\ominus})\right)\right)\right|.$$

Analogously, we define

$$\mathscr{D}_j^\oplus(z,r;\varepsilon) = \left|\frac{\partial}{\partial z}\log\left(1-\frac{1}{2^{r_\circ}}\Gamma\left(\varepsilon(\boldsymbol{\eta}_{\bullet,1,1},\dots\boldsymbol{\eta}_{\bullet,1,r_\bullet})\right)\Gamma\left(\varepsilon(\boldsymbol{\eta}_{\oplus,1,1},\dots,\boldsymbol{\eta}_{\oplus,1,j-1},z,\boldsymbol{\eta}'_{\oplus,1,j+1},\dots,\boldsymbol{\eta}'_{\oplus,1,r_\oplus})\right)\Gamma\left(\varepsilon(\boldsymbol{\eta}'_{\ominus,1,1},\dots\boldsymbol{\eta}'_{\ominus,1,r_\ominus})\right)\right)\right|,$$

$$\mathscr{D}_\ell^\ominus(z,r;\varepsilon) = \left|\frac{\partial}{\partial z}\log\left(1-\frac{1}{2^{r_\circ}}\Gamma\left(\varepsilon(\boldsymbol{\eta}_{\bullet,1,1},\dots\boldsymbol{\eta}_{\bullet,1,r_\bullet})\right)\Gamma\left(\varepsilon(\boldsymbol{\eta}_{\oplus,1,1},\dots,\boldsymbol{\eta}_{\oplus,1,r_\oplus})\right)\Gamma\left(\varepsilon(\boldsymbol{\eta}_{\ominus,1,1},\dots,\boldsymbol{\eta}_{\ominus,1,\ell-1},z,\boldsymbol{\eta}'_{\ominus,1,\ell+1},\dots,\boldsymbol{\eta}'_{\oplus,1,r_\ominus})\right)\right)\right|.$$

With the above notation in place, we are now ready to bound $W_1(\hat\rho_\oplus,\hat\rho'_\oplus)$. For each of the pairs of distributions $(\rho_\bullet,\rho'_\bullet)$, $(\rho_\oplus,\rho'_\oplus)$, and $(\rho_\ominus,\rho'_\ominus)$, fix an arbitrary coupling among its coordinates.

Lemma 7.4. *$W_1(\hat\rho_\oplus,\hat\rho'_\oplus)$ is upper bounded by*

$$\frac{d/2}{1-e^{-\frac{d}{2}}}\cdot\mathbb{E}\left[\sum_{i=1}^{\boldsymbol{r}_{\bullet,1}}\int_{\boldsymbol{\eta}^\wedge_{\bullet,1,i}}^{\boldsymbol{\eta}^\vee_{\bullet,1,i}}\mathscr{D}_i^\bullet(w_i,\boldsymbol{r}_1;+1)\mathrm{d}w_i+\sum_{j=1}^{\boldsymbol{r}_{\oplus,1}}\int_{\boldsymbol{\eta}^\wedge_{\oplus,1,j}}^{\boldsymbol{\eta}^\vee_{\oplus,1,j}}\mathscr{D}_j^\oplus(y_j,\boldsymbol{r}_1;+1)\mathrm{d}y_j+\sum_{\ell=1}^{\boldsymbol{r}_{\ominus,1}}\int_{\boldsymbol{\eta}^\wedge_{\ominus,1,\ell}}^{\boldsymbol{\eta}^\vee_{\ominus,1,\ell}}\mathscr{D}_\ell^\ominus(z_\ell,\boldsymbol{r}_1;+1)\mathrm{d}z_\ell\right]. \tag{7.8}$$

Proof. Let us write $\Xi'_{i,j}(\varepsilon,r)$ for the expression in the r.h.s. of (2.19) where distribution ρ' is used instead of ρ, i..,

$$\Xi'_{i,j}(\varepsilon,r) = 1-\frac{1}{2^{r_\circ}}\Gamma\left(\varepsilon(\boldsymbol{\eta}'_{\bullet,4i+j,1},\dots,\boldsymbol{\eta}'_{\bullet,4i+j,r_\bullet})\right)\Gamma\left(\varepsilon(\boldsymbol{\eta}'_{\oplus,4i+j,1},\dots,\boldsymbol{\eta}'_{\oplus,4i+j,r_\oplus})\right)\Gamma\left(\varepsilon(\boldsymbol{\eta}'_{\ominus,4i+j,1},\dots,\boldsymbol{\eta}'_{\ominus,4i+j,r_\ominus})\right). \tag{7.9}$$

By identically coupling the number of clauses and the types of the children variables of each clause in $\hat\rho_\oplus,\hat\rho'_\oplus$, we see that by the definition of the W_1 norm,

$$W_1(\hat\rho_\oplus,\hat\rho'_\oplus)\le\mathbb{E}\left[\left|-\sum_{i=1}^{\boldsymbol{d}^\star_+}\log\frac{\Xi_{i,3}(+1,\boldsymbol{r}_{4i+3})}{\Xi'_{i,3}(+1,\boldsymbol{r}_{4i+3})}\right|\right].$$

Applying Wald's lemma, we further obtain

$$W_1(\hat\rho_\oplus,\hat\rho'_\oplus)\le\frac{d/2}{1-e^{-d/2}}\cdot\mathbb{E}\left[\left|\log\frac{\Xi_{1,3}(+1,\boldsymbol{r}_7)}{\Xi'_{1,3}(+1,\boldsymbol{r}_7)}\right|\right]=\frac{d/2}{1-e^{-d/2}}\cdot\mathbb{E}\left[\left|\log\frac{\Xi_{0,1}(+1,\boldsymbol{r}_1)}{\Xi'_{0,1}(+1,\boldsymbol{r}_1)}\right|\right]. \tag{7.10}$$

Let us now focus on the expectation in the r.h.s. of (7.10). Recalling the definition of Ξ in (2.19), and the definition of Ξ' in (7.9), we expand

$$\log\frac{\Xi_{0,1}(+1,\boldsymbol{r}_1)}{\Xi'_{0,1}(+1,\boldsymbol{r}_1)}=\log\frac{1-2^{-r_\circ}\cdot\Gamma\left(\boldsymbol{\eta}_{\bullet,1,1},\dots,\boldsymbol{\eta}_{\bullet,1,\boldsymbol{r}_{\bullet,1}}\right)\cdot\Gamma\left(\boldsymbol{\eta}_{\oplus,1,1},\dots,\boldsymbol{\eta}_{\oplus,1,\boldsymbol{r}_{\oplus,1}}\right)\cdot\Gamma\left(\boldsymbol{\eta}_{\ominus,1,1},\dots,\boldsymbol{\eta}_{\ominus,1,\boldsymbol{r}_{\ominus,1}}\right)}{1-2^{-r_\circ}\cdot\Gamma\left(\boldsymbol{\eta}'_{\bullet,1,1},\dots,\boldsymbol{\eta}'_{\bullet,1,\boldsymbol{r}_{\bullet,1}}\right)\cdot\Gamma\left(\boldsymbol{\eta}'_{\oplus,1,1},\dots,\boldsymbol{\eta}'_{\oplus,1,\boldsymbol{r}_{\oplus,1}}\right)\cdot\Gamma\left(\boldsymbol{\eta}'_{\ominus,1,1},\dots,\boldsymbol{\eta}'_{\ominus,1,\boldsymbol{r}_{\ominus,1}}\right)}. \tag{7.11}$$

Telescoping over the arguments of the functions Γ in the r.h.s of (7.11), invoking the fundamental theorem of calculus for each term, and applying the triangle inequality we further obtain

$$\left|\log\frac{\Xi_{0,1}(+1,\boldsymbol{r}_1)}{\Xi'_{0,1}(+1,\boldsymbol{r}_1)}\right|\le\sum_{i=1}^{\boldsymbol{r}_{\bullet,1}}\left|\int_{\boldsymbol{\eta}'_{\bullet,1,i}}^{\boldsymbol{\eta}_{\bullet,1,i}}\mathscr{D}_i^\bullet(w_i,\boldsymbol{r}_1;+1)\mathrm{d}w_i\right|+\sum_{j=1}^{\boldsymbol{r}_{\oplus,1}}\left|\int_{\boldsymbol{\eta}'_{\oplus,1,j}}^{\boldsymbol{\eta}_{\oplus,1,j}}\mathscr{D}_j^\oplus(y_j,\boldsymbol{r}_1;+1)\mathrm{d}y_j\right|+\sum_{\ell=1}^{\boldsymbol{r}_{\ominus,1}}\left|\int_{\boldsymbol{\eta}'_{\ominus,1,\ell}}^{\boldsymbol{\eta}_{\ominus,1,\ell}}\mathscr{D}_\ell^\ominus(z_\ell,\boldsymbol{r}_1;+1)\mathrm{d}z_\ell\right|.$$

Plugging the above into (7.10) gives the result. □

Following the same steps as above, but replacing '+1' with '−1', yields the corresponding bound for $W_1(\hat\rho_\ominus,\hat\rho'_\ominus)$.

Lemma 7.5. *$W_1(\hat\rho_\ominus,\hat\rho'_\ominus)$ is upper bounded by*

$$\frac{d/2}{1-e^{-\frac{d}{2}}}\cdot\mathbb{E}\left[\sum_{i=1}^{\boldsymbol{r}_{\bullet,1}}\int_{\boldsymbol{\eta}^\wedge_{\bullet,1,i}}^{\boldsymbol{\eta}^\vee_{\bullet,1,i}}\mathscr{D}_i^\bullet(w_i,\boldsymbol{r}_1;-1)\mathrm{d}w_i+\sum_{j=1}^{\boldsymbol{r}_{\oplus,1}}\int_{\boldsymbol{\eta}^\wedge_{\oplus,1,j}}^{\boldsymbol{\eta}^\vee_{\oplus,1,j}}\mathscr{D}_j^\oplus(y_j,\boldsymbol{r}_1;-1)\mathrm{d}y_j+\sum_{\ell=1}^{\boldsymbol{r}_{\ominus,1}}\int_{\boldsymbol{\eta}^\wedge_{\ominus,1,\ell}}^{\boldsymbol{\eta}^\vee_{\ominus,1,\ell}}\mathscr{D}_\ell^\ominus(z_\ell,\boldsymbol{r}_1;-1)\mathrm{d}z_\ell\right]. \tag{7.12}$$

Exploiting the signs of the variables with types $\oplus$ and $\ominus$, we obtain the following bounds for each of the $\mathscr{D}$-functions. For $\lambda \in (0,1]$, we define the real function $\psi_\lambda : [0,1] \to \mathbb{R}$ as

$$\psi_\lambda(w) = \frac{\lambda \cdot w}{1-\lambda \cdot w} \cdot (1-w) \ . \tag{7.13}$$

It is easy to check that $\psi_{\lambda'}(w) \leq \psi_\lambda(w)$, for every $\lambda' \leq \lambda$.

Claim 7.6. *For every $r = (r_\bullet, r_\oplus, r_\ominus, r_\bigcirc)$, and $i \in [r_\bullet]$ we have that*

$$\mathscr{D}_i^\bullet(w_i, r; +1) \leq \psi_{2^{-r_\bigcirc - r_\ominus}}\left(\frac{1+\tanh(w_i/2)}{2}\right) \qquad \text{and} \qquad \mathscr{D}_i^\bullet(w_i, r; -1) \leq \psi_{2^{-r_\bigcirc - r_\oplus}}\left(\frac{1-\tanh(w_i/2)}{2}\right) \ . \tag{7.14}$$

Similarly, we also have that for $j \in [r_\oplus]$,

$$\mathscr{D}_j^\oplus(y_j, r; +1) \leq \psi_{2^{-r_\bigcirc - r_\ominus}}\left(\frac{1+\tanh(y_j/2)}{2}\right) \qquad \text{and} \qquad \mathscr{D}_j^\oplus(y_j, r; -1) \leq \psi_{2^{-r_\bigcirc - (r_\oplus - 1)}}\left(\frac{1-\tanh(y_j/2)}{2}\right) \ , \tag{7.15}$$

and for $\ell \in [r_\ominus]$

$$\mathscr{D}_\ell^\ominus(z_\ell, r; +1) \leq \psi_{2^{-r_\bigcirc - (r_\ominus - 1)}}\left(\frac{1+\tanh(z_\ell/2)}{2}\right) \qquad \text{and} \qquad \mathscr{D}_\ell^\ominus(z_\ell, r; -1) \leq \psi_{2^{-r_\bigcirc - r_\oplus}}\left(\frac{1-\tanh(z_\ell/2)}{2}\right) \ . \tag{7.16}$$

Proof. We only prove the first inequality of (7.14) as the rest of them follow in a similar manner. A straightforward calculation shows that for $z \in \mathbb{R}^q$, $\varepsilon \in \{\pm 1\}$, and $i \in [q]$ we have

$$\frac{\partial}{\partial z_i}\Gamma(\varepsilon \cdot z) = \varepsilon \cdot \frac{1-\tanh(\varepsilon \cdot z_i/2)}{2} \cdot \Gamma(\varepsilon \cdot z) \ . \tag{7.17}$$

Writing $K = 2^{-r_\bigcirc}\Gamma\left(\boldsymbol{\eta}_{\bullet,1,1}, \ldots, \boldsymbol{\eta}_{\bullet,1,i-1}, \boldsymbol{\eta}'_{\bullet,1,i+1}, \ldots \boldsymbol{\eta}'_{\bullet,1,r_\bullet}\right)\Gamma\left(\boldsymbol{\eta}'_{\oplus,1,1}, \ldots, \boldsymbol{\eta}'_{\oplus,1,r_\oplus}\right)\Gamma\left(\boldsymbol{\eta}'_{\ominus,1,1}, \ldots, \boldsymbol{\eta}'_{\ominus,1,r_\ominus}\right)$, applying the chain rule, and using (7.17), we see that

$$\mathscr{D}_i^\bullet(w_i, r; +1) = \psi_K\left(\frac{1+\tanh(w_i/2)}{2}\right) \ . \tag{7.18}$$

Using the fact that $\rho'_\ominus$ is supported in $(-\infty, 0]$, and that $\Gamma \leq 1$, we obtain $K \leq 2^{-r_\bigcirc - r_\ominus}$. The monotonicity of ψ_λ with respect to the parameter λ concludes the proof. ☐

Using Claim 7.6, and maximising each of the functions ψ_λ appearing in (7.14)–(7.16), we can recover the bounds of [50]. To obtain sharper bounds, a natural idea is to optimise groups of summands, instead of optimising each $\mathscr{D}$-summand of $W_1(\hat{\rho}_\oplus, \hat{\rho}'_\oplus) + W_1(\hat{\rho}_\ominus, \hat{\rho}'_\ominus)$ in isolation. In particular, it is tempting to pair terms of the form $\mathscr{D}(\cdot, -1)$ with corresponding terms of the form $\mathscr{D}(\cdot, +1)$, as Lemma 7.7 suggests.

Lemma 7.7. *Let $\phi_\lambda : [0,1] \to \mathbb{R}$ to be the function $\phi_\lambda(w) = \psi_\lambda(w) + \psi_\lambda(1-w)$. For every $\lambda \in (0,1]$, we have that $\phi_\lambda(w) \leq \phi_\lambda(1/2) = \frac{\lambda/2}{1-\lambda/2}$, for all $w \in [0,1]$.*

Proof. For $\lambda = 1$, we have that $\psi_\lambda(w) = w$ implying $\phi_\lambda(w) = 1$, and thus, the result holds trivially. Let now $\lambda \in (0,1)$. Differentiating ψ_λ gives

$$\psi'_\lambda(w) = \frac{\lambda^2 w^2 - 2\lambda w + \lambda}{\lambda^2 w^2 - 2\lambda w + 1} = 1 - \frac{1-\lambda}{(1-\lambda w)^2} \ .$$

Therefore,

$$\phi'_\lambda(w) = 1 - \frac{1-\lambda}{(1-\lambda w)^2} - \left(1 - \frac{1-\lambda}{(1-\lambda(1-w))^2}\right) = \frac{1-\lambda}{(1-\lambda(1-w))^2} - \frac{1-\lambda}{(1-\lambda w)^2} \ .$$

It is straightforward to check that the above expression has only one root at $w = 1/2$, being non-negative for $w \in [0,1/2)$, and non-positive for $w \in (1/2,1]$. Therefore, $\phi_\lambda(1/2) = \frac{\lambda/2}{1-\lambda/2}$ is the maximum value of ϕ_λ. ☐

However, directly applying Lemma 7.7 to $W_1(\hat{\rho}_\oplus, \hat{\rho}'_\oplus) + W_1(\hat{\rho}_\ominus, \hat{\rho}'_\ominus)$ seems hopeless, since in the bounds supplied by Claim 7.6, the parameters of the functions ψ bounding $\mathscr{D}(\cdot, r, +1)$-terms in $W_1\left(\hat{\rho}_\oplus, \hat{\rho}'_\oplus\right)$ are quite different from the parameters of the functions ψ bounding the corresponding $\mathscr{D}(\cdot, r, -1)$-terms in $W_1\left(\hat{\rho}_\ominus, \hat{\rho}'_\ominus\right)$.

The following lemma reveals, a somewhat unexpected, symmetry between $W_1(\hat{\rho}_\oplus, \hat{\rho}'_\oplus)$ and $W_1(\hat{\rho}_\ominus, \hat{\rho}'_\ominus)$, that facilitates our pairing strategy.

Some additional notation is in order. We denote with $\mathscr{R}(k)$ for the set of all vectors $r = (r_\bullet, r_\oplus, r_\ominus, r_\bigcirc)$ of non-negative integer entries which sum to $k-1$. For every $r \in \mathscr{R}(k)$ we use the shorthand

$$P(r) = \frac{(k-1)!}{r_\bullet! r_\oplus! r_\ominus! r_\bigcirc!} \cdot p_\bullet^{r_\bullet} p_\oplus^{r_\oplus} p_\ominus^{r_\ominus} p_\bigcirc^{r_\bigcirc} \ ,$$

where $p_\bullet, p_\oplus, p_\ominus, p_\bigcirc$ are the probabilities defined in (2.18). Finally, we define

$$E_\bullet = \sum_{\substack{r\in\mathscr{R}(k)\\ r_\bullet \geq 1}} P(r) \cdot r_\bullet \cdot \mathbb{E}\left[\int_{\boldsymbol{\eta}^\wedge_{\bullet,1,1}}^{\boldsymbol{\eta}^\vee_{\bullet,1,1}} \phi_{2^{-r_\bigcirc - r_\ominus}}\left(\frac{1+\tanh(w/2)}{2}\right) \mathrm{d}w\right] \ , \tag{7.19}$$

$$E_\oplus = \sum_{\substack{r\in\mathscr{R}(k)\\ r_\oplus \geq 1}} P(r) \cdot r_\oplus \cdot \mathbb{E}\left[\int_{\boldsymbol{\eta}^\wedge_{\oplus,1,1}}^{\boldsymbol{\eta}^\vee_{\oplus,1,1}} \phi_{2^{-r_\bigcirc - r_\ominus}}\left(\frac{1+\tanh(y/2)}{2}\right) \mathrm{d}y\right] \ , \tag{7.20}$$

$$E_\ominus = \sum_{\substack{r\in\mathscr{R}(k)\\ r_\ominus \geq 1}} P(r) \cdot r_\ominus \cdot \mathbb{E}\left[\int_{\boldsymbol{\eta}^\wedge_{\ominus,1,1}}^{\boldsymbol{\eta}^\vee_{\ominus,1,1}} \phi_{2^{-r_\bigcirc - r_\oplus}}\left(\frac{1+\tanh(z/2)}{2}\right) \mathrm{d}z\right] \ . \tag{7.21}$$

Lemma 7.8. *We have that*

$$W_1(\hat\rho_\oplus, \hat\rho'_\oplus) + W_1(\hat\rho_\ominus, \hat\rho'_\ominus) \leq \frac{d/2}{1-e^{-d/2}}\left(E_\bullet + E_\oplus + E_\ominus\right) \ . \tag{7.22}$$

Proof. Expanding the expectation in (7.8) with respect to $\boldsymbol{r} = \left(\boldsymbol{r}_\bullet, \boldsymbol{r}_\oplus, \boldsymbol{r}_\ominus, \boldsymbol{r}_\bigcirc\right)$, and using the shorthand

$$E^\pm_\bullet(r) = \mathbb{E}\left[\int_{\boldsymbol{\eta}^\wedge_{\bullet,1,1}}^{\boldsymbol{\eta}^\vee_{\bullet,1,1}} \mathscr{D}_1^\bullet(w, r; \pm 1)\,\mathrm{d}w\right], \ E^\pm_\oplus(r) = \mathbb{E}\left[\int_{\boldsymbol{\eta}^\wedge_{\oplus,1,1}}^{\boldsymbol{\eta}^\vee_{\oplus,1,1}} \mathscr{D}_1^\oplus\left(y, r; \pm 1\right) \mathrm{d}y\right], \ E^\pm_\ominus(r) = \mathbb{E}\left[\int_{\boldsymbol{\eta}^\wedge_{\ominus,1,1}}^{\boldsymbol{\eta}^\vee_{\ominus,1,1}} \mathscr{D}_1^\ominus(z, r; \pm\mathbf{1})\,\mathrm{d}z\right],$$

we see that

$$\begin{aligned} W_1(\hat\rho_\oplus, \hat\rho'_\oplus) &\leq \frac{d/2}{1-e^{-d/2}}\left(\sum_{r\in\mathscr{R}(k)} P(r)\cdot r_\bullet \cdot E^+_\bullet(r) + \sum_{r\in\mathscr{R}(k)} P(r)\cdot r_\oplus \cdot E^+_\oplus(r) + \sum_{r\in\mathscr{R}(k)} P(r)\cdot r_\ominus \cdot E^+_\ominus(r)\right) \\ &= \frac{d/2}{1-e^{-d/2}}\left(\sum_{\substack{r\in\mathscr{R}(k)\\ r_\bullet\geq 1}} P(r)\cdot r_\bullet \cdot E^+_\bullet(r) + \sum_{\substack{r\in\mathscr{R}(k)\\ r_\oplus\geq 1}} P(r)\cdot r_\oplus \cdot E^+_\oplus(r) + \sum_{\substack{r\in\mathscr{R}(k)\\ r_\ominus\geq 1}} P(r)\cdot r_\ominus \cdot E^+_\ominus(r)\right) \ . \end{aligned} \tag{7.23}$$

In a similar manner, we derive

$$W_1(\hat\rho_\ominus, \hat\rho'_\ominus) \leq \frac{d/2}{1-e^{-d/2}}\left(\sum_{\substack{r\in\mathscr{R}(k)\\ r_\bullet\geq 1}} P(r)\cdot r_\bullet \cdot E^-_\bullet(r) + \sum_{\substack{r\in\mathscr{R}(k)\\ r_\oplus\geq 1}} P(r)\cdot r_\oplus \cdot E^-_\oplus(r) + \sum_{\substack{r\in\mathscr{R}(k)\\ r_\ominus\geq 1}} P(r)\cdot r_\ominus \cdot E^-_\ominus(r)\right) \ . \tag{7.24}$$

Let us now consider the bound on the sum $W_1(\hat\rho_\oplus, \hat\rho'_\oplus) + W_1(\hat\rho_\ominus, \hat\rho'_\ominus)$ obtained by summing (7.23), (7.24). We next group each of the three sums in (7.23) with the corresponding sum in (7.24), carefully pairing their terms. Specifically, for the $\bullet$–sums we match the term of $\sum_r P(r)\cdot r_\bullet \cdot E^+_\bullet(r)$ corresponding to $r = (r_\bullet, r_\oplus, r_\ominus, r_\bigcirc)$ with the term of $\sum_{r'} P(r')\cdot r'_\bullet \cdot E^-_\bullet(r')$ that corresponds to $r' = (r_\bullet, r_\ominus, r_\oplus, r_\bigcirc)$. Since $r \mapsto r'$ is a bijection of $\mathscr{R}(k) \cap \{r : r_\bullet \geq 1\}$, and $r'_\bullet = r_\bullet$, and $P(r') = P(r)$ we see that

$$\sum_{\substack{r\in\mathscr{R}(k)\\ r_\bullet\geq 1}} P(r)\cdot r_\bullet\left(E^+_\bullet(r) + E^-_\bullet(r)\right) = \sum_{\substack{r\in\mathscr{R}(k)\\ r_\bullet\geq 1}} P(r)\cdot r_\bullet\left(E^+_\bullet(r) + E^-_\bullet(r')\right) \ . \tag{7.25}$$

Invoking the bounds (7.14) of Claim 7.6, and recalling the definitions of ϕ, $E_\bullet$, we upper bound the r.h.s. of (7.25) by

$$\sum_{\substack{r\in\mathscr{R}(k)\\ r_\bullet\geq 1}} P(r) r_\bullet \left(\mathbb{E}\left[\int_{\boldsymbol{\eta}^\wedge_{\bullet,1,1}}^{\boldsymbol{\eta}^\vee_{\bullet,1,1}} \psi_{2^{-r_\bigcirc - r_\ominus}}\left(\frac{1+\tanh(w/2)}{2}\right) \mathrm{d}w\right] + \mathbb{E}\left[\int_{\boldsymbol{\eta}^\wedge_{\bullet,1,1}}^{\boldsymbol{\eta}^\vee_{\bullet,1,1}} \psi_{2^{-r_\bigcirc - r_\ominus}}\left(\frac{1-\tanh(w/2)}{2}\right) \mathrm{d}w\right]\right) = E_\bullet \ . \tag{7.26}$$

The matchings between the terms for the $\oplus, \ominus$–sums of (7.23), (7.24) are more delicate. In particular, for the $\oplus$–sum it turns out that we can pull off the same trick as above by pairing the term of $\sum_r P(r)\cdot r_\oplus \cdot E^+_\oplus(r)$ corresponding to the vector $r=(r_\bullet, r_\oplus, r_\ominus, r_\bigcirc)$ with the term of $\sum_{r''} P(r'')\cdot r''_\oplus \cdot E^-_\oplus(r'')$ that corresponds to $r''=(r_\bullet, r_\ominus+1, r_\oplus-1, r_\bigcirc)$. To see this, note that the mapping $r\mapsto r''$ is a bijection of $\mathscr{R}(k)\cap\{r: r_\oplus\geq 1\}$, leaving the quantity $P(r)\cdot r_\oplus$ invariant as

$$P(r)\cdot r_\oplus = \frac{(k-1)!}{r_\bullet!(r_\oplus-1)!r_\ominus!r_\bigcirc!}\cdot p_\bullet^{r_\bullet} p_\oplus^{r_\oplus} p_\ominus^{r_\ominus} p_\bigcirc^{r_\bigcirc} = \frac{(k-1)!}{r_\bullet!(r_\ominus+1)!(r_\oplus-1)!r_\bigcirc!}\cdot p_\bullet^{r_\bullet} p_\oplus^{r_\oplus} p_\ominus^{r_\ominus} p_\bigcirc^{r_\bigcirc}\cdot(r_\ominus+1) = P(r'')\cdot r''_\oplus \ .$$

Invoking the bounds (7.15) of Claim 7.6, recalling the definitions of ϕ, $E_\oplus$, and arguing as above, we obtain

$$\sum_{\substack{r\in\mathscr{R}(k)\\ r_\oplus\geq 1}} P(r)\cdot r_\oplus\left(E^+_\oplus(r)+E^-_\oplus(r'')\right)\leq E_\oplus \ . \tag{7.27}$$

Similarly, using the mapping $r\mapsto r'''$, with $r'''=(r_\bullet, r_\ominus-1, r_\oplus+1, r_\bigcirc)$, and following the same steps as above, we get

$$\sum_{\substack{r\in\mathscr{R}(k)\\ r_\ominus\geq 1}} P(r)\cdot r_\ominus\left(E^+_\ominus(r)+E^-_\ominus(r''')\right)\leq E_\ominus \ . \tag{7.28}$$

Summing (7.26)–(7.28) concludes the proof. □

In light of the above, we are now ready to finish the proof of Proposition 2.15.

Proof of Proposition 2.15. Applying Lemma 7.7 on the function ϕ in the r.h.s. of (7.19) gives

$$\phi_{2^{-r_\bigcirc-r_\ominus}}\left(\frac{1+\tanh(w/2)}{2}\right)\leq\frac{2^{-r_\ominus-r_\bigcirc-1}}{1-2^{-r_\ominus-r_\bigcirc-1}}\leq\left(\frac{1}{2}\right)^{r_\ominus+r_\bigcirc} \ . \tag{7.29}$$

Plugging the above into (7.19) and applying the binomial theorem, further yields

$$E_\bullet\leq(k-1)\cdot p_\bullet\left(1-\frac{e^{-\frac{d}{2}}}{2}\right)^{k-2}\mathbb{E}\left[|\boldsymbol{\eta}_{\bullet,1,1}-\boldsymbol{\eta}'_{\bullet,1,1}|\right] \ . \tag{7.30}$$

Working in a similar manner, we obtain

$$E_\oplus\leq(k-1)\cdot p_\oplus\left(1-\frac{e^{-\frac{d}{2}}}{2}\right)^{k-2}\mathbb{E}\left[|\boldsymbol{\eta}_{\oplus,1,1}-\boldsymbol{\eta}'_{\oplus,1,1}|\right], \quad\text{and}\quad E_\ominus\leq(k-1)\cdot p_\ominus\left(1-\frac{e^{-\frac{d}{2}}}{2}\right)^{k-2}\mathbb{E}\left[|\boldsymbol{\eta}_{\ominus,1,1}-\boldsymbol{\eta}'_{\ominus,1,1}|\right] \ . \tag{7.31}$$

Finally, plugging the bounds (7.30), and (7.31) into (7.22) we see that $W_1(\hat\rho_\oplus,\hat\rho'_\oplus)+W_1(\hat\rho_\ominus,\hat\rho'_\ominus)$ is upper bounded by

$$\frac{d\cdot(k-1)}{2}\cdot\left(1-\frac{e^{-\frac{d}{2}}}{2}\right)^{k-2}\left[\left(1-e^{-\frac{d}{2}}\right)\mathbb{E}\left[|\boldsymbol{\eta}_{\bullet,1,1}-\boldsymbol{\eta}'_{\bullet,1,1}|\right]+e^{-\frac{d}{2}}\mathbb{E}\left[|\boldsymbol{\eta}_{\oplus,1,1}-\boldsymbol{\eta}'_{\oplus,1,1}|\right]+e^{-\frac{d}{2}}\mathbb{E}\left[|\boldsymbol{\eta}_{\ominus,1,1}-\boldsymbol{\eta}'_{\ominus,1,1}|\right]\right] \ . \tag{7.32}$$

Recall that we established (7.32) assuming an arbitrary coupling between the coordinates of each pair of distributions $(\rho_\bullet,\rho'_\bullet)$, $(\rho_\oplus,\rho'_\oplus)$, and $(\rho_\ominus,\rho'_\ominus)$. Therefore, the definition of W_1 norm and (7.32), imply the first inequality below, while (7.33) follows by the definition (2.22) of dist_d

$$\begin{aligned}W_1(\hat\rho_\oplus,\hat\rho'_\oplus)+W_1(\hat\rho_\ominus,\hat\rho'_\ominus)&\leq\frac{d(k-1)}{2}\left(1-\frac{e^{-\frac{d}{2}}}{2}\right)^{k-2}\left[\left(1-e^{-\frac{d}{2}}\right)W_1(\rho_\bullet,\rho'_\bullet)+e^{-\frac{d}{2}}W_1(\rho_\oplus,\rho'_\oplus)+e^{-\frac{d}{2}}W_1(\rho_\ominus,\rho'_\ominus)\right]\\&\leq\frac{d(k-1)}{2}\left(1-\frac{e^{-\frac{d}{2}}}{2}\right)^{k-2}\mathrm{dist}_d(\rho,\rho') \ .\end{aligned}\tag{7.33}$$

Moreover, as per the triangle inequality we see that

$$W_1(\hat\rho_\bullet,\hat\rho'_\bullet)\leq W_1(\hat\rho_\oplus,\hat\rho'_\oplus)+W_1(\hat\rho_\ominus,\hat\rho'_\ominus)\leq\frac{d(k-1)}{2}\left(1-\frac{e^{-\frac{d}{2}}}{2}\right)^{k-2}\mathrm{dist}_d(\rho,\rho') \ . \tag{7.34}$$

Plugging the bounds (7.33) and (7.34) into the expression of $\mathrm{dist}_d\left(\hat\rho,\hat\rho'\right)$ yields

$$\mathrm{dist}_d\left(\hat\rho,\hat\rho'\right)=(1-e^{-d/2})\cdot W_1(\hat\rho_\bullet,\hat\rho'_\bullet)+e^{-d/2}\left(W_1(\hat\rho_\oplus,\hat\rho'_\oplus)+W_1(\hat\rho_\ominus,\hat\rho'_\ominus)\right)\leq\frac{d(k-1)}{2}\left(1-\frac{e^{-\frac{d}{2}}}{2}\right)^{k-2}\mathrm{dist}_d(\rho,\rho') \ .$$

Recalling the definition of d_{con}, we see that for $d<d_{\mathrm{con}}(k)$, the operator $\mathrm{LL}^\star_{d,k}$ contracts with respect to the metric dist_d, as desired. □

7.4. **Proof of Proposition 2.13.** To get a handle on the $\boldsymbol{\eta}_x^{(\ell)}$ from (2.12), we show that these quantities can be calculated by propagating the extremal boundary condition $\boldsymbol{\sigma}^+$ bottom-up toward the root of the tree. Specifically, we consider the operator

$$\Lambda^+_{\mathbb{T}^{(\ell)}} : (-\infty,\infty]^{V(\mathbb{T}^{(\ell)})} \to (-\infty,\infty]^{V(\mathbb{T}^{(\ell)})} \ , \qquad\qquad \eta \mapsto \hat{\eta} = \Lambda^+_{\mathbb{T}^{(\ell)}}(\eta) \ ,$$

defined as follows. For all $x \in \partial^{2\ell}\mathfrak{x}$ we set $\hat{\eta}_x = \infty$. Moreover, for a variable $x \in \partial^{2q}\mathfrak{x}$ with $q < \ell$ having children clauses $a_1,\dots,a_t$, and grandchildren variables $y_{1,1},\dots,y_{1,(k-1)},\dots,y_{t,1},\dots,y_{t,(k-1)}$ we define

$$\hat{\eta}_x = -\sum_{i=1}^{t} \boldsymbol{\tau}^+(x)\mathrm{sign}(x,a_i)\cdot\log\left(1-\Gamma\left(\boldsymbol{\tau}^+(x)\mathrm{sign}(x,a_i)\cdot(\eta_{y_{i,1}},\dots,\eta_{y_{1,(k-1)}})\right)\right) \ . \tag{7.35}$$

It may not be apparent that the sum above is well-defined as $-\infty$ summands may manifest. The following lemma rules out such possibility and shows that the ℓ-fold iteration of $\Lambda^{+(\ell)}_{\mathbb{T}^{(\ell)}}$, initiated all-$(+\infty)$ yields $\boldsymbol{\eta}^{(\ell)} = (\boldsymbol{\eta}^{(\ell)}_x)_{x\in V(\mathbb{T}^{(\ell)})}$.

Lemma 7.9. *The operator $\Lambda^+_{\mathbb{T}^{(\ell)}}$ is well-defined and $\Lambda^{+(t)}_{\mathbb{T}^{(\ell)}}(+\infty,\dots,+\infty) = \boldsymbol{\eta}^{(\ell)}$ for every $t \ge \ell$.*

Proof. To show that $\Lambda^+_{\mathbb{T}^{(\ell)}}$ is well defined we verify that, in the notation of (7.35), $\hat{\eta}_x \in (-\infty,\infty]$ for all x. Indeed, in the expression on the r.h.s. of (7.35) a $\pm\infty$ summand can arise only from variables $y_{i,j}$ with $\eta_{y_{i,j}} = \infty$. But the definition of $\boldsymbol{\tau}^+$ ensures that such $y_{i,j}$ either render a zero summand if $\boldsymbol{\tau}^+(x)\mathrm{sign}(x,a_i) = -1$, or a $+\infty$ summand if $\boldsymbol{\tau}^+(x)\mathrm{sign}(x,a_i) = 1$. Thus, the sum is well-defined and $\hat{\eta}_x \in (-\infty,\infty]$.

Further, to verify the identity $\boldsymbol{\eta}^{(\ell)} = \Lambda^{+(\ell)}_{\mathbb{T}^{(\ell)}}(\infty,\dots,\infty)$, consider a variable x of $\mathbb{T}^{(\ell)}$. Let $a_1^+,\dots,a_g^+$ be the children (clauses) of x with $\mathrm{sign}(x,a_i^+) = \boldsymbol{\tau}^+(x)$. Also let $y_{11},\dots,y_{1(k-1)},\dots,y_{g1},\dots,y_{g(k-1)}$ be the children of $a_1^+,\dots,a_g^+$. Similarly, let $a_1^-,\dots,a_h^-$ be the children of x with $\mathrm{sign}(x,a_i^-) = -\boldsymbol{\tau}^+(x)$ and let $z_{11},\dots,z_{1(k-1)},\dots,z_{h1},\dots,z_{h(k-1)}$ be their children. Then (7.2), and (7.3) yield

$$\begin{aligned}\boldsymbol{\eta}_x^{(\ell)} &= -\sum_{i=1}^{g}\log\left(1-\prod_{q=1}^{k-1}\frac{Z(\mathbb{T}^{(\ell)}_{y_{iq}},\boldsymbol{\tau}^+,\boldsymbol{\tau}^+(y_{iq}))}{Z(\mathbb{T}^{(\ell)}_{y_{iq}},\boldsymbol{\tau}^+)}\right)+\sum_{j=1}^{h}\log\left(1-\prod_{q=1}^{k-1}\frac{Z(\mathbb{T}^{(\ell)}_{z_{jq}},\boldsymbol{\tau}^+,-\boldsymbol{\tau}^+(z_{jq}))}{Z(\mathbb{T}^{(\ell)}_{z_{jq}},\boldsymbol{\tau}^+)}\right)\\ &= -\sum_{i=1}^{g}\log\left(1-\Gamma\left(\mathrm{sign}(x,a_i^+)\boldsymbol{\tau}^+(x)\cdot\left(\boldsymbol{\eta}^{(\ell)}_{y_{i1}},\dots,\boldsymbol{\eta}^{(\ell)}_{y_{i(k-1)}}\right)\right)\right)+\sum_{j=1}^{h}\log\left(1-\Gamma\left(\mathrm{sign}(x,a_i^-)\boldsymbol{\tau}^+(x)\cdot\left(\boldsymbol{\eta}^{(\ell)}_{z_{j1}},\dots,\boldsymbol{\eta}^{(\ell)}_{z_{j(k-1)}}\right)\right)\right) \ .\end{aligned}$$

The assertion follows because $\mathrm{sign}(x,a_i^+)\boldsymbol{\tau}^+(x) = 1$ and $\mathrm{sign}(x,a_i^-)\boldsymbol{\tau}^+(x) = -1$. □

The next aim is to approximate the ℓ-fold iteration of $\Lambda^+_{\mathbb{T}^{(\ell)}}$, and more specifically the distribution of $\boldsymbol{\eta}^{(\ell)}_{\mathfrak{x}}$, using a non-random operator. To this end, we need to cope with the $\pm\infty$-entries of the vector $\boldsymbol{\eta}^{(\ell)}$. This is addressed by Lemma 2.14, proven in Section 7.2, which provides a bound on $\boldsymbol{\eta}^{(\ell)}_x$ for variables x near the root of the tree.

In the following we continue to write c and $(\varepsilon_t)_t$ for the number and the sequence supplied by Lemma 2.14. Guided by Lemma 2.14 we consider the vector $\boldsymbol{\eta}^{(\ell)}_{\wedge t}$ of truncated log-likelihood ratios

$$\left(\boldsymbol{\eta}^{(\ell)}_{\wedge t}\right)_x = \begin{cases} -2t^c & \text{if } x \in \partial^{2t}\mathfrak{x} \text{ and } \boldsymbol{\eta}_x^{(\ell)} < -2t^c \ , \\ 2t^c & \text{if } x \in \partial^{2t}\mathfrak{x} \text{ and } \boldsymbol{\eta}_x^{(\ell)} > 2t^c \ , \\ \boldsymbol{\eta}_x^{(\ell)} & \text{otherwise} \ . \end{cases} \tag{7.36}$$

Further, let $\boldsymbol{\eta}^{(\ell,t)}$ be the result of t iterations of $\Lambda^+_{\mathbb{T}^{(\ell)}}(\cdot)$ starting from $\boldsymbol{\eta}^{(\ell)}_{\wedge t}$. The following corollary is a direct consequence of Lemma 7.9 and Lemma 2.14.

Corollary 7.10. *For any $\ell > ct^c$ we have $\mathbb{P}[\boldsymbol{\eta}^{(\ell,t)}_{\mathfrak{x}} \neq \boldsymbol{\eta}^{(\ell)}_{\mathfrak{x}}] < \varepsilon_t$.*

Proof. Due to Lemma 2.14, the truncation in (7.36) is inconsequential with probability at least $1-\varepsilon_t$, in which case

$$\boldsymbol{\eta}^{(\ell,t)} = \Lambda^{+(t)}_{\mathbb{T}^{(\ell)}}\left(\boldsymbol{\eta}^{(\ell)}_{\wedge t}\right) = \Lambda^{+(t)}_{\mathbb{T}^{(\ell)}}(\boldsymbol{\eta}^{(\ell)}) = \Lambda^{+(\ell+t)}_{\mathbb{T}^{(\ell)}}(+\infty,\dots,+\infty) = \boldsymbol{\eta}^{(\ell)} \ ,$$

where the last equality follows from Lemma 7.9. □

Recall that we defined the non-random operator $\mathrm{LL}^\star_{d,k}$ from (2.17), mimicking $\Lambda^+_{\mathbb{T}^{(\ell)}}$. To make the connection between the random operator $\Lambda^+_{\mathbb{T}^{(\ell)}}$ and $\mathrm{LL}^\star_{d,k}$ precise, we introduce the following concepts. Given a tree formula T we write $V_\bullet(T)$, for the set of x variables of T that appear both as positive and negative literals in the sub-tree T_x comprising x and its the progeny. We define $V_\circ(T)$, $V_\oplus(T)$, and $V_\ominus(T)$ similarly. Note that the above sets constitute a partition of $V(T)$. We use $\mathrm{tp}: V(T) \to \{\bullet, \oplus, \ominus, \circ\}$ to indicate the part each vertex belongs to. We denote with $\mathbb{T}^{(\ell)}_{\bullet}$ the random Galton-Watson formula $\mathbb{T}$ conditioned on the root satisfying $\mathrm{tp}(\mathfrak{r}) = \bullet$. We define $\mathbb{T}^{(\ell)}_{\oplus}$, and $\mathbb{T}^{(\ell)}_{\ominus}$ analogously. Degenerately, we also write $\mathbb{T}^{(\ell)}_{\circ}$ for the formula comprised by a single variable $\mathfrak{r}$. Let us denote with $\hat{\eta}^{(\ell,t)}_{\bullet}$ the distribution of $(\boldsymbol{\eta}^{(\ell)}_{\wedge t})_{\mathfrak{r}}$ in $\mathbb{T}^{(\ell)}_{\bullet}$. Moreover, let $\bar{\eta}^{(\ell-t)}_{\bullet}$ be the distribution of

$$\boldsymbol{\eta}^{(\ell-t)}_{\mathfrak{r}} \cdot \mathbb{1}\left\{|\boldsymbol{\eta}^{(\ell-t)}_{\mathfrak{r}}| \le 2t^c\right\} + 2t^c \cdot \mathbb{1}\left\{\boldsymbol{\eta}^{(\ell-t)}_{\mathfrak{r}} > 2t^c\right\} - 2t^c \cdot \mathbb{1}\left\{\boldsymbol{\eta}^{(\ell-t)}_{\mathfrak{r}} < -2t^c\right\} \ ,$$

i.e., the truncation of $\boldsymbol{\eta}^{(\ell-t)}_{\mathfrak{r}}$ in $\mathbb{T}^{(\ell)}_{\bullet}$. Analogously we define the distributions $\hat{\eta}^{(\ell,t)}_{\oplus}, \hat{\eta}^{(\ell,t)}_{\ominus}$, and $\bar{\eta}^{(\ell-t)}_{\oplus}, \bar{\eta}^{(\ell-t)}_{\ominus}$. Notice that, degenerately, $\hat{\eta}^{(\ell,t)}_{\circ} = \bar{\eta}^{(\ell-t)}_{\circ} = \delta_0$.

Lemma 7.11. *For $\ell > ct^c$ we have that* $\left(\hat{\eta}^{(\ell,t)}_{\bullet}, \hat{\eta}^{(\ell,t)}_{\oplus}, \hat{\eta}^{(\ell,t)}_{\ominus}\right) = \mathrm{LL}^{\star}_{d,k}\left(\bar{\eta}^{(\ell-t)}_{\bullet}, \bar{\eta}^{(\ell-t)}_{\oplus}, \bar{\eta}^{(\ell-t)}_{\ominus}\right)$.

Proof. We use induction on t. Specifically, let $\nu = \left(\nu_{\bullet}, \nu_{\oplus}, \nu_{\ominus}\right)$ be any triplet in $\mathscr{P}(-\infty,\infty] \times \mathscr{P}[0,+\infty] \times \mathscr{P}[-\infty,0]$, and $\nu^{(t)} = \mathrm{LL}^{\star(t)}_{d,k}(\nu)$ be the outcome of the t-fold application of $\mathrm{LL}^{\star}_{d,k}$. Moreover, let $(\boldsymbol{\eta}_x)_{x \in V(\mathbb{T}^{(t)})}$ be a vector of independent samples with $\boldsymbol{\eta}_x \sim \nu_{\mathrm{tp}(x)}$. We claim that root value $\boldsymbol{\eta}^{(t)}_{\mathfrak{r}}$ of the random operator $\Lambda^{+(t)}_{\mathbb{T}^{(t)}}$, coincides with $\nu_{\mathrm{tp}(\mathfrak{r})}$. Indeed, for $t = 1$ the claim follows readily from the definitions. For the inductive step, we notice that the t-fold application of $\mathrm{LL}^{\star}_{d,k}$ is obtained by applying $\mathrm{LL}^{\star}_{d,k}$ to the $(t-1)$-fold application. Per the induction hypothesis

$$\left(\Lambda^{+(t-1)}_{\mathbb{T}^{(t-1)}}(\boldsymbol{\eta}_x)_x\right)_{\mathfrak{r}} \sim \nu^{(t-1)}_{\mathrm{tp}(\mathfrak{r})} \ . \tag{7.37}$$

Applying $\mathrm{LL}^{\star}_{d,k}$ to $\nu^{(t-1)}$ implies the result as the first layer of $\mathbb{T}^{(t)}$ is independent of the subtrees rooted at the grandchildren $\partial^2\mathfrak{r}$ of the root, which are distributed as i.i.d. copies of $\mathbb{T}^{(t-1)}$. The lemma follows from applying the above identity to $\nu = \left(\bar{\eta}^{(\ell-t)}_{\bullet}, \bar{\eta}^{(\ell-t)}_{\oplus}, \bar{\eta}^{(\ell-t)}_{\ominus}\right)$. □

Refining the definition of the $\mathrm{BP}_{d,k}$ operator in (1.3), we write $\mathrm{BP}^{\bullet}_{d,k}$ for the operator obtained from $\mathrm{BP}_{d,k}$ upon conditioning on $\boldsymbol{d}^+, \boldsymbol{d}^- \ge 1$. Similarly $\mathrm{BP}^{\oplus}_{d,k}$ and $\mathrm{BP}^{\ominus}_{d,k}$ are obtained from $\mathrm{BP}_{d,k}$ upon conditioning on $\boldsymbol{d}^+ \ge 1, \boldsymbol{d}^- = 0$, and $\boldsymbol{d}^+ = 0, \boldsymbol{d}^- \ge 1$, respectively. We define

$$\pi^{\bullet}_{d,k} = \mathrm{BP}^{\bullet}_{d,k}\left(\pi_{d,k}\right), \quad \pi^{\oplus}_{d,k} = \mathrm{BP}^{\oplus}_{d,k}\left(\pi_{d,k}\right), \quad \pi^{\ominus}_{d,k} = \mathrm{BP}^{\ominus}_{d,k}\left(\pi_{d,k}\right) \ .$$

Let us write γ, γ^{-1} for the continuous and mutually inverse real functions

$$\gamma : \mathbb{R} \to (0,1), \ \ z \mapsto (1 + \tanh(z/2))/2, \qquad \gamma^{-1} : (0,1) \to \mathbb{R}, \ \ p \mapsto \log(p/(1-p)) \ . \tag{7.38}$$

Let $\rho^{\bullet}_{d,k} = \gamma^{-1}(\pi^{\bullet}_{d,k})$, and define $\rho^{\oplus}_{d,k}, \rho^{\ominus}_{d,k}$ similarly.

Claim 7.12. *The vector* $\left(\rho^{\bullet}_{d,k}, \rho^{\oplus}_{d,k}, \rho^{\ominus}_{d,k}\right)$ *is a fixed point of the operator* $\mathrm{LL}^{\star}_{d,k}$.

Proof. Let $\rho_{d,k} = \gamma^{-1}\left(\pi_{d,k}\right)$. First, we claim that

$$\mathrm{LL}^{\star}_{d,k}\left(\rho_{d,k}, \rho_{d,k}, \rho_{d,k}\right) = \left(\rho^{\bullet}_{d,k}, \rho^{\oplus}_{d,k}, \rho^{\ominus}_{d,k}\right) \ . \tag{7.39}$$

Indeed, since all input distributions are the same, by Proposition 2.1, the two summands in the left term of (2.21) corresponding to $\boldsymbol{d}^{\star}_{+}$ and $\boldsymbol{d}^{\star}_{-}$ are identically distributed, and also identically distributed to the sums that appear in the other two terms. Therefore, (7.39) follows directly from the definitions of $\mathrm{BP}^{\bullet}_{d,k}, \mathrm{BP}^{\oplus}_{d,k}$, and $\mathrm{BP}^{\ominus}_{d,k}$. The claim now follows from (7.39), the definition of the operator $\mathrm{LL}^{\star}_{d,k}$, and the law of total probability. □

Let $\rho^{(\ell)}$ be the distribution of the log-likelihood ratio $\boldsymbol{\eta}^{(\ell)}_{\mathfrak{r}}$.

Corollary 7.13. *For $d < d_{\mathrm{con}}(k)$ the sequence $\left(\gamma\left(\rho^{(\ell)}\right)\right)_{\ell}$ converges weakly to $\pi_{d,k}$.*

Proof. The result follows by combining Corollary 7.10, Lemma 7.11, Proposition 2.15, Claim 7.12, and applying the continuous mapping theorem and the law of total probability. □

Proof of Proposition 2.13. Recall that we write $\Lambda^{+(\ell)}_{\mathbb{T}^{(\ell)}}$ for the ℓ-fold iteration of the operator $\Lambda^{+}_{\mathbb{T}}$. Let us write $\boldsymbol{\theta}^{(\ell)}_{\mathfrak{r}} = \left(\Lambda^{+(\ell)}_{\mathbb{T}^{(\ell)}}(0,\ldots,0)\right)_{\mathfrak{r}}$. Using arguments similar to Fact 4.2, we can show that $\boldsymbol{\theta}^{(\ell)}_{\mathfrak{r}}$ is nothing but the distribution of the random variable $\gamma^{-1}(\mathbb{P}\left[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \mathbb{T}\right])$. Therefore,

$$\mathbb{P}[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \mathbb{T}] \sim \gamma(\boldsymbol{\theta}^{(\ell)}_{\mathfrak{r}}) \ , \qquad \text{and} \qquad \mathbb{P}[\boldsymbol{\tau}^{(\ell)}(\mathfrak{r}) = 1 \mid \forall y \in \partial^{2\ell}\mathfrak{r} : \boldsymbol{\tau}^{(\ell)}(y) = \boldsymbol{\tau}^{+}(y), \mathbb{T}] \sim \gamma(\boldsymbol{\eta}^{(\ell)}_{\mathfrak{r}}) \ .$$

Due to Lemma 2.12, $0 \le \gamma(\boldsymbol{\theta}_{\mathfrak{x}}^{(\ell)}) \le \gamma(\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)}) \le 1$. Moreover, from Lemma 7.11, Proposition 2.15, and Claim 7.12, we see that for $d < d_{\mathrm{con}}(k)$ the sequence $\big(\gamma(\boldsymbol{\theta}_{\mathfrak{x}}^{(\ell)})\big)_\ell$ converges weakly to $\pi_{d,k}$. Finally, Corollary 7.13 implies that $\big(\gamma(\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)})\big)_\ell$ also converges weakly to $\pi_{d,k}$, and thus,

$$\lim_{\ell\to\infty} \mathbb{E}\left|\gamma(\boldsymbol{\theta}_{\mathfrak{x}}^{(\ell)}) - \gamma(\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)})\right| = \lim_{\ell\to\infty}\left|\mathbb{E}\left[\gamma(\boldsymbol{\theta}_{\mathfrak{x}}^{(\ell)})\right] - \mathbb{E}\left[\gamma(\boldsymbol{\eta}_{\mathfrak{x}}^{(\ell)})\right]\right| = 0 \ ,$$

implying the assertion. □

Acknowledgements

We would like to thank the anonymous referees for thoroughly reviewing our paper and for suggesting valuable corrections and improvements.

Amin Coja-Oghlan is supported by DFG CO 646/3, DFG CO 646/5 and DFG CO 646/6. Catherine Greenhill is supported by ARC DP250101611. Vincent Pfenninger is supported by the Austrian Science Fund (FWF) [10.55776 / 16502]. Pavel Zakharov is supported by DFG CO 646/6. Kostas Zampetakis is supported by DFG CO 646/5. For open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[1] E. Abbe, A. Montanari: On the concentration of the number of solutions of random satisfiability formulas. Random Structures and Algorithms **45** (2014) 362–382.

[2] D. Achlioptas, A. Coja-Oghlan: Algorithmic barriers from phase transitions. Proc. 49th FOCS (2008) 793–802.

[3] D. Achlioptas, A. Coja-Oghlan, M. Hahn-Klimroth, J. Lee, N. Müller, M. Penschuck, G. Zhou: The number of satisfying assignments of random 2-SAT formulas. Random Structures and Algorithms **58** (2021) 609–647.

[4] D. Achlioptas, A. Coja-Oghlan, F. Ricci-Tersenghi: On the solution-space geometry of random constraint satisfaction problems. Random Structures and Algorithms **38** (2011) 251–268.

[5] D. Achlioptas, C. Moore: Random k-SAT: two moments suffice to cross a sharp threshold. SIAM Journal on Computing **36** (2006) 740–762.

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

[7] D. Achlioptas, Y. Peres: The threshold for random k-SAT is $2^k \ln 2 - O(k)$. Journal of the AMS **17** (2004) 947–973.

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

[9] D. Aldous, J. Steele: The objective method: probabilistic combinatorial optimization and local weak convergence. In: H. Kesten (ed.): Probability on Discrete Structures. Springer (2004).

[10] N. Alon, J. Spencer: The probabilistic method. Wiley (2016).

[11] V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann, D. Vilenchik: The condensation phase transition in random graph coloring. Communications in Mathematical Physics **341** (2016) 543–606.

[12] M. Bayati, D. Gamarnik, P. Tetali: Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. Ann. Probab. **41** (2013) 4080–4115.

[13] R. Biswas, W. Chen, A. Sen: On the replica symmetric solution in general diluted spin glasses. arXiv:2410.15599 (2024).

[14] S. Boucheron, G. Lugosi, P. Massart: Concentration Inequalities: A Nonasymptotic Theory of Independence. OUP Oxford (2013).

[15] G. Bresler, B. Huang: The algorithmic phase transition of random k-sat for low degree polynomials. Proc. 62th FOCS (2021) 298–309.

[16] A. Broder, A. Frieze, E. Upfal: On the satisfiability and maximum satisfiability of random 3-CNF formulas. Proc. 4th SODA (1993) 322–330.

[17] A. Chatterjee, A. Coja-Oghlan, N. Müller, C. Riddlesden, M. Rolvien, P. Zakharov, H. Zhu: The number of random 2-SAT solutions is asymptotically log-normal. Proc. 28th RANDOM (2024) #39.

[18] P. Cheeseman, B. Kanefsky, W. Taylor: Where the *really* hard problems are. Proc. IJCAI (1991) 331–337.

[19] Z. Chen, A. Galanis, L. A. Goldberg, H. Guo, A. Herrera-Poyatos, N. Mani, A. Moitra: Fast sampling of satisfying assignments from random k-SAT with applications to connectivity. SIAM J. Disc. Math. **38** (2024) 2750–2811.

[20] Z. Chen, A. Lonkar, C. Wang, K. Yang, Y. Yin: Counting random k-SAT near the satisfiability threshold. arXiv 2411.02980v1 (2024).

[21] V. Chvátal, B. Reed: Mick gets some (the odds are on his side). Proc. 33th FOCS (1992) 620–627.

[22] A. Coja-Oghlan: A better algorithm for random k-SAT. SIAM Journal on Computing **39** (2010) 2823–2864.

[23] A. Coja-Oghlan: Belief Propagation fails on random formulas. Journal of the ACM **63** (2017) #49.

[24] A. Coja-Oghlan, T. Kapetanopoulos, N. Müller: The replica symmetric phase of random constraint satisfaction problems. Combinatorics, Probability and Computing **29** (2020) 346-422.

[25] A. Coja-Oghlan, F. Krzakala, W. Perkins, L. Zdeborová: Information-theoretic thresholds from the cavity method. Advances in Mathematics **333** (2018) 694–795.

[26] A. Coja-Oghlan, N. Müller, J. Ravelomanana: Belief Propagation on the random k-SAT model. Annals of Applied Probability **32** (2022) 3718–3796.

[27] A. Coja-Oghlan, K. Panagiotou: The asymptotic k-SAT threshold. Advances in Mathematics **288** (2016) 985–1068.

[28] A. Coja-Oghlan, W. Perkins: Belief Propagation on replica symmetric random factor graph models. Annales de l'institut Henri Poincare D **5** (2018) 211–249.

[29] A. Coja-Oghlan, N. Wormald: The number of satisfying assignments of random regular k-SAT formulas. Combinatorics, Probability and Computing **27** (2018) 496–530.

[30] A. Dembo, A. Montanari: Gibbs measures and phase transitions on sparse random graphs. Brazilian Journal of Probability and Statistics **24** (2010) 137–211.
[31] A. Dembo, A. Montanari: Ising models on locally tree-like graphs. Annals of Applied Probability **20** (2010) 565–592.
[32] A. Dembo, A. Montanari, N. Sun: Factor models on locally tree-like graphs. Annals of Probability **41** (2013) 4162–4213.
[33] J. Ding, A. Sly, N. Sun: Proof of the satisfiability conjecture for large k. **20** Annals of Mathematics **196** (2022) 1–388.
[34] O. Dubois, J. Mandler: The 3-XORSAT threshold. Proc. 43rd FOCS (2002) 769–778.
[35] S. Franz, M. Leone: Replica bounds for optimization problems and diluted spin systems. J. Stat. Phys. **111** (2003) 535–564.
[36] E. Friedgut: Sharp thresholds of graph properties, and the k-SAT problem. Journal of the AMS **12** (1999) 1017–1054.
[37] A. Frieze, S. Suen: Analysis of two simple heuristics on a random instance of k-SAT. Journal of Algorithms **20** (1996) 312–355.
[38] A. Galanis, L. A. Goldberg, H. Guo, K. Yang. Counting solutions to random SAT formulas. SIAM J. Comput. **50** (2021) 1701–1738.
[39] H.-O. Georgii: Gibbs measures and phase transitions. De Gruyter (1988).
[40] A. Goerdt: A threshold for unsatisfiability. J. Comput. Syst. Sci. **53** (1996) 469–486
[41] F. Guerra: Broken replica symmetry bounds in the mean field spin glass model. Comm. Math. Phys. **233** (2003) 1–12.
[42] M. Hajiaghayi, G. Sorkin: The satisfiability threshold of random 3-SAT is at least 3.52. IBM Research Report RC22942 (2003).
[43] K. He, K. Wu, K. Yang: Improved bounds for sampling solutions of random SAT formulas. Proc. 34th SODA (2023) 3330–3361.
[44] S. Hetterich: Analysing Survey Propagation Guided Decimationon Random Formulas. Proc. 43rd ICALP (2016) #65.
[45] S. Janson, T. Luczak, A. Ruciński: Random Graphs. Wiley (2000).
[46] A. Kaporis, L. Kirousis, E. Lalas: The probabilistic analysis of a greedy satisfiability algorithm. Random Structures and Algorithms **28** (2006) 444–480.
[47] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborová: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. National Academy of Sciences **104** (2007) 10318–10323.
[48] L. Lovász: Large networks and graph limits. AMS (2012).
[49] S. Mertens, M. Mézard, Riccardo Zecchina: Threshold values of random K-SAT from the cavity method. Random Structures and Algorithms **28** (2006) 340–373.
[50] M. Mézard, A. Montanari: Information, physics and computation. Oxford University Press (2009).
[51] M. Mézard, G. Parisi, R. Zecchina: Analytic and algorithmic solution of random satisfiability problems. Science **297** (2002) 812–815.
[52] A. Moitra. Approximate counting, the Lovasz local lemma, and inference in graphical models. J. ACM **66** #10 (2019).
[53] M. Molloy: Cores in random hypergraphs and Boolean formulas. Random Structures and Algorithms **27** (2005) 124–135.
[54] R. Monasson, R. Zecchina: The entropy of the k-satisfiability problem. Phys. Rev. Lett. **76** (1996) 3881.
[55] R. Monasson, R. Zecchina: Statistical mechanics of the random K-SAT model. Phys. Rev. E **56** (1997) 1357–1370.
[56] A. Montanari, D. Shah: Counting good truth assignments of random k-SAT formulae. Proc. 18th SODA (2007) 1255–1264.
[57] D. Panchenko: The Sherrington-Kirkpatrick model. Springer (2013).
[58] D. Panchenko: Spin glass models from the point of view of spin distributions. Annals of Probability **41** (2013) 1315–1361.
[59] D. Panchenko: On the replica symmetric solution of the K-sat model. Electron. J. Probab. **19** (2014) #67.
[60] D. Panchenko, M. Talagrand: Bounds for diluted mean-fields spin glass models. Probab. Theory Relat. Fields **130** (2004) 319–336.
[61] A. Sly: Computational transition at the uniqueness threshold. Proc. 51st FOCS (2010) 287–296.
[62] M. Talagrand: The high temperature case for the random K-sat problem. Probab. Theory Related Fields **119** (2001) 187–212.
[63] L. Valiant: The complexity of enumeration and reliability problems. SIAM Journal on Computing **8** (1979) 410–421.
[64] C. Wang Y. Yin: A sampling Lovasz local lemma for large domain sizes. Proc. 65th FOCS (2024) 129–150.
[65] A. Coja-Oghlan, W. Perkins: Bethe states of random factor graphs. Communications in Mathematical Physics 366 (2019) 173–201.

ARNAB CHATTERJEE, `arnab.chatterjee@tu-dortmund.de`, TU DORTMUND, FACULTY OF COMPUTER SCIENCE, 12 OTTO-HAHN-ST, DORTMUND 44227, GERMANY.

AMIN COJA-OGHLAN, `amin.coja-oghlan@tu-dortmund.de`, TU DORTMUND, FACULTY OF COMPUTER SCIENCE AND FACULTY OF MATHEMATICS, 12 OTTO-HAHN-ST, DORTMUND 44227, GERMANY.

CATHERINE GREENHILL, `c.greenhill@unsw.edu.au`, SCHOOL OF MATHEMATICS AND STATISTICS, UNSW SYDNEY, NSW 2052, AUSTRALIA.

VINCENT PFENNINGER, `pfenninger@math.tu-graz.at`,TU GRAZ, INSTITUTE OF DISCRETE MATHEMATICS, STEYRERGASSE 30, 8010 GRAZ, AUSTRIA.

MAURICE ROLVIEN, `maurice.rolvien@uni-hamburg.de`, UNIVERSITY OF HAMBURG, FACULTY OF MATHEMATICS, INFORMATICS AND NATURAL SCIENCES, DEPARTMENT OF INFORMATICS, VOGT-KÖLLN-STR. 30, 22527 HAMBURG, GERMANY.

PAVEL ZAKHAROV, `pavel.zakharov@tu-dortmund.de`,TU DORTMUND, FACULTY OF COMPUTER SCIENCE AND FACULTY OF MATHEMATICS, 12 OTTO-HAHN-ST, DORTMUND 44227, GERMANY.

KOSTAS ZAMPETAKIS, `konstantinos.zampetakis@tu-dortmund.de`,TU DORTMUND, FACULTY OF COMPUTER SCIENCE, 12 OTTO-HAHN-ST, DORTMUND 44227, GERMANY.